On the Role of University Size in Assessment of NSERC Discovery Grant Applications

A Report to the Committee on Grants and Scholarships

Douglas Morris

Group Chair, Evolution and Ecology

20 April 2011

Introduction

Assessment of NSERC Discovery Grant Applications is based on only three equally-weighted criteria: Excellence of the Researcher (EoR), Merit of the Proposal (MoP) and Contributions to the Training of Highly Qualified Personnel (HQP). Success and funding rates are nevertheless higher for applicants from 'large' institutions than for applicants from 'medium' and 'small' universities. Several factors might help to explain the pattern. Large universities might attract, recruit, retain, or produce applicants with superior records of achievement. Large institutions might also provide more direct and indirect support of research, place higher emphasis on graduate-level education, attract more and higher-quality graduate students, demand higher productivity from faculty, or provide more time to pursue research.

Several applicants and officials from smaller institutions suspect two other causes, 1, that application review is prejudiced against smaller institutions, and 2, that applicants from smaller institutions are not receiving proper credit for the value of their HQP training. Concerns about 'prejudice' emerge from the Peer Review Manual's statement "*NSERC particularly cautions members against any prior judgement of an application based on the size or reputation of the applicant's institution*" (Section 6.8.1.2). Although well-intentioned, some applicants suspect that simply highlighting institution size increases the likelihood that it will influence the evaluation. Concerns over credit for HQP training at smaller institutions have been raised over the statement "*A researcher working at a university without a graduate program* ***should not be ranked lower due to limited or no graduate student supervision***" (Section 6.8.1.3). The implication is that the researcher might very well be ranked lower for other undefined reasons associated with the absence of a graduate program. The same concern arises from the statement "*Some applicants conduct their work with little interaction with graduate students, as they are appointed to departments that lack master's or PhD programs and may principally supervise undergraduate students;* ***these researchers should not be automatically ranked lower for not supervising master's or PhD students****; however, their application should still provide evidence of high quality training*" (Section 6.8.3.2).

The indicators table is also a concern to those who feel that undergraduate instruction, in particular, is undervalued. High rankings for HQP training require contributions to "***quality, original research***" (very strong); "***high-quality research***" (outstanding), or "***top-quality research***" (exceptional). These 'achievements' are obviously more difficult to attain for an applicant training primarily undergraduate and Master's students than one with several post-doctoral fellows.

It is thus rather important to provide an initial assessment of whether an institution's size plays a disproportionate role in the evaluation and success of Discovery Grant applications, and particularly so with respect to HQP.

Methods

Denis Leclerk kindly provided an anonymous listing of all 3,460 applicants assessed in the February 2011 Discovery Grant competition. Data for each applicant included the evaluation group, stage of applicant (experienced researcher or early career), size of the institution (large, medium, or small), whether the applicant was funded or not, the bin level, and the ranking for each of the three evaluation criteria (rankings for one applicant were missing; this subject was excluded from the analysis). As the main objective was to evaluate potential bias in HQP assessments by university size, I restricted the analysis to experienced researchers (3,002 valid subjects).

I reasoned that bias by institution size could be excluded if the ranking of any evaluation criterion was independent of institution size. I then searched for the best model explaining lack of independence among the three evaluation criteria and institution size with a hierarchical log-linear analysis (SPSS 18). The analysis uses a backward elimination procedure to reduce a multi-dimensional contingency table to only those main effects and sub-sets of interactions among factors accounting for the lack of independence in the data. Interactions that remain represent the 'partial associations' of their overall contribution to the interdependencies among the factors. I grouped the three criteria into two classes (strong and lower; very strong and higher) in order to obtain reliable expected cell frequencies required for the analysis (some of the 81cells are empty; the expected values are also zero and could bias the tests of significance [e.g., no applicants from small universities who ranked I or M on both EoR and MoP ranked VS or higher on HQP]).

Results

The rankings of applicants by the three criteria were not independent (all two-way interactions were highly significant, Table 1).

