



# Turning the spit: / Tournons la broche : A perspective on the NSERC Discovery Grant review process / Une perspective sur le processus d'évaluation des subventions à la découverte du CRSNG

Paul D. McNicholas

The NSERC Discovery Grants program remains the single most important source of research funding for statistics faculty in Canada. I had the privilege of being the statistics section chair for the 2015 and 2016 Discovery Grant competitions. Before that, I was incoming section chair for the 2014 competition. I am writing this article to share the perspective I have gained, and to offer some small pieces of advice to applicants and reviewers. But first, some background.

Discovery Grants fund research in the natural sciences and engineering (NSE), and are evaluated by discipline-specific evaluation groups (EGs). As statisticians, we are part of EG 1508: Mathematics and Statistics, which consists of three sections: pure math, applied math, and statistics. The EG meets in Ottawa each February and the statistics section of EG 1508 has around 11 members. Five people evaluate each application: two reviewers and three readers. An application is reviewed based on past record (Excellence of Researcher; EoR), proposed research (Merit of Proposal; MoP), and previous and planned training (Highly Qualified Personnel; HQP). In each of the three categories, applications are rated as exceptional (E), outstanding (O), very strong (V), strong (S), moderate (M), or insufficient (I) in accordance with the merit indicators<sup>1</sup>. Based on these ratings, applications are placed into one of 16 “bins” denoted A (corresponding to EEE), B (e.g., EEO), C (e.g., EOO), down to P (corresponding to III). At the end of the competition week, funding is allocated to each bin in decreasing (or non-increasing, at least) amounts. Note that I write about the evaluation of applications rather than applicants; this is important, it is the application and not the applicant that is evaluated.

My first year on the EG was the 2014 competition, i.e., for applications submitted to NSERC in fall 2013. I must admit to some surprise at being asked to join the EG and even more surprise at being asked to be incoming section chair – which means being section chair the following year. I thoroughly enjoyed the preparation for the competition

<sup>1</sup> [http://www.nserc-crsng.gc.ca/\\_doc/Professors-Professeurs/DG\\_Merit\\_Indicators\\_eng.pdf](http://www.nserc-crsng.gc.ca/_doc/Professors-Professeurs/DG_Merit_Indicators_eng.pdf)

Le programme des subventions à la découverte du CRSNG demeure la plus importante source de financement de la recherche pour les professeurs de statistique au Canada. J'ai eu le privilège de présider la section de statistique pour les concours des subventions à la découverte 2015 et 2016. Auparavant, j'avais été futur président du groupe lors du concours 2014. J'ai rédigé cet article pour partager l'expérience que j'ai acquise et offrir quelques conseils aux candidats et aux évaluateurs. Mais d'abord, un peu de contexte.

Les subventions à la découverte financent la recherche en sciences naturelles et en génie (SNG) et sont évaluées par des groupes d'évaluation spécifiques à la discipline (GE). Nous statisticiens faisons partie du GE 1508 : Mathématiques et statistiques, qui inclut trois sections : maths pures, maths appliquées et statistique. Le GE se réunit à Ottawa chaque février et la section de statistique du GE 1508 compte environ 11 membres. Cinq personnes évaluent chaque demande : deux évaluateurs et trois lecteurs. Les demandes sont évaluées sur la base du dossier antérieur (Excellence du chercheur : EC), de la recherche proposée (Mérite de la proposition : MP) et la formation antérieure et prévue (Personnel hautement qualifié : PHQ). Dans chacune de ces trois catégories, les demandes sont cotées « Exceptionnel » (E), « Remarquable » (R), « Très élevé » (T), « Élevé » (É), « Moyen » (M) ou « Insuffisant » (I), selon des indicateurs de mérite<sup>1</sup>. En fonction de ces cotes, les demandes sont placées dans l'une de seize « catégories » appelées A (correspondant à EEE), B (p. ex., EER), C (p. ex., ERR), jusqu'à P (correspondant à III). À la fin de la semaine de concours, un financement est affecté à chaque catégorie par montants décroissants (ou, du moins, non croissants). Notez que je parle d'évaluer des demandes et non des candidats. C'est important : c'est la demande, et non le candidat, qui est évaluée.

Ma première année au GE fut celle du concours 2014, à savoir celui qui évaluait les demandes soumises au CRSNG à l'automne 2013. Je dois admettre que j'ai été quelque peu surpris de me voir invité à rejoindre le GE et plus encore surpris de me voir demander d'être le futur président de la section – ce qui

<sup>1</sup> [http://www.nserc-crsng.gc.ca/\\_doc/Professors-Professeurs/DG\\_Merit\\_Indicators\\_fra.pdf](http://www.nserc-crsng.gc.ca/_doc/Professors-Professeurs/DG_Merit_Indicators_fra.pdf)



week; the opportunity to read so many statistics applications was wonderful and gave me a window into some of the really interesting research going on across Canada. From the point of view of meeting and getting to know colleagues from home and abroad, that first competition week was very enjoyable. The actual deliberations, however, were less enjoyable. As incoming chair, I was encouraged to spend some time in the pure and applied math rooms during the deliberations; I even served as a reader on some of their files as well as a few files from different EGs. This was very beneficial because it gave me an appreciation of the similarities and differences between the sections within our EG and, to a lesser degree, between EGs. Alas, the within EG differences did not cast us statisticians in a positive light.

One of only two people to win both a Nobel prize and an Oscar, George Bernard Shaw, reputedly said:

*Put an Irishman on the spit and you can always get another Irishman to turn him.*

Unfortunately, I think there is a bit of this mentality in the statistics community and, most importantly, perhaps more of this in our community than in the pure or applied math communities. At least, this is how it seemed to me as I moved between rooms during the 2014 deliberations. This mentality was apparent both during deliberations and from reading external reviewers' reports; in fact, reading these reports was one of the most disappointing aspects of my work on the EG. Some external reviewer reports are just excessively negative and, in practice, are less of a concern because they are often ignored altogether, and formally so. However, excessively negative reports are relatively rare. More common are those that, while not coming across as excessively negative, focus disproportionately on the weaker aspects of an application. This latter sort can be dangerous and seems to be born out of a "let's see what I can criticize about this" mentality rather than out of a more balanced approach.

