“Score Talk” and its Impact: A Conversation Analytic Investigation of Reviewer Interaction during NIH Study Section Meetings

Background: The ability to secure independent research funding is a crucial component of career advancement in academic science. Research exploring funding outcomes in peer review for the National Institutes of Health (NIH), the “gold standard” for major research funding programs, has identified numerous areas for potential bias in the review process, finding disparities in funding correlated with applicant gender (Ley & Hamilton, 2008; Polhaus et al., 2011), applicant race and ethnicity (Ginther et al., 2011, 2012), and the application’s focus on clinical versus laboratory research (Kotchen et al., 2004, 2006). Recent research aimed at opening up the "black box" of peer review has shown that the language used in written application critiques for NIH grants is a potential indicator for reviewer bias (Kaatz et al., 2015). Extending this focus on language as a site for the potential introduction of bias in peer review, we investigate linguistic and interactional processes in review panel meetings (or “study sections”) of NIH reviewers.

Method: The research team organized four “constructed” study sections of experienced NIH reviewers evaluating proposals that were previously reviewed between 2012 and 2015 by study sections within NIH’s National Cancer Institute. Our goal was to emulate the norms and practices of NIH in all aspects of study design, and methodological decisions were informed by consultation with staff from NIH’s Center for Scientific Review and a retired NIH scientific review officer who assisted the research team in recruiting grants, reviewers, and chairpersons. These constructed study sections were videorecorded from start to finish and reviewer dialogue transcribed by the research team. Using a grounded theory approach to the data and employing the methods of Conversation Analysis (Atkinson & Heritage, 1984)—an empirical approach to examining the structures of communicative social interaction—we identified linguistic and interactional structures that lead to changes in reviewer scores, and thus serve as a potential site for the introduction of bias during the peer review process.

Results: We identified a communicative genre that we identify as *score talk*, in which reviewers halt discussion of the grant currently under review and move into talk about grant scoring criteria. Although NIH provides reviewers with a general rubric for how to score applications (employing a reverse nine-point scale in which 1 corresponds to “Outstanding” and 9 corresponds to “Poor”), an analysis of the transcripts from these meetings shows that there is often inter-reviewer disagreement about the meaning of a score. During score talk, reviewers frequently engage in real-time *score calibration*. Such calibration often entails the use of joking, laughter, and teasing from members of the group; in such cases, reviewers who are the target of these processes may respond by immediately and publicly changing their score. Score talk thus provides an opportunity for one or more participants to alter the final scores that another reviewer has previously assigned to a grant, which may in turn influence group-wide shifts in final score assignments for the grant under discussion.

Conclusion: This exploratory study suggests that score talk provides reviewers with a key opportunity to influence the scoring practices of other reviewers, and may provide a site for the introduction of reviewer bias into the review process.

References

Ginther, D. K., Schaffer, W. T., Schnell, J., Basimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science* 333(6045), 1015–1019.

Ginther D. K., Haak, L. L., Schaffer, W. T., & Kington, R. (2012). Are race, ethnicity, and medical school affiliation associated with NIH R01 type 1 award probability for physician investigators? *Academic Medicine* 87(11), 1516–1524.

Kaatz, A., Magua, W., Zimmerman, D. R., & Carnes, M. (2015). A quantitative linguistic analysis of National Institutes of Health R01 application critiques from investigators at one institution. *Academic Medicine* 90(1), 69-75.

Kotchen T. A., Lindquist, T., Malik, K., & Ehrenfeld, E. (2004). NIH peer review of grant applications for clinical research. *JAMA* 291(7), 836–843.

Kotchen T. A., Lindquist, T., Miller Sostek, A., Hoffmann, R., Malik, K., & Stanfield, B. (2006). Outcomes of National Institutes of Health peer review of clinical grant applications. *Journal of Investigative Medicine* 54, 13–19.

Ley T. J. & Hamilton, B. H. (2008). The gender gap in NIH grant applications. *Science* 322(5907), 1472–1474.

Pohlhaus J. R., Jiang, H., Wagner, R. M., Schaffer, W. T., & Pinn, V. W. (2011). Sex differences in application, success, and funding rates for NIH extramural programs. *Academic Medicine* 86 (6), 759–767.