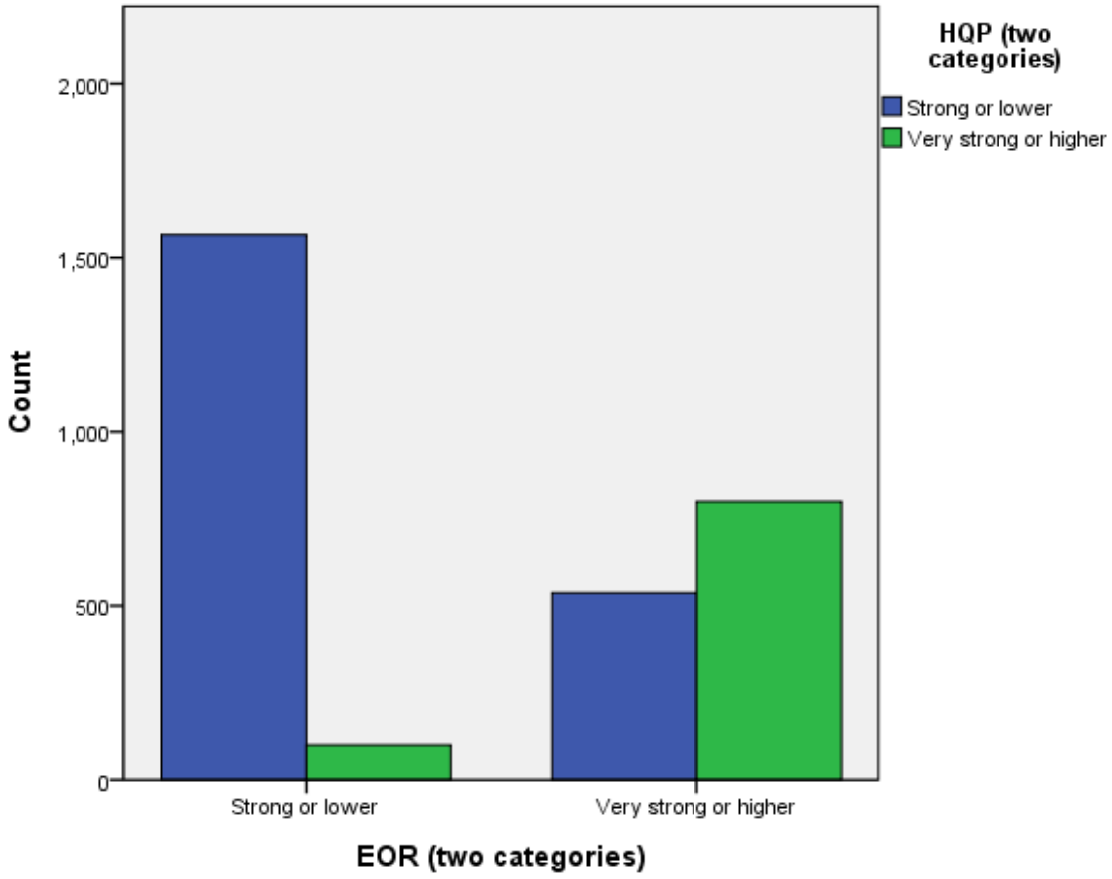
Table 1. Significant interactions retained (tests of partial association) by a hierarchical log-linear analysis evaluating the independence of NSERC Discovery Grant criteria from university size (February 2011 competition).

Interaction	**ΔChi-square***	**Df**	***p***
HQP×EoR	524.7	1	< 0.001
EoR×MOP	346.8	1	<0.001
HQP×MOP	127.1	1	<0.001
Size×EoR	62.3	2	<0.001
Size×MOP	12.9	2	0.002
Size×HQP	12.6	2	0.002

*Change in chi-square if the interaction was deleted from the model.