On the last day of the 2014 competition week, most EG members went home but the Executive Committee stayed around to study the bin distributions and recommend funding levels (with the constraint of an overall fixed budget). From a statistician's perspective, the relative bin distributions were not pleasant viewing (Figure 1). I distinctly remember my flight home from Ottawa after that meeting. I took a flight one hour earlier than the one I was booked on. I had to pay for the privilege but it was worth it: I just wanted to get out of there. I felt awful. Everybody involved did his or her best. As far as I could tell, our ratings were very self-consistent (i.e., consistent within the section) throughout the week. There were none but good intentions and yet, in the end, we produced a result that did not do our community justice. I concluded that this was nobody's fault; it was a matter of

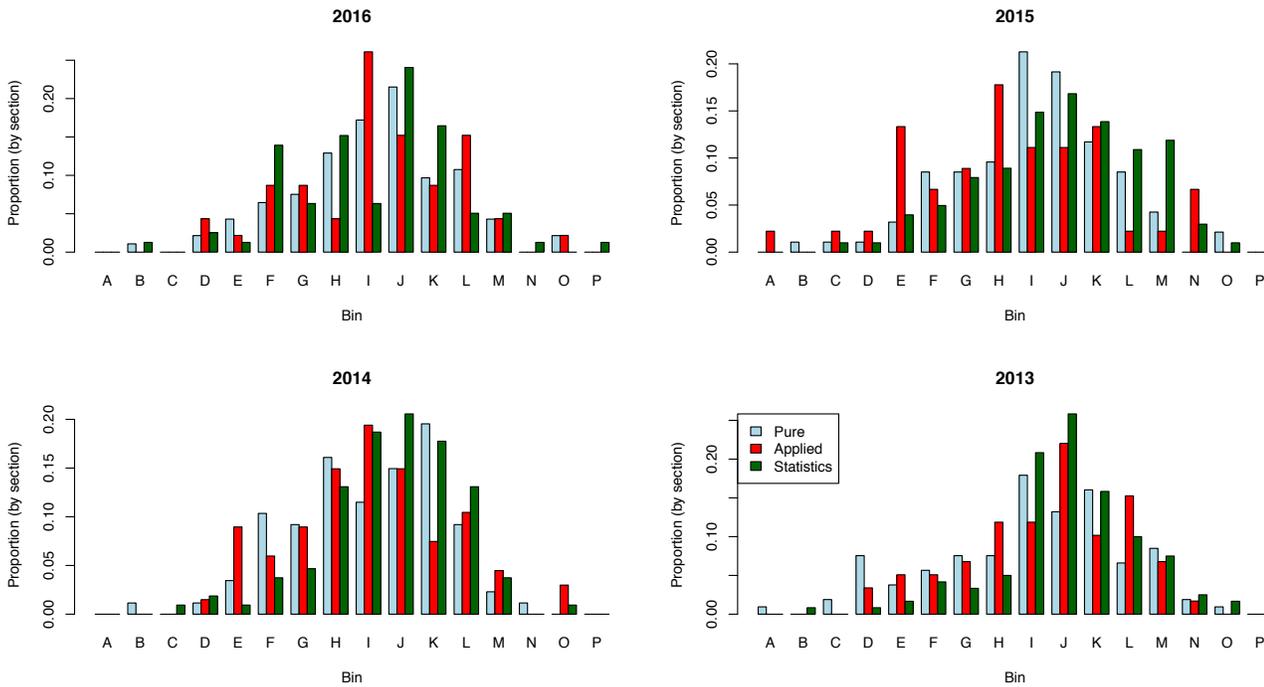
signifiait de la présider l'année suivante. J'ai beaucoup aimé la préparation à la semaine du concours, cette fabuleuse opportunité de lire tant de demandes en statistique, qui m'ont donné un aperçu de la recherche si intéressante qui a lieu partout au Canada. L'occasion de rencontrer et d'apprendre à connaître des collègues canadiens et étrangers a aussi rendu cette première semaine de concours très agréable. Les délibérations proprement dites, par contre, n'ont pas été aussi agréables. À titre de futur président, on m'a encouragé à passer du temps dans les salles de maths pures et appliquées pendant leurs délibérations; j'ai même été lecteur de certaines de leurs demandes, ainsi que pour quelques demandes d'autres GE. Cela m'a été très profitable car j'ai vu les similarités et les différences qui pouvaient exister entre les sections de notre GE et, dans une certaine mesure, entre GE. Malheureusement, les différences entre nos sections ne donnent pas une perception très positive de nous, statisticiens.

George Bernard Shaw, l'un des deux individus à avoir jamais remporté un prix Nobel et un Oscar, aurait dit un jour :

*Mettez un Irlandais sur la broche et vous trouverez toujours un autre Irlandais pour la tourner.*

Malheureusement, je pense que l'on retrouve un peu de cette mentalité dans la communauté statistique et, plus important encore, peut-être davantage chez nous que dans les communautés de maths pures ou appliquées. C'est du moins l'impression que j'ai eue en passant d'une salle à l'autre pendant les délibérations de 2014. Cette mentalité était apparente à la fois pendant les délibérations et à la lecture des rapports des évaluateurs externes : en fait, la lecture de ces rapports a été l'un des aspects les plus désagréables de mon travail au GE. Certains rapports d'évaluateurs externes sont excessivement négatifs et posent donc, dans la pratique, moins de problèmes car ils sont tout simplement – et formellement – ignorés. Cependant, les rapports excessivement négatifs sont relativement rares. Bien plus communs sont ceux qui, sans sembler excessivement négatifs, mettent disproportionnellement l'accent sur les points faibles d'une demande. Ce type de rapports peut être dangereux et semble issu d'une mentalité « voyons ce que je peux bien critiquer ici » plutôt que d'une approche plus équilibrée.

Le dernier jour de la semaine du concours 2014, la plupart des membres du GE sont rentrés chez eux, mais le comité exécutif devait rester pour étudier les répartitions entre catégories et recommander les niveaux de financement (dans la limite d'un budget global fixe). Du point de vue d'un statisticien, les répartitions relatives entre catégories n'étaient pas belles à voir (Figure 1). Je me souviens très distinctement de mon vol de retour après cette réunion. J'ai pris un vol une heure plus tôt que prévu pour quitter Ottawa. J'ai dû payer la différence,



**Figure 1: Bin distributions from the last four Discovery Grant competitions, broken down by section / Répartitions entre catégories des quatre derniers concours de Subventions à la découverte, par section**

culture, a culture that led to the statistics section having a notably different bin distribution to pure or applied math.

Contrary to some arguments I heard before joining the EG – and I will elaborate on some of these later – the mathematicians were not the problem. It was us, and us alone. While I observed quite a bit of movement of EG members between the pure and applied math rooms (each section has its own room), relatively few cases in the statistics room involved members from the pure or applied rooms – and, when they did, the application was often in financial math. For all intents and purposes, statistics applications are reviewed by statisticians. That this was a self-inflicted injury was both disappointing and encouraging – if the problem was based within the statistics section then it could be solved within the statistics section.

Of the self-inflicted wounds, the most striking was the distribution for MoP; compared to the math sections, we were disproportionately rating proposals as strong. In fact, we rated over 60% of all proposals as strong. I later discovered that this tendency was not peculiar to the 2014 competition (cf. the corresponding distribution for 2013 in Figure 2). I also discovered that not once in the history of the current evaluation system, which was first used in the 2009

mais cela en valait la peine : je voulais partir dès que possible. Je me sentais très mal. Tous les participants à la réunion avaient fait de leur mieux. Pour autant que je sache, nos cotes avaient été tout à fait cohérentes (pour la section) pendant toute la semaine. Tout le monde était de bonne foi et pourtant, au final, nous avons produit un résultat qui ne rendait pas justice à notre communauté. J'en ai conclu que ce n'était de la faute à personne; c'était plutôt un problème de culture, une culture qui a donné à la section de statistique une répartition entre catégories très différente de celles de maths pures ou appliquées.

