

## Opening the black box of evaluation: How quality is recognized by peer review panels

Michèle Lamont, Department of Sociology and African and African-American Studies, Harvard University  
Katri Huutoniemi, Department of Social Research, University of Helsinki

*In two recent contributions to the study of peer review we went beyond stating the obvious that peer review produces valid judgments. We draw on in-depth analyses of five fellowship competitions in the United States, and of four grant panels organized by the Academy of Finland. We analyze and compare the intersubjective understandings academic experts create and maintain in making collective judgments on research quality. More specifically, we analyze the social conditions that lead panelists to an understanding of their choices as fair and legitimate, and to a belief that they are able to identify the best and less good proposals.*

Our studies contest the common notion that one can separate cognitive from non-cognitive aspects of evaluation, as we describe the evaluative process as deeply interactional, emotional, and cognitive, and as mobilizing the self-concept of evaluators as much as their expertise. Studies of the internal functioning of peer review have revealed various “intrinsic biases” in peer review like “cognitive particularism” (Travis and Collins 1991), “favoritism for the familiar” (Porter and Rossini 1985), or “peer bias” (Chubin and Hackett 1990; Fuller 2002). These effects show that peer review is not a socially disembodied quality-assessing process in which a set of objective criteria is applied consistently by various reviewers. In fact, the particular cognitive and professional lenses through which evaluators understand proposals necessarily shape evaluation. It is in this context that the informal rules peer reviewers follow become important, as are the lenses through which they understand proposals and the emotions they invest in particular topics and research styles. Thus, instead of contrasting “biased” and “unbiased” evaluation, we aim to capture how evaluation unfolds, as it is carried off and understood by emotional, cognitive, and social beings who necessarily interact with the world through specific frames, narratives, and conventions, but who nevertheless develop expert views concerning what de-

fines legitimate and illegitimate assessments, as well as excellent and less stellar research.

In the two studies, we interviewed academic professionals serving on peer review panels that evaluate fellowship or grant proposals. During the interviews, panelists were asked to describe the arguments they made about a range of proposals, to contrast their arguments with those of other panelists, and to explain what happened in each case. Throughout the interviews, we asked panelists to put themselves in the role of privileged insider and to explain to us how “it” works. They were encouraged to take on the role of the native describing to the observer the rules of the universe in which they operate. We also had access to the preliminary evaluations produced before panel deliberations by individual panelists and to the list of awards given.

### Pragmatic fairness is produced by informal rules

“How Professors Think” came out more than a year ago and has been debated within various academic communities, as it takes on several aspects of the evaluation in multidisciplinary panels in the social sciences and humanities. It is based on an analysis of twelve funding panels organized by important national funding competitions: those of the Social Science Research Council, the American Council for Learned Societies, the Woodrow Wilson Fellowship Foundation, a Society of Fellow at an Ivy League university and an important social science foundation in the social sciences. It draws on 81 interviews with panelists and program officers, as well as on observation of three panels.

A first substantive chapter describes how panels are organized. A second one concerns the evaluative culture of various disciplines, ranging from philosophy to literary studies, history, political science, and economics. A third chapter considers how multidisciplinary panels reach consensus despite variations in disciplinary evaluative cultures. This is followed by two chapters that focus on criteria of evaluation. One analyzes the formal criteria of evaluation provided by the funding organization to panelists (originality, significance, feasibility, etc.) as well as informal criteria (elegance, display of cultural capital, fit between theory and data, etc.). The following chapter considers how cognitive criteria are meshed with extra-cognitive ones (having to do with diversity and interdisciplinarity). We discover that institutional and disciplinary diversity loom much larger than gender and racial diversity in decision making. A concluding chapter considers the implications of the study of evaluation cultures across national contexts, including in Europe.