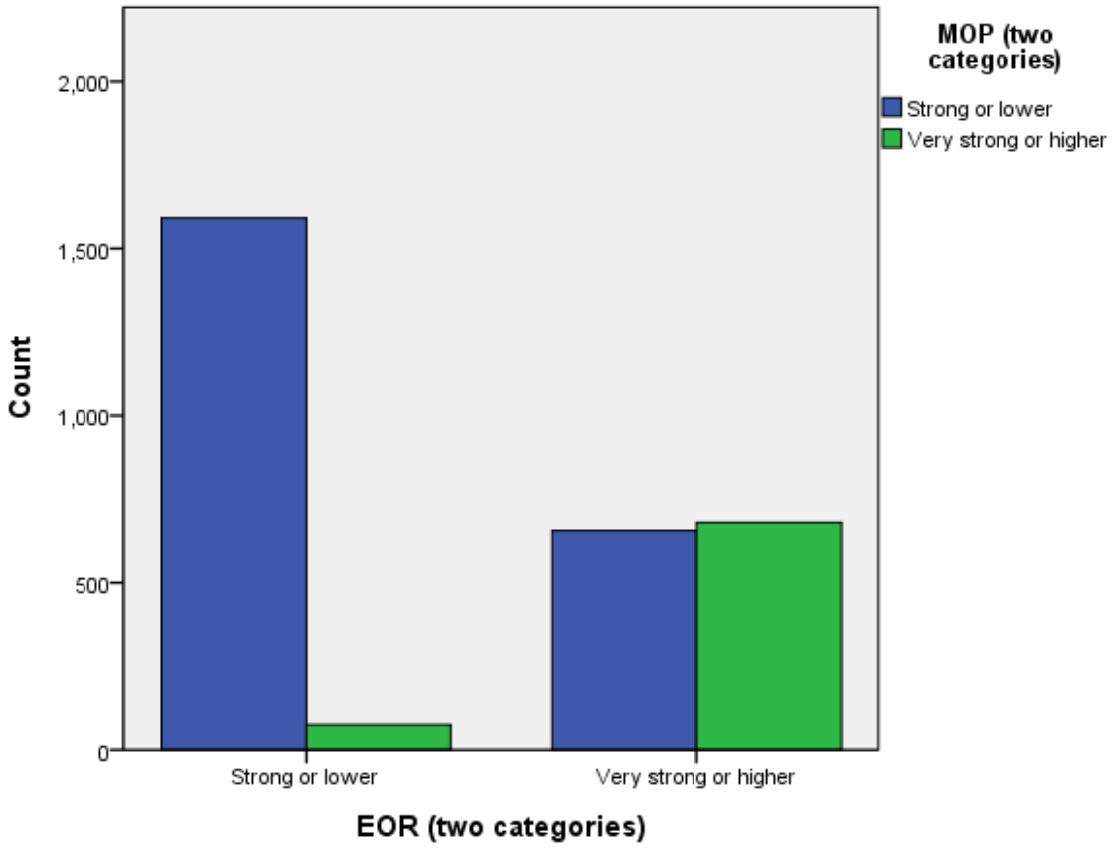
We can gain insight into the interactions by exploring each one sequentially. Figure 1 reveals that few applicants who scored strong or lower for EoR obtained a higher score for HQP. The majority of applicants ranked very strong or higher for EoR also ranked very strong or higher for HQP. Rankings for HQP tended to be lower than those for EoR.

Figure 1. Association between HQP and Excellence of the Researcher.