Contrairement à certains arguments que j'avais entendus avant de rejoindre le GE – sur lesquels je reviendrai – ce ne sont pas les mathématiciens qui posent problème. C'est nous, et nous seuls, qui sommes responsables. Alors que j'ai observé beaucoup de mouvement de membres des GE entre les salles de maths pures et appliquées (chaque section a sa propre salle), relativement peu de demandes en statistique avaient des liens avec des membres des salles de maths – et il s'agissait alors généralement de demandes en maths financières. À toutes fins pratiques, les demandes en statistique sont évaluées par des statisticiens. Le fait qu'il s'agissait d'un préjudice auto-infligé est à la fois décevant et encourageant : si le problème a ses origines dans la section de statistique, il peut être résolu par la



competition, did the statistics section evaluate a proposal as exceptional. Just think about some of the people who would have written proposals during those years. Either none of them wrote exceptional proposals or something is wrong with the way the proposals were evaluated; the former is hard to believe. Because no proposal was rated exceptional, no statistician had ever ended up in Bin A, and it was 2013 before the first one made it into Bin B; again, think about some of the people who would have submitted applications during those years.

I admit that I took what happened in the 2014 competition a bit personally. For me, the mission became to make sure

section elle-même.

De tous les préjudices auto-infligés, le plus flagrant concerne la répartition du MP : par rapport aux sections de maths, nous cotons de manière disproportionnée les demandes comme « Élevé ». En fait, nous avons noté plus de 60 % de toutes les demandes « Élevé ». J’ai découvert par la suite que cette tendance n’était pas limitée au concours 2014 (voyez, en Figure 2, la répartition correspondante pour 2013). J’ai également découvert que jamais, de toute l’histoire du système d’évaluation actuel, qui a été inauguré lors du concours 2009, la section de statistique n’a jugé une seule demande « Exceptionnel ». Pensez un instant à toutes les personnes qui

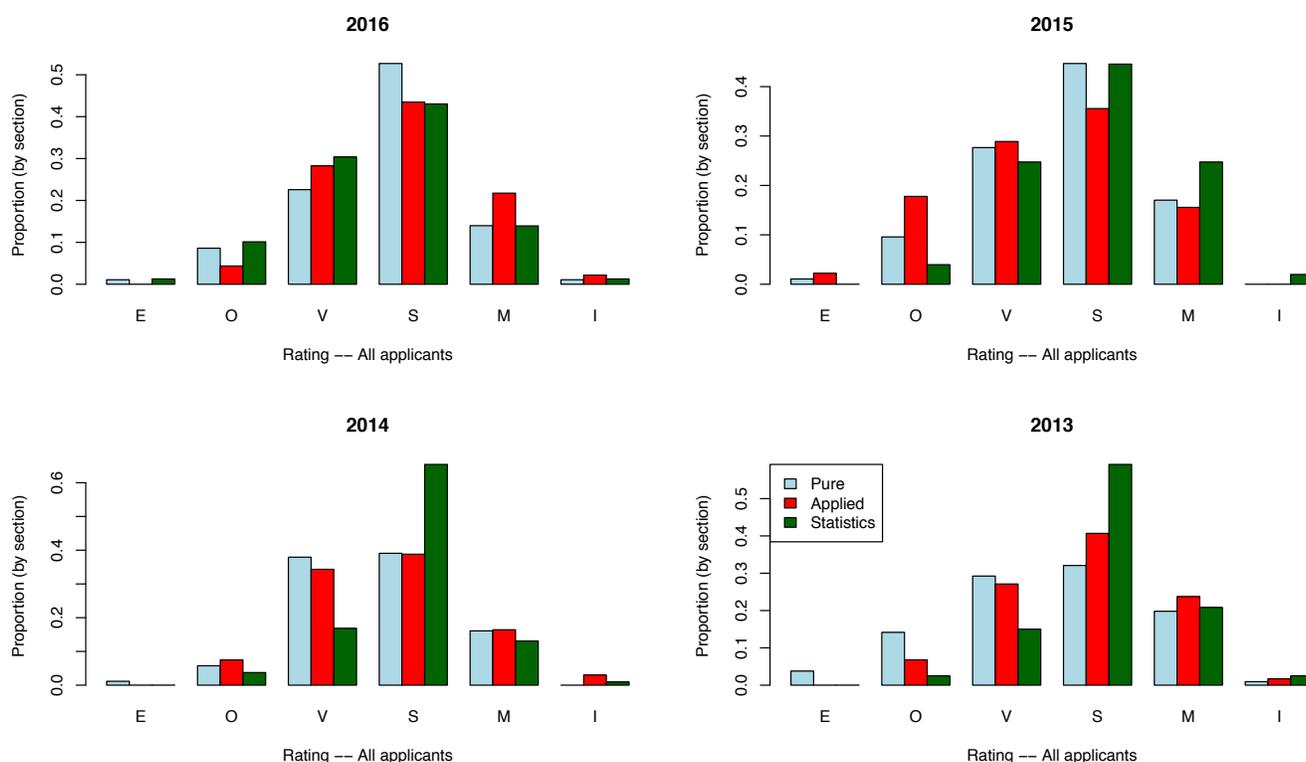


Figure 2 : MoP ratings for all applications from the last four Discovery Grant competitions / Cotes de MP des demandes des quatre derniers concours de Subventions à la découverte

that we (the statistics section) got our fair share in the 2015 and 2016 competitions or, more accurately, to convince my colleagues in the statistics section that this should be the case. This would mean that our deliberations, and use of the scale, had to be such that the resulting distribution was similar to the math sections. Crucially, that is not to say that we had to use the very same standards; this is not even the expectation from NSERC – that is one reason why there are different sections in the first place. Comparing distributions across the three sections is complicated. There

ont soumis des demandes depuis cette date. Soit aucune d’entre elle n’a rédigé de proposition exceptionnelle, soit il y a un problème dans la manière dont les demandes ont été évaluées; il est bien difficile de croire à la première hypothèse. Étant donné qu’aucune demande n’a jamais été jugée « Exceptionnel », aucun statisticien ne s’est jamais retrouvée en catégorie A et il a fallu attendre 2013 avant que la première demande ne tombe en catégorie B; pensez une fois de plus à toutes les personnes qui ont soumis des demandes ces années-là.



are several reasons why, and I will attempt to explain using the distributions from the 2014 competition. For better or worse, applications from early career researchers (ECRs) are evaluated in the same way, using the same merit indicators, as those from established researchers (ERs) with only one exception. The one exception is that, for HQP, it is possible for an ECR application to receive a moderate rating without any training experience, provided that the training plan is satisfactory. Note that an ECR application could only receive a rating above moderate for HQP if it was merited in the same way as for an ER application, and this is very rare.