48 The book is concerned not only with disciplinary compromise, but also with the pragmatic rules that panelists say they abide by, which lead them to believe that the process is fair (this belief is shared by the vast majority of academics interviewed). "How Professors Think" details a range of rules, which include for instance the notion that one should defer to expertise, and that methodological pluralism should be respected.

#### Rules vary across evaluation settings

In her forthcoming article with Huutoniemi, Lamont explores whether these customary rules apply across contexts, and how they vary with how panels are set up. Thus, "Comparing Customary Rules of Fairness" is based on a dialogue between "How Professors Think" and a parallel study conducted by Huutoniemi of the four panels organized by the Academy of Finland. These panels concern: Social Sciences; Environment and Society; Environmental Sciences; and Environmental Ecology. Unlike Lamont's study, this analysis was explicitly concerned with the effects of the mix of expertise on panels on how customary rules were enacted. The idea was to compare panels with varying degrees of specialization (unidisciplinary - multidisciplinary panels) and with different kinds of expertise (specialist experts - generalists). However, in the course of comparing results from the two studies, other points of comparison beyond expert composition emerged - whether panelists "rate" or "rank" proposals, have an advisory or decisional role, come from the social sciences and humanities fields or from more scientific fields, etc. Our exploratory analysis points to some important similarities and differences in the internal dynamics of evaluative practices that have gone unnoticed to date and that shed light on how evaluative settings enable and constrain various types of evaluative conventions.

Among the most salient customary rules of evaluation, deferring to expertise and respecting disciplinary sovereignty manifest themselves differently based on the degree of specialization of panels: we find that there is less deference in unidisciplinary panels where the expertise of panelists more often overlap. Overlapping expertise makes it more difficult for any one panelist to convince others of the value of a proposal when opinions differ; unlike in multidisciplinary panels, insisting on sovereignty would result in intense conflict for scientific authority. There is also less respect of disciplinary sovereignty in panels composed of generalists rather than experts specialized in particular disciplines, and panels concerned with topics such as Environment and Society that are of interest to wider audiences. In such panels, we find more explicit reference to general arguments and to the role of intuition in grounding decision-making.

While there is a rule against the conspicuous display of alliances across all panels, strategic voting and so-called "horse-trading" appear to be less frequent in panels that "rate" as opposed to "rank" proposals, and in those that have an advisory as opposed to a decisional role. The evaluative technique imposed by the funding agency thus influences the behavior of panelists. Moreover, the customary rules of methodological pluralism and cognitive contextualism are more salient in the humanities and social science panels than they are in the pure and applied science panels, where disciplinary identities may be unified around the notion of scientific consensus, including the definition of shared indicators of quality. Finally, a concern for the use of consistent criteria and the bracketing of idiosyncratic taste is more salient in the sciences than in the social sciences and humanities, due in part to the fact that in the latter disciplines evaluators may be more aware of the role played by (inter)subjectivity in the evaluation

process. While the analogy of democratic deliberation appears to describe well the work of the social sciences and humanities panels, the science panels may be best described as functioning as a court of justice, where panel members present a case to a jury.

#### Conclusion: Practices matter

The customary rules of fairness are part of "epistemic cultures" and essential to the process of collective attribution of significance. In this context, considering reasons offered for disagreement, how those are negotiated, as well as how panelists interpret agreement is crucial to capture fairness as a collective accomplishment. Together, these studies demonstrate the necessity for more comparative studies of evaluative processes and evaluative culture. This remains a largely unexplored but promising aspect of the field of higher education, especially in a context where European research organizations and universities aim to standardize evaluative practices. We look forward to interacting with colleagues as this research area develops.

This article was first published in Q&A magazine, nr. 2, september 2010, pp. 7-10

- monograph "How Professors Think" by Michèle Lamont (Harvard University Press 2009)
- article "Comparing Customary Rules of Fairness: Evidence of Evaluative Practices in Peer Review Panels", by Michèle Lamont and Katri Huutoniemi (forthcoming in an edited volume "Social Science in the Making", University of Chicago Press)