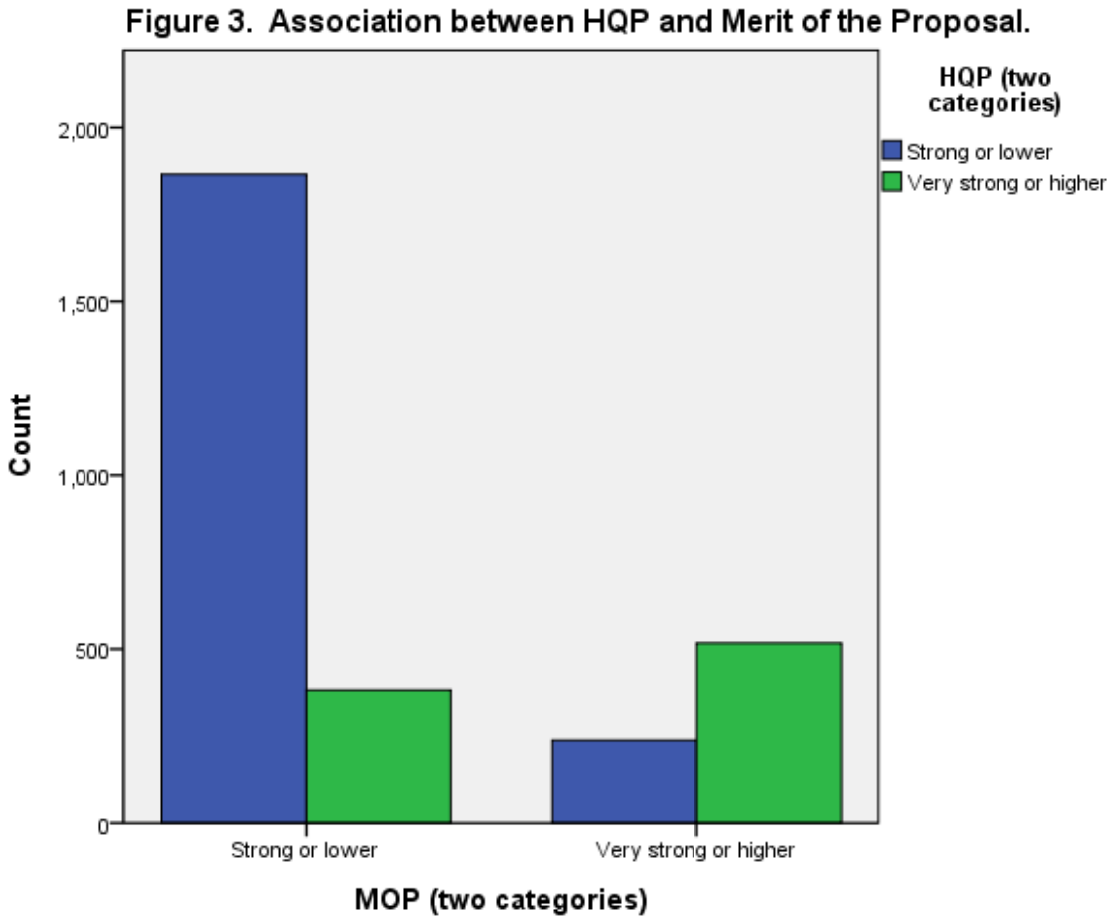


A similar pattern occurred between the rankings for merit of the proposal and excellence of the researcher (Figure 2), although in each instance a smaller proportion of applicants were ranked very strong or higher for MoP than was the case for HQP.

Figure 2. Association between Excellence of the Researcher and Merit of the Proposal.