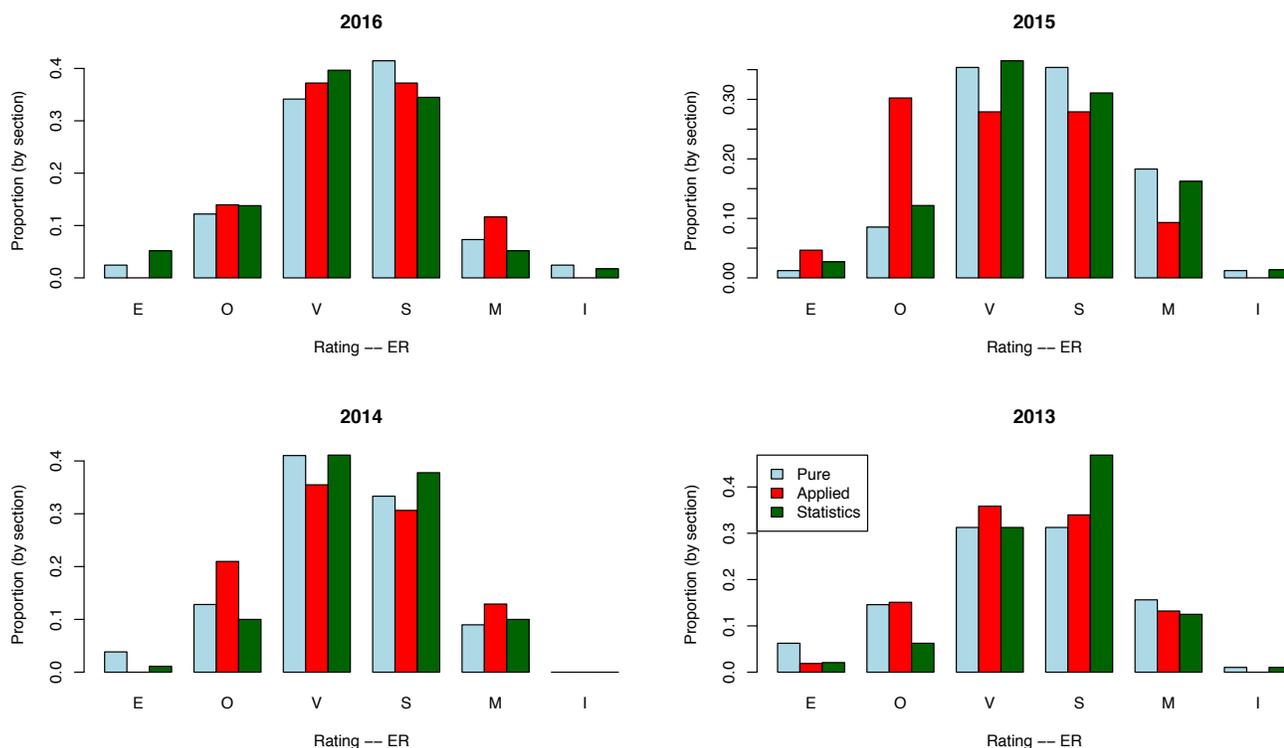
A typical ECR application to one of the math sections will come from an applicant who spent several years as a postdoc prior to her/his first faculty appointment. This means that the six-year window for a math ECR will typically contain only time spent as a PhD student, a postdoctoral fellow, or an assistant professor. A statistics ECR, however, will often have some time spent as a master's and/or undergraduate student as part of the six-year window, e.g., when I first applied, three of my six years were spent as an undergraduate. This means that applications from math ECRs have a shot at being rated anywhere on the scale for EoR, whereas a rating above very strong is almost impossible for a statistics ECR – in the past four years, not one statistics ECR application was rated above very strong for EoR. Accordingly, the bin distribution for statistics ECRs is expected to look inferior to the bin distribution for math ECRs. From my point of view, there was nothing the statistics section could do about this; it is just a consequence of how applications from ECRs have to be evaluated under the current system. Also, because the number of ECRs applying in a given year is relatively small, this issue was not going to prevent us from achieving a similar overall bin distribution to the math sections.

Looking at the distribution of EoR for ERs across the three sections is much more instructive (Figure 3). In 2014, the distributions were not drastically different but it looks as if the statistics section was relatively reluctant to give ratings above very strong. The corresponding figure for HQP (Figure 4) reveals that the statistics section rated more than half of the applications as strong. Finally, recalling the MoP distribution for all applications (Figure 2) – being an ECR should not really matter here – there was an obvious tendency for the statistics section to rate applications as strong. With over 50% of statistics applications receiving a rating of strong for HQP and more than 60% rated as strong for MoP, it is no wonder that the overall bin distribution was disappointing. Over the subsequent two competitions, i.e., 2015 and 2016, the section worked to improve the situation. A big part of doing this was addressing the aforementioned issues, which boiled down to using the entire scale for all three categories. I think our work was successful (cf. Figures 1-4), and more

Je dois dire que j'ai été quelque peu offensé des résultats du concours 2014. Pour moi, il s'agissait désormais de tout faire pour que nous (section de statistique) obtenions notre juste part lors des concours 2015 et 2016 ou, plus exactement, de convaincre mes collègues de la section de statistique que cela devait être le cas. Pour cela, nos délibérations et notre recours à l'échelle devaient garantir une répartition similaire à celles de sections de mathématiques. Il est important de noter que cela ne signifie pas que nous devons utiliser des normes identiques; le CRSNG ne l'exige nullement – c'est là justement l'une des raisons pour lesquelles il existe trois sections distinctes. Il est compliqué de comparer les répartitions d'une section à l'autre et ce pour diverses raisons, que je vais tâcher d'expliquer à l'aide des répartitions du concours 2014. Pour le meilleur ou pour le pire, les demandes des chercheurs en début de carrière (CDC) sont évaluées de la même manière, à l'aide des mêmes indicateurs de mérite, que celles des chercheurs établis (CE) à une exception près : pour le PHQ, il est possible pour une demande de CDC d'obtenir la cote « Moyen » sans aucune expérience de formation, si le plan de formation est satisfaisant. Notez qu'une demande de CDC ne peut recevoir une cote supérieure à « Moyen » pour le PHQ que si elle répond aux mêmes critères qu'une demande de CE, ce qui est très rare.

Dans les sections de maths, une demande de CDC typique provient d'un candidat qui a bénéficié d'une bourse postdoctorale pendant plusieurs années avant sa première nomination universitaire. Cela signifie que pendant les six dernières années, le CDC en maths aura généralement été étudiant au doctorat, boursier postdoctoral ou professeur adjoint. Un CDC en statistique, par contre, aura souvent été étudiant de maîtrise ou de premier cycle pendant cette même période de six ans : lorsque j'ai soumis ma première demande, j'avais été étudiant sous-gradué trois années sur six. Cela signifie que les demandes soumises par les CDC en maths ont une chance d'obtenir la cote « Exceptionnel » pour l'EC, tandis qu'il est presque impossible pour un CDC en statistique d'obtenir une cote supérieure à T – depuis quatre ans, aucune demande de CDC en statistique n'a reçue de cote supérieure à T pour l'EC. Par conséquent, la répartition entre catégories pour les CDC en statistique sera forcément inférieure à celle des CDC en maths. Selon moi, il n'y a rien que la section de statistique puisse changer à cela : c'est simplement une conséquence de la manière dont les demandes des CDC doivent être évaluées dans le système actuel. Cependant, étant donné que relativement peu de CDC soumettent de demandes chaque année, ce problème n'allait pas nous empêcher d'arriver à une répartition entre catégories similaire à celle des sections de maths.