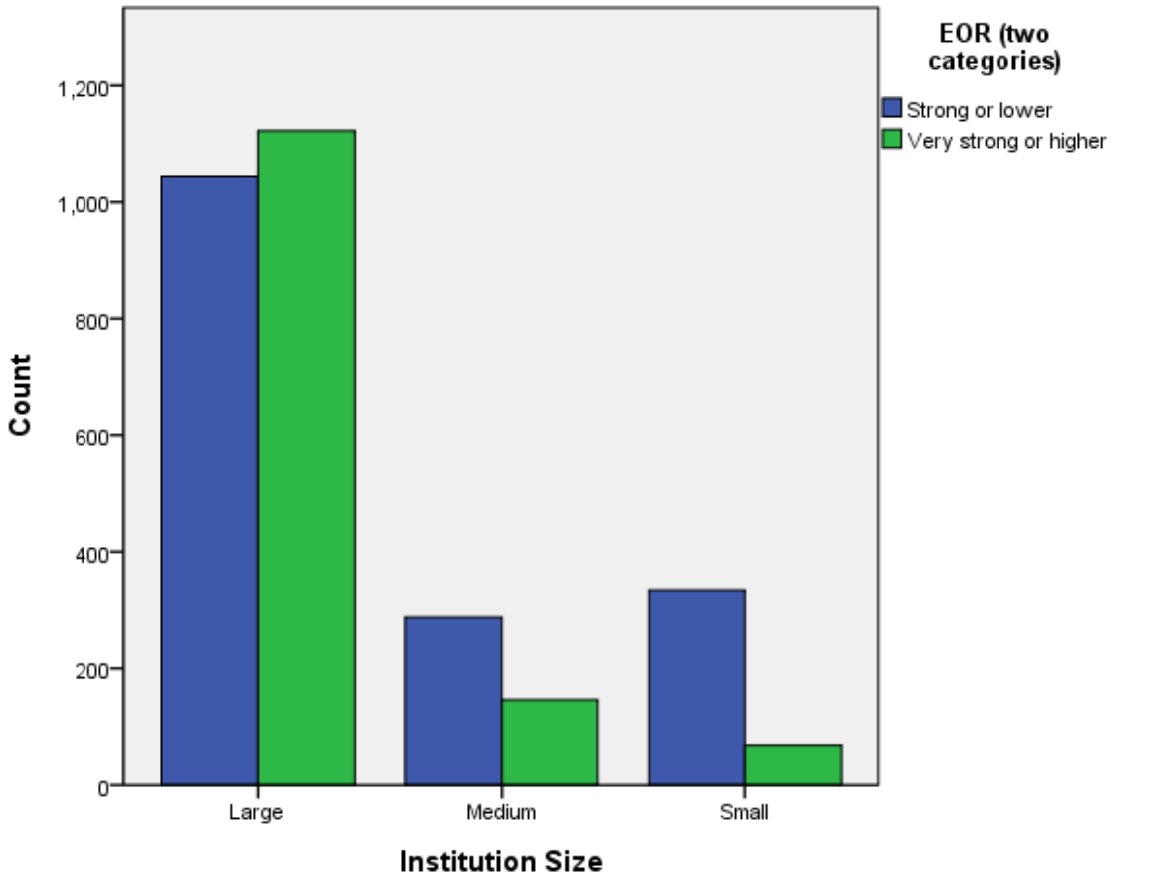


Although fewer applicants scored very strong or higher for merit of the proposal than for excellence of the researcher, a greater proportion of those with very strong or higher rankings for MoP also scored very strong or higher for HQP (Figure 3).



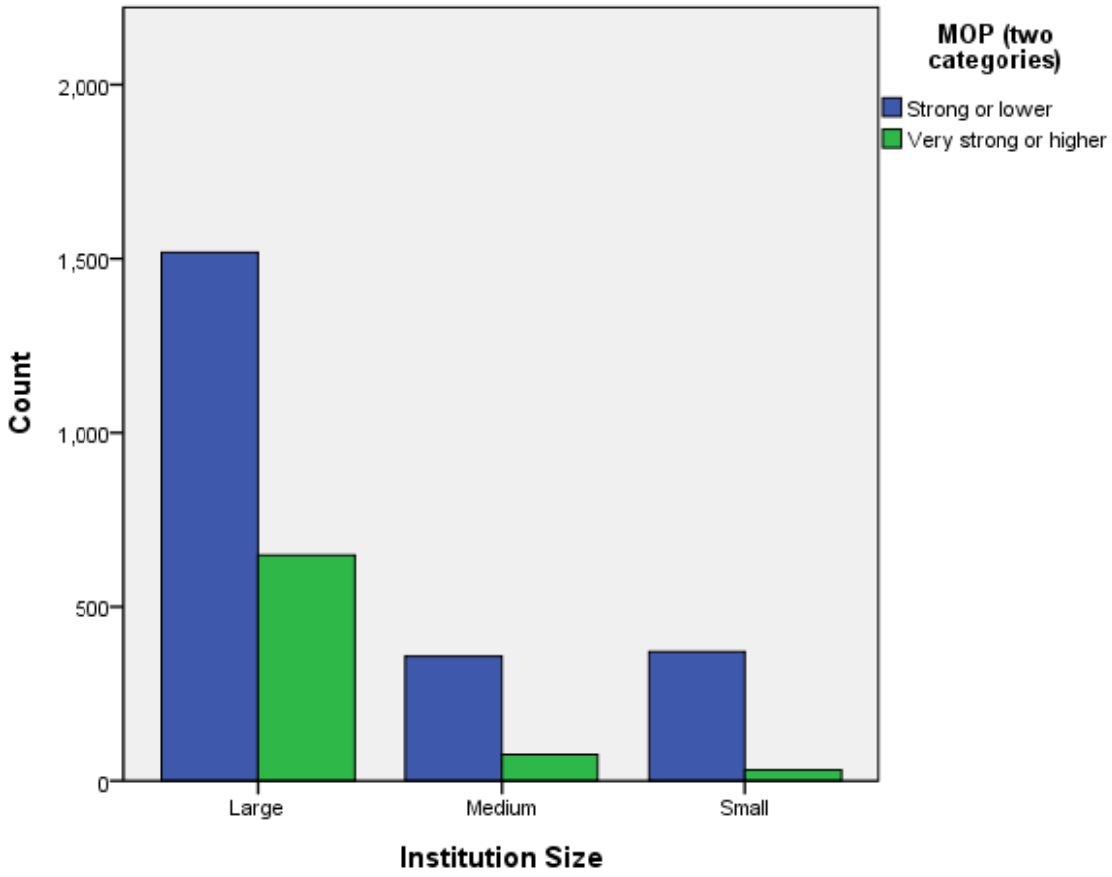
None of the criteria was independent of institution size. The majority of applicants from large institutions were ranked very strong or higher for EoR. The trend was opposite in medium-sized institutions, and the vast majority of applicants from small institutions ranked strong or lower (Figure 4).

Figure 4. Association between Excellence of the Researcher and Institution Size.



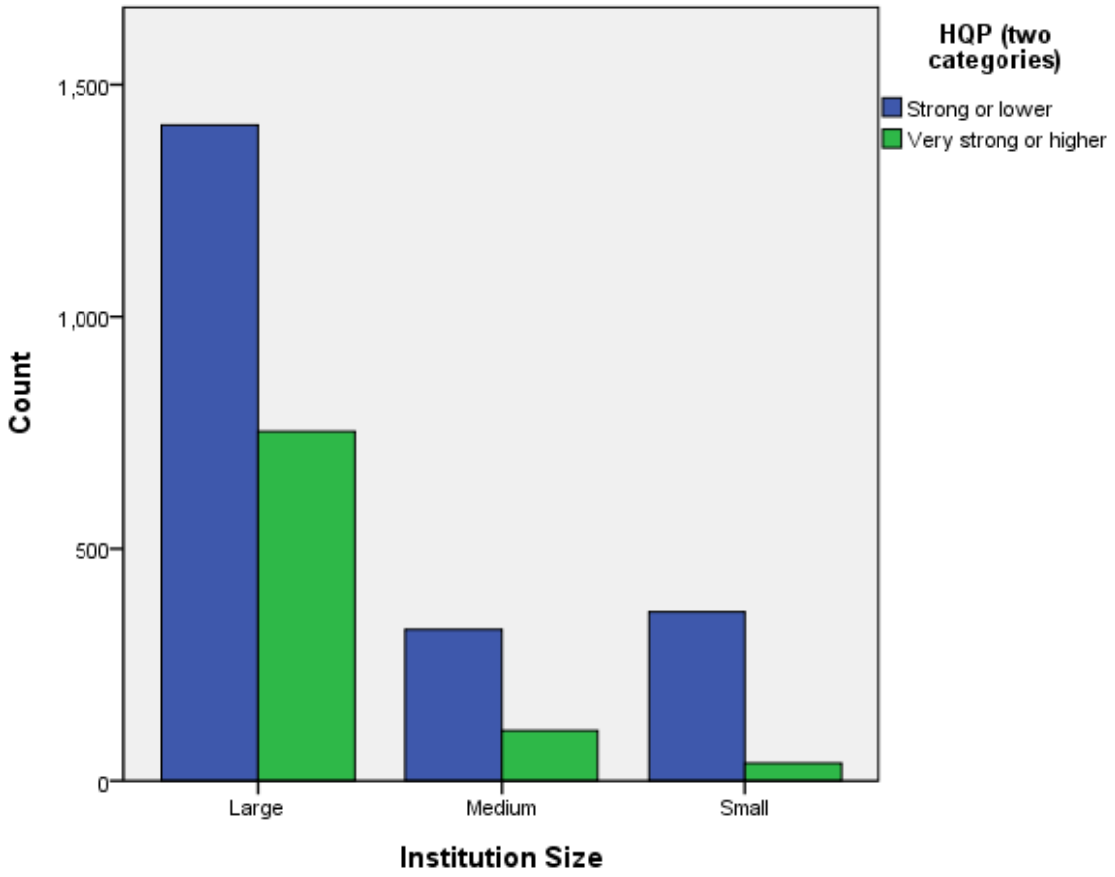
The majority of applicants from all institutions scored strong or lower for MoP reflecting the relatively lower rankings for MoP than for EoR. Nevertheless, the proportion scoring strong or lower was less for medium-sized institutions and least of all for small institutions (Figure 5).