Il est beaucoup plus instructif d'étudier la répartition de l'EC pour les CE des trois sections (Figure 3). En 2014, les répartitions n'étaient pas radicalement différentes, mais il semble que la section de statistique hésitait à donner une cote supérieure à « Très élevé ». La figure correspondante



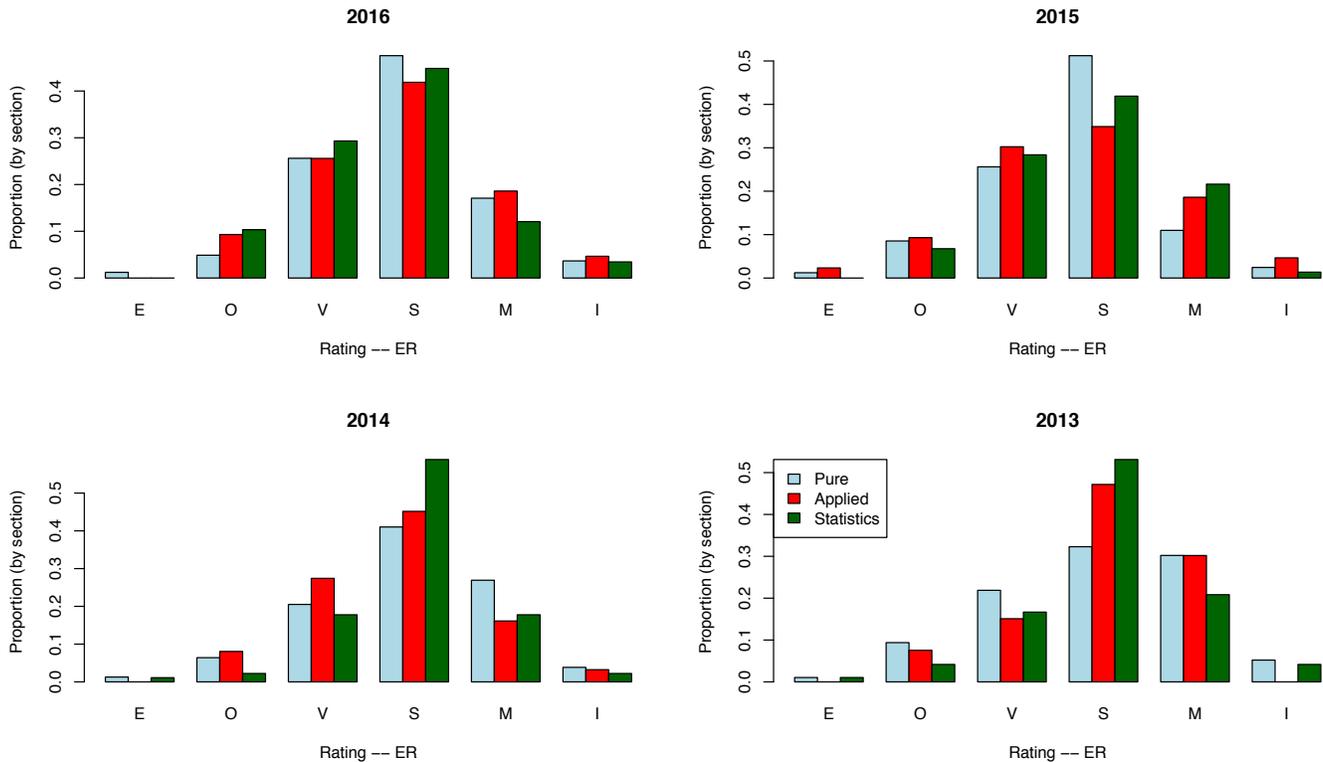
**Figure 3 : EoR ratings for ERs from the last four Discovery Grant competitions / Cotes d'EC pour les CE dans les quatre derniers concours de Subventions à la découverte**

successful in 2016 than 2015 – largely because getting MoP “right” took two years (Figure 2).

I will avoid the temptation to wax lyrical about the differences between the distributions for 2015 and 2016 compared to 2013 and 2014. However, before leaving it to the reader to draw their own conclusions on the finer points, I should mention two issues. First, the applied math distribution for 2015 looks out of place; for what it is worth, less than 20% of applications to the EG in 2015 were to the applied math section, and I am told that a lot of their “stars” applied in 2015. Anyway, steps were taken in 2016 to make sure that we would be better able to spot across-section differences such as this early in the competition week. Second, the statistics section suffers from the “non-NSE” phenomenon, i.e., the tendency of applications that are clearly not within the NSE to be dealt with by the statistics section. I need to make up a somewhat ridiculous example here, to avoid a past applicant linking it to their application, so imagine an application that consists only of working out the mean length of a certain brand of jellybeans and comparing lengths across colours. There is nothing new in the NSE here, and the associated notice of intent will trigger a warning, but if the application comes in, it has to be

pour le PHQ (Figure 4) révèle que la section de statistique a coté plus de la moitié des demandes « Élevé ». Enfin, si l'on se souvient de la répartition du MP pour toutes les demandes (Figure 2) – car le fait d'être CDC ne devrait pas vraiment avoir d'importance ici – on voit une nette tendance pour la section de statistique de coter les demandes « Élevé ». Avec plus de 50 % des demandes en statistique cotés « Élevé » pour le PHQ et plus de 60 % cotées « Élevé » pour le MP, il n'est pas étonnant que la répartition globale soit décevante. Sur les deux concours suivants, 2015 et 2016, la section a travaillé à améliorer la situation. En grande partie, elle a cherché à traiter les problèmes susmentionnés, à savoir à utiliser toute l'échelle de notation pour les trois catégories. Je crois que nous avons réussi (voir Figures 1-4), et ce davantage en 2016 qu'en 2015 – notamment parce qu'il a fallu deux ans pour « bien noter » le MP (Figure 2).

J'éviterai la tentation des envolées lyriques sur les différences entre les répartitions pour 2015 et 2016 par rapport à 2013 et 2014. Cependant, avant de laisser le lecteur tirer ses propres conclusions concernant les détails, je tiens à mentionner deux points. Tout d'abord, la répartition des maths appliquées pour 2015 ne semble pas logique; sachez qu'en 2015, moins de 20 % des demandes au GE étaient des demandes de maths



**Figure 4 : HQP ratings for ERs from the last four Discovery Grant competitions / Cotes de PHQ pour les CE dans les quatre derniers concours de Subventions à la découverte**

evaluated and it will inevitably end up with the statistics section. The pure math section does not see non-NSE applications like this and they are very rare in the applied math section. Therefore, these non-NSE applications are creating an artificial difference in overall bin distributions and in the distributions of ratings for each category. To give an example, consider the bin distribution for the 2015 application year and my best guess at the same distribution without the non-NSE cases (Figure 5). Note that there is no official recording of non-NSE cases, and I could not be in the room for every case (e.g., because of conflict of interest), and so it can only be my best guess.

I want to emphasize that I am not using non-NSE as a synonym for biostatistics here. Some of the non-NSE applications I saw did have medical applications but it was not a majority. Over my three years on the EG, I did not observe a single case where a biostatistics application was treated unfavorably within the basic rules. Of course, some may consider the basic rules, e.g., that NSERC funds research in the NSE, ipso facto unfair but that is another article altogether, written by someone else altogether. From Figure 5, the pure math and statistics bin distributions for

appliquées – et l’on m’a dit qu’un certain nombre d’étoiles du domaine étaient de ce nombre. Quoiqu’il en soit, des mesures ont été prises en 2016 pour permettre la détection précoce de telles différences entre sections dès le début de la semaine de concours. Deuxièmement, la section de statistique souffre du phénomène « non-SNG », à savoir la tendance à soumettre à la section de statistique des demandes qui ne sont clairement pas des demandes SNG. Il me faut ici inventer un exemple un peu ridicule, pour éviter qu’un ancien candidat ne se reconnaisse. Imaginez donc une demande qui consisterait uniquement à calculer la longueur moyenne d’une certaine marque de bonbons et d’en comparer la longueur pour chaque couleur. Rien de nouveau en ce qui concerne les SNG dans cette demande et l’avis d’intention associé devrait déclencher un avertissement, mais si la demande est soumise, elle doit être évaluée et elle finira inévitablement dans la section de statistique. La section de maths pures n’a jamais à évaluer de demandes non-SNG de la sorte et elles sont très rares dans la section de maths appliquées. Par conséquent, ces demandes non-SNG créent une différence artificielle dans les répartitions entre catégories globales et dans les répartitions des cotes pour chaque catégorie. Ainsi, comparez la répartition entre catégories pour l’année de demandes 2015 et mon estimation de ces mêmes

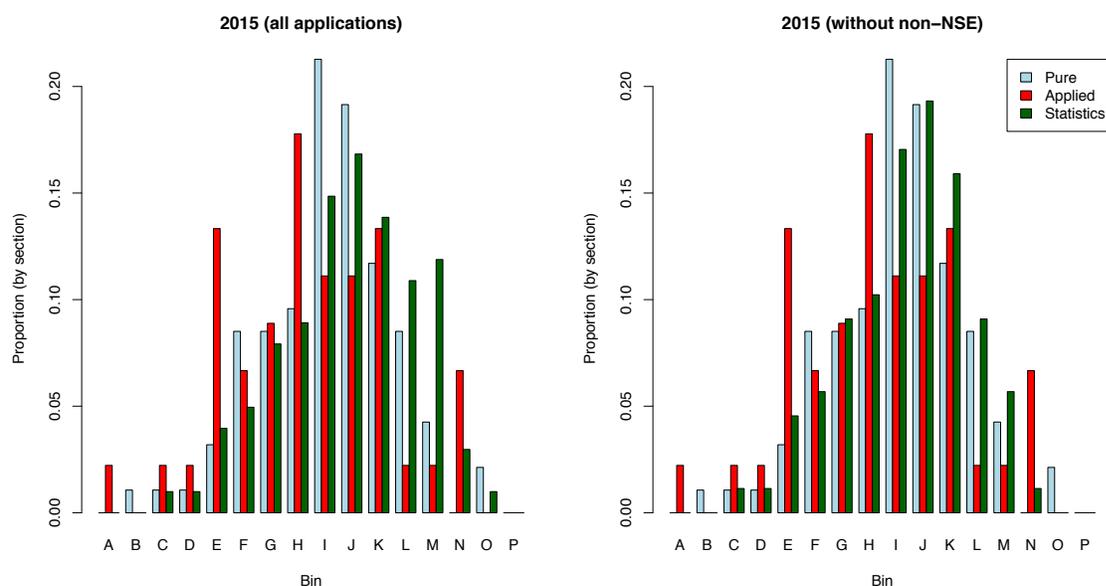


2015 are much more similar after my ad hoc correction for non-NSE applications – again, 2015 was an unusual year for applied math. There are other factors that one could correct for when comparing the bin distributions across sections, e.g., statistics tends to have more ECRs. However, I promised not to wax lyrical and, besides, the 2016 bin distribution (Figure 1) illustrates what is possible without any mitigation.

The main takeaway from my ramblings about the past three years should be that we, the statistics section, had a problem and (I think) we fixed it. Crucially, the mathematicians did not cause the problem. In fact, the group chairs (who

résultats si l'on ignore les cas non-SNG (Figure 5). Notez qu'il n'existe aucun registre officiel des cas non-SNG et je n'étais pas présent pour toutes les demandes (notamment pour des raisons de conflit d'intérêt), si bien qu'il ne s'agit là que d'une estimation.

J'insiste sur le fait que je n'utilise pas ici le terme non-SNG comme synonyme de la biostatistique. Certaines demandes non-SNG que j'ai vues avaient des applications médicales, mais pas dans la majorité des cas. En trois années au GE, je n'ai pas observé un seul cas où une demande en biostatistique a été traitée défavorablement compte tenu des règles fondamentales.



**Figure 5 : Bin distributions from the 2015 competition for all applications (left) and a best guess without non-NSE applications / Répartitions entre catégories pour le concours 2015 pour toutes les demandes (gauche) et une estimation sans les demandes non-SNG**

were mathematicians), math section chairs, and incoming math section chairs were very supportive of the various initiatives that we used or requested to resolve our problem. I understand that some may find it more comforting to blame mathematicians, or perhaps NSERC, as an alternative to introspection, but that just does not mesh with the reality of the past three years. I should also note that the various steps we took to correct our problem led to more work for the NSERC staff; yet, NSERC was consistently willing, at all relevant levels, to help us bring our bin distribution into line with the math distributions. I think this point should not be lost on the community.

In preparation for writing this piece, I reread the open letter by Christian Léger (2012). Written on behalf of the SSC by

Bien entendu, certains pourront considérer que les règles fondamentales, p. ex. le fait que le CRSNG finance la recherche en SNG, sont ipso facto injustes, mais c'est à quelqu'un d'autre d'en parler ailleurs. Il apparaît clairement dans la Figure 5 que les répartitions en maths pures et en statistique pour 2015 sont bien plus similaires après ma correction ad hoc pour les demandes non-SNG – en rappelant, une fois encore, que 2015 était une année inhabituelle en maths appliquées. On pourrait appliquer d'autres facteurs de correction pour comparer les répartitions entre sections, p. ex. le fait que la statistique a tendance à avoir plus de demandes de CDC. Cependant, j'ai promis d'éviter les envolées lyriques et la répartition entre catégories pour 2016 (Figure 1) illustre ce qui est possible sans aucune mitigation.