Figure 5. Association between Merit of the Proposal and Institution Size.



A similar trend was evident in the pattern of association between institution size and HQP, again reflecting the lower overall rankings of HQP relative to EoR (Figure 6).

Figure 6. Association between HQP and Institution Size.



Interpretation

The rankings that an applicant receives for the three evaluation criteria were not independent of one another, although their evaluation should be. Applicants who ranked very strong or higher in terms of excellence of the researcher also ranked highly for successful training of HQP and meritorious proposals. These synergistic effects are expected among the best applications.

The associations of the rankings for the three merit criteria among institutions of varying sizes are more problematical. Applicants from smaller institutions consistently ranked lower on all three criteria than did applicants from large institutions. It is crucially important to note that these effects are not redundant with one another. Removal of the Size×HQP interaction, for example, results in a highly significant reduction in the chi-square of the overall log-linear

model. Thus, the partial association of an institution's size with the ranking of its applicants for HQP is significant even when one accounts for all other two-way interactions and main effects in the model.

It is thereby quite possible that applicants involved primarily in undergraduate or MSc training at smaller institutions are receiving less credit for HQP than are applicants from institutions with PhD and post-doctoral programs. This interpretation is consistent with the viewpoint expressed by applicants and administrators from smaller institutions, as well as with statements made during competition week, that Evaluation Groups value BSc training lower than that of PhD or post-doctoral training. One must also wonder whether lower rankings for HQP training also influence the lower rankings that applicants from smaller institutions receive on the other criteria.

Implications

Low rankings in smaller institutions have the potential for dramatic and damaging influences on science in Canada and equality of opportunities for Canada's university students in Natural Sciences and Engineering (NSE). Many of the smaller institutions are located outside of the largest cities and serve different regions and different cohorts of students than do large institutions. There are several major consequences if these trends continue.

- Scientific and engineering expertise will be increasingly concentrated in larger centres and be either unavailable, or inappropriate, to regional priorities.
- Declining success and funding rates in small institutions reduce the opportunities for substantial numbers of otherwise promising recruits into graduate studies and employment in NSE.
- Low success and funding rates for medium-sized institutions may cause them to fall into the 'small-institution' category, and further expand the gap between the 'have' larger institutions and the 'have not' small ones.
- Each effect reinforces the others and will increase the difficulties faced by smaller institutions to recruit and retain high-quality faculty, thus further eroding expertise and NSE opportunities within the regions.

Continued decline in success and funding rates for Discovery Grants in smaller institutions is detrimental to the progress of science in Canada and to opportunities for Canadians to contribute to NSE. Much, but not all, of the responsibility for reversing the trend is appropriately targeted towards the institutions themselves. Smaller institutions will need to explore new and effective means to improve their culture of NSE research and student instruction.

Ultimately, NSERC and other federal and provincial agencies will need to evaluate the expected and appropriate roles for all institutions involved in NSE research. In the short term, it is essential that NSERC removes any potential forms of bias that may account for rankings of evaluation criteria associated with institution size.

Recommendations

1. All references to institution size should be removed from the peer review guidelines.
2. The guidelines should be revised with clear expectations on outcomes of successful HQP training at each level in the HQP pipeline.
3. The indicators table for HQP training should be revised to be consistent with the stated outcomes.
4. The training summary on form 100 should be revised to enable an assessment of the number of students supervised annually rather than a total count by level of training.
5. Any revisions to HQP reporting on forms 100 and 101 must be independent of institution size.
6. NSERC should undertake a formal and detailed analysis of the causes and consequences for differential success and funding rates among institutions of different sizes and with different academic and post-graduate programs.