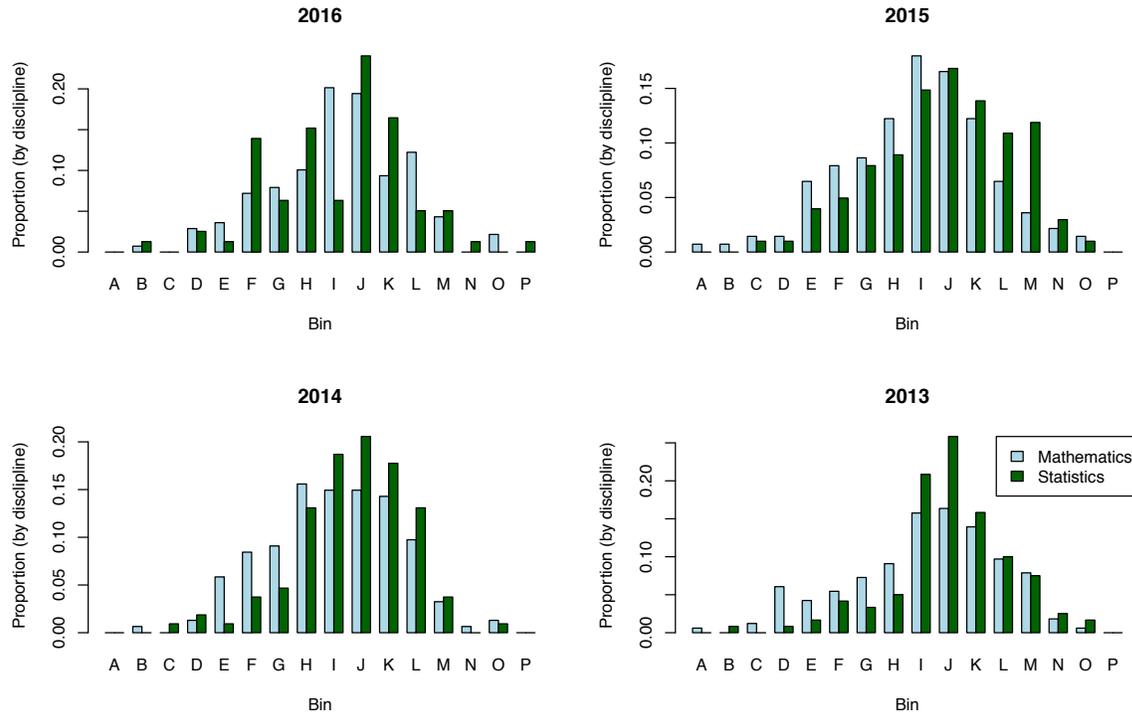
the then president, and referring to results from the 2009-2012 competitions, Léger (2012) writes:

*In each year there is disparity between the distributions for the Mathematics and the Statistics applications; on average applications in Mathematics receive a better rating than applications in Statistics.*

Bin distributions, broken down by discipline (rather than section, as in Figure 1), are provided by Léger (2012) to back up this argument, and they do so very convincingly. For completeness, the corresponding plots for 2013-2016 are provided in Figure 6; thankfully, the situation has improved substantially over the years.

Léger (2012) gives a clear account of the negative financial consequences for the statistics community arising from the 2009-2012 results. While these are real, and quite probably

La principale conclusion de mes divagations sur les trois dernières années doit être que nous, section de statistique, avons un problème et que nous l'avons (je crois) résolu. Chose essentielle, ce ne sont pas les mathématiciens qui ont causé le problème. En fait, les présidents du groupe (qui étaient des mathématiciens), les présidents des sections de maths et les futurs présidents de ces sections étaient tous très favorables aux diverses initiatives que nous avons lancées ou demandées pour résoudre notre problème. Je comprends que certains peuvent trouver réconfortant d'accuser les mathématiciens, ou peut-être le CRSNG, plutôt que de se livrer à une introspection, mais cela ne cadre pas avec la réalité des trois dernières années. Je tiens également à noter que les diverses mesures que nous avons prises pour corriger notre problème a créé plus de travail pour le personnel du CRSNG; et pourtant, le CRSNG a toujours été prêt, à tous les niveaux nécessaires, à nous aider à mettre notre répartition entre catégories en conformité avec celles des



**Figure 6 : Bin distributions, broken down by discipline, for the last four Discovery Grant competitions**  
**Répartitions entre catégories, par discipline, pour les quatre derniers concours de Subventions à la découverte**

continued into 2013 and 2014, the situation in 2015 and 2016 is different. In other words, some of the funds that might be considered “transferred” out of our community have come back in and I am optimistic that this will continue in the coming years. From reading Léger (2012) and listening

sections de maths. C’est là un point que la communauté ne devrait pas perdre de vue.

En préparation à cet article, j’ai relu la lettre ouverte de Christian Léger (2012). Rédigée pour le compte de la SSC par



to many voices within the community, one may take the view that NSERC is to blame for these ills. Based on my experiences from the past three competitions, I wonder if we should accept some portion of the blame? Could some of the energy expended criticizing NSERC have instead been focused on fixing our own bin distribution? Were exceptional proposals really entirely absent until the 2016 competition? Would we really be better off in the long run if we had been kept separate from math, e.g., as our own EG? It certainly would have made it easier for us to keep turning the spit, but to what end? A negative end, to put it mildly, I think. I could continue in this vein but I will not. I did not write this piece to praise NSERC or to criticize colleagues – the latter would, of course, entail criticizing myself, as part of the 2014 EG. I just wanted to advocate for us taking some responsibility for our own bin distributions, both past and future.

As promised, I now provide some advice for applicants and reviewers. For applicants, I strongly suggest looking at the merit indicators; they really are used by the EG, and for every case at that. I also suggest reading the instructions<sup>2</sup> and following them when preparing the application. The most common of the easily avoidable issues I have seen centre around different aspects not being fully explained in the application. Here are some examples. The roles in co-supervisions should be clear. The order of authors on publications needs to be explained. The relationship to other research support should be clarified. The quality of the contributions to HQP training should be outlined and the outcomes clarified; the HQP criterion is much more than a numbers game. Etcetera. For external reviewers, some similar advice applies: the merit indicators are helpful and there are instructions. However, I advise against using a copy-and-paste approach, e.g., from the merit indicators. I note also that a proposal itself does not need to be “groundbreaking” to merit a rating of outstanding or exceptional; rather, the language of the merit indicators speaks to leading or contributing to groundbreaking advances, which is certainly not the same thing. Above all, I implore reviewers to prepare balanced reviews. Keep in mind that a proposal is not a research paper, where some focus on the negative points is expected during peer review. Whereas a research paper generally contains complete work, a research proposal cannot – if it did, it would not be a proposal.

I feel upbeat about our prospects, vis-à-vis NSERC Discovery Grant funding, in the short term. With the current committee membership and momentum, I am optimistic about our bin distributions for the next couple of competitions. Beyond that, however, it would be quite easy for us to regress; this would be nothing short of tragic and we should do everything we can to make sure we stay on course. With this in mind,

<sup>2</sup> [http://www.nserc-crsng.gc.ca/ResearchPortal-PortailDeRecherche/Instructions-Instructions/DG-SD\\_eng.asp](http://www.nserc-crsng.gc.ca/ResearchPortal-PortailDeRecherche/Instructions-Instructions/DG-SD_eng.asp)

son président de l'époque, et en référence aux résultats des concours 2009-2012, Léger (2012) écrit :

*À chaque année, il y a un écart entre la répartition des résultats pour mathématique et celle pour statistique; en moyenne les demandes en mathématiques reçoivent de meilleures évaluations que celles en statistique.*

Les répartitions entre catégories, par discipline (plutôt que par section, comme dans la Figure 1), sont fournies par Léger (2012) pour appuyer cet argument, et elles le font de manière convaincante. À des fins d'exhaustivité, les chiffres correspondants pour 2013-2016 sont indiqués en Figure 6; heureusement, la situation s'est nettement améliorée au fil des années.

Léger (2012) donne un compte-rendu clair des conséquences financières négatives pour la communauté statistique des résultats de 2009-2012. Même si celles-ci sont réelles et qu'elles ont probablement perduré en 2013 et 2014, la situation en 2015 et 2016 est différente. En d'autres termes, certains des fonds qui pouvaient avoir été « transférés » hors de notre communauté lui ont été rendus et je suis optimiste que cela continuera dans les années à venir. À lire Léger (2012) et à écouter de nombreuses voix dans la communauté, on pourrait croire que le CRSNG est responsable de ces maux. Sur la base de mes expériences des trois derniers concours, je me demande si nous devrions accepter une part de responsabilité. Une partie de l'énergie dépensée à critiquer le CRSNG n'aurait-elle pas été mieux employée à résoudre notre propre problème de répartition? N'y avait-il réellement aucune demande exceptionnelle jusqu'au concours 2016? Aurions-nous vraiment été en meilleure posture à long terme si nous avions été séparés des maths dans notre propre GE? Cela nous aurait certainement aidé à continuer à tourner la broche, mais à quelle fin? Une fin négative, pour dire le moins, je pense. Je pourrais continuer dans cette veine, mais je ne le ferai pas. Je n'ai pas rédigé cet article pour glorifier le CRSNG ni pour critiquer les collègues – dans ce dernier cas je devrais, bien entendu, me critiquer moi-même pour avoir participé au GE 2014. Je voulais simplement défendre le fait que nous ayons assumé une certaine responsabilité concernant nos répartitions entre catégories, à la fois hier et demain.

Comme promis, je vais maintenant proposer quelques conseils à l'intention des candidats et des évaluateurs. Aux candidats, je suggère vivement d'étudier les indicateurs de mérite : le GE s'en sert vraiment pour toutes les demandes. Je leur suggère aussi de lire les instructions<sup>2</sup> et de les suivre pour préparer leur demande. L'erreur la plus commune et facilement évitable consiste à ne pas pleinement expliquer certains aspects de la demande. En voici quelques exemples. Les rôles des co-superviseurs doivent être clairement présentés. L'ordre des auteurs des publications doit être expliqué. Les relations avec d'autres sources de

<sup>2</sup> [http://www.nserc-crsng.gc.ca/ResearchPortal-PortailDeRecherche/Instructions-Instructions/DG-SD\\_fra.asp](http://www.nserc-crsng.gc.ca/ResearchPortal-PortailDeRecherche/Instructions-Instructions/DG-SD_fra.asp)



I encourage people to say “yes”, if at all possible, when approached to join the EG. It is a big commitment but it is so important for the community. As section chair in 2015 and 2016, I was surprised at how various constraints, when added together, can make it very tricky to fill vacancies. If the changes to our bin distribution in recent years tell us anything, it is that a small group of people can make a big difference.

### Reference

Léger (2012), “The Statistical Society of Canada (SSC) response to the NSERC Consultation on the Evaluation of the Discovery Grants Program”, *Liaison* 27(2), 12-21.

### Acknowledgement

I am grateful to several colleagues and friends, from across Canada, for their valuable comments and suggestions. I am grateful to NSERC for providing the data I used to generate the figures herein and for doing so with only the minor proviso that within-section proportions be used rather than raw numbers.

*Paul McNicholas was a member of EG 1508 for the 2014, 2015, and 2016 Discovery Grant competitions. He was statistics section chair for the 2015 and 2016 competitions, and incoming statistics section chair for the 2014 competition. He holds the Canada Research Chair in Computational Statistics at McMaster University, where he is a Professor in the Department of Mathematics and Statistics.*



*Paul McNicholas a été membre du GÉ 1508 pour les concours des Subventions à la découverte 2014, 2015 et 2016. Il a présidé la section de statistiques lors des concours 2015 et 2016 après en avoir été le président élu en 2014. Il est titulaire de la Chaire de recherche du Canada en statistiques computationnelles à McMaster University, où il est professeur au Département de mathématiques et de statistique.*

financement doivent être clarifiées. La qualité des contributions à la formation du PHQ doit être exposée et les résultats clarifiés : le critère du PHQ n'est pas simplement quantitatif, loin de là. Etcetera. Aux évaluateurs externes, certains conseils similaires s'appliquent : les indicateurs de mérite sont utiles et des instructions existent pour vous guider. Toutefois, évitez d'utiliser une approche « copier-coller » des indicateurs de mérite. Notez aussi que la demande elle-même n'a pas à être « originale et novatrice » pour mériter une cote « Remarquable » ou « Exceptionnel »; le libellé des indicateurs de mérite parle plutôt de mener ou de contribuer à des percées importantes, ce qui n'est certainement pas la même chose. Par dessus tout, j'implore les évaluateurs de préparer des évaluations équilibrées. Gardez à l'esprit qu'une demande n'est pas un article de recherche, pour lequel les évaluateurs se doivent de relever les points négatifs. Un article de recherche contient généralement des travaux finalisés, ce qui n'est pas le cas pour une proposition de recherche, sans quoi il ne s'agirait pas d'une proposition.

Je suis optimiste quant à nos perspectives à court terme en ce qui concerne le financement des Subventions à la découverte par le CRSNG. Avec nos membres et notre élan actuels, je suis optimiste quant à nos répartitions entre catégories pour les prochains concours. À plus long terme, cependant, il serait facile pour nous de retomber dans de mauvaises habitudes; ce serait tragique et nous devons tout faire pour garder le cap. Dans cette optique, j'encourage les gens, dans la mesure du possible, à dire « oui » si l'on vous demande de rejoindre le GE. C'est un engagement majeur, mais c'est si important pour la communauté. En tant que président de la section en 2015 et 2016, j'ai été surpris de voir combien les diverses restrictions, une fois combinées, compliquent le fait de combler les postes vacants. Si l'évolution de nos répartitions ces dernières années peut nous apprendre quelque chose, c'est bien qu'un petit groupe de personnes peut faire toute la différence.

### Référence

Léger (2012), « Réponse de la Société statistique du Canada (SSC) à la consultation du CRSNG concernant le programme de subventions à la découverte », *Liaison* 27(2), 12-21.

### Remerciements

Je suis très reconnaissant à mes collègues et amis de tout le Canada pour leurs précieux commentaires et suggestions. Je remercie le CRSNG de m'avoir fourni les données utilisées ici, en me demandant simplement de ne publier que les proportions par section plutôt que les données brutes.