Rankings of Canadian Universities, 2000: Buyer Beware

Stewart Page
Ken Cramer

We present a critical perspective on the rankings of Canadian universities presented in *Maclean's* magazine, November 20, 2000. Data from several statistical analyses, including cluster analysis, show that in many cases universities actually resemble each other in a manner different from their classification or final ranking by *Maclean's*, and that the magazine's classifications are often decided by variables not realistically measurable by or amenable to students. We summarize several pitfalls in the ranking procedures for the 2000 data, and in ranking exercises more generally. In their present form, *Maclean's* data cannot be logically or empirically useful to students.

Les auteurs critiquent le classement des universités canadiennes présenté dans le *Maclean's* le 20 novembre 2000. Les données tirées de plusieurs analyses statistiques, dont l'analyse typologique, indiquent que, dans bien des cas, les universités se ressemblent sur d'autres fronts que ce que laisse supposer leur position dans le palmarès. Les auteurs résument les écueils méthodologiques liés au classement pour les données de 2000. Dans leur forme actuelle, les classifications du magazine reposent sur des variables qui ne sont ni empiriquement mesurables ni accessibles aux étudiants.

Maclean's is a major Canadian mass circulation magazine that emphasizes Canadian and, secondarily, North American content. Its content in general and its circulation are similar to the content and audience of *Time* and *Newsweek*. In its November 20, 2000 issue, entitled "Measuring Excellence," *Maclean's* published its ninth annual rankings (described fully in that issue pp. 62ff) that readers can use to "take the measure" of Canadian universities. This approach is similar to that used by publications such as *Consumer Reports*, which ranks goods and services and then publishes overall standings based on these data. Perhaps because of the increased popularity of notions such as cost effectiveness, efficiency, and value for one's "educational dollar," the creation of ranks or ratings of universities, based on a variety of supposed performance indicators, has become an increasingly popular method with which to assess higher education throughout North America (Bruneau & Savage, 2001). As reported in the *Toronto Globe and*

Mail of January 22, 2001 (Bertin, 2001), for example, the *Financial Times* of London recently used these procedures for its recent international rankings of business (MBA) schools.

CANADIAN UNIVERSITIES' RANKING DATA, 2000

For 2000, *Maclean's* classified universities into three main types: Medical / Doctoral universities ($N = 15$), which have medical schools and large graduate departments; Comprehensive universities ($N = 11$), which do not have medical schools but may have graduate departments at the master's or doctoral level; and Primarily Undergraduate universities ($N = 21$), which generally do not have graduate programs.

Following procedures we used in previous research (e.g., Page, 1998, 2000; Page, & Cramer, 2001; Page, Cramer & Page, 2001), we compiled the *Maclean's* 2000 data according to its six main measures, each composed of several indices (as detailed in *Part 5* of the "Results" section on page 297). Summed over the six measures for the 2000 data, *Maclean's* used a total of 22 indices for Medical / Doctoral universities, 21 for Comprehensive universities, and 20 for Primarily Undergraduate universities. *Maclean's* converts preliminary (raw) data, across several indices, to ranks (first, second, etc.), then constructs from these an overall rank for each university. From these, a rank ordering is constructed for the universities within each type.

In previous studies, we examined the *Maclean's* ranking parameters and data to assess the degree to which the parameters were correlated with each other and with the overall final rankings assigned. We also examined the extent to which higher-ranking universities were reliably different, in terms of the *Maclean's* parameters, from lower-ranking universities. In those studies we invariably found, through analyses of the magazine's published data, that the *Maclean's* measures and component indices were not internally consistent and not reliably related either to each other or to final ranks (Page, 1998, 2000; Page & Cramer, 2001; Page, Cramer, & Page, 2001). In homespun terms, our previous studies showed that the notion of rank has been unreliably correlated with easy labels about "good" or "bad" universities, as these terms are often used in common parlance.

In this article, we report on several statistical and other aspects of the 2000 data, as described specifically in each section below. In our opinion, especially in light of previous studies and of the increased marketing and use of rank-based data (Bruneau & Savage, 2001), it is important to examine and update the statistical aspects of published data such as those produced and circulated by *Maclean's*. Examining this type of data is also important in terms of evaluating claims that rank-ordering procedures will assist students in selecting a university.

RESULTS

We divided the results for the 2000 *Maclean's* data into five parts, as reported in sequence below.[1] Part 1 concerned the extent to which the indices underlying the ranking data predicted the final ranking results. Part 2 concerned the extent to which higher-ranking ("better") universities were empirically different from lower-ranking universities. Part 3 concerned the extent to which the ranking parameters (indices) were discriminably different across the three major *Maclean's* university classifications. Part 4 concerned the extent to which the indices used were internally consistent among themselves, that is, to what extent a student or other observer could assume that the measurement indices would be consistent in their description and portrayal of a "good" (higher-ranking) or "bad" (lower-ranking) university. Part 5 used cluster analysis to identify groupings (clusters) into which highly similar universities could be placed and from which dissimilar ones could be excluded. Another purpose was to identify the extent to which universities might be clustered differently in comparison with their grouping by *Maclean's* into its three main university categories.

Part 1. For the 2000 data, many indices used in the six main measures were unrelated to the final rankings by *Maclean's.* We computed Spearman *rho* correlations between each university's final rank, as derived by *Maclean's*, and its rank on each of the indices comprising the six main measures. The proportion of indices significantly related (at $p < .05$) to final rank, collapsing over the three university types, was .45.[2] We note that with an alpha level set at $p < .05$ for these, as well as for other sets of correlations computed in the present data, approximately 5% can be expected to be significant by chance. These results showed generally that most of the *Maclean's* indices were not strongly or significantly related to final university rankings.

For Medical/Doctoral universities, 8 of 21 indices (38%) correlated significantly to final rank. For Comprehensive universities, 10 of 21 (47%) correlated significantly to final rank. For Undergraduate universities, 10 of 20 (50%) correlated significantly to final rank. Regardless of university type, mean (average) ranks on the six main measures were also not strongly related to final rankings.

Part 2. As in previous studies, we ascertained the extent to which lower-ranking universities differed from higher-ranking ones in terms of the indices used. To do so, we examined and compared the top and bottom subgroups (halves) of the universities within each type, using the Wilcoxon Rank Sum test. This test, which may be interpreted similarly to the point-biserial correlation or Mann-Whitney U-test, examines the significance of differences in ranked data on a specified parameter taken from two

independent samples of subjects (in this case, universities). In this case, we thus compared universities by assessing the extent to which the rank scores of indices for the top differed significantly from those of the bottom half.

For Medical / Doctoral universities, the Wilcoxon tests showed that the top and bottom subgroups (halves) differed significantly, at p (z) < .05, on 12 of the 22 (54%) individual indices.[3] For Comprehensive universities, the top and bottom halves differed significantly on 12 of 21 (57%) of these indices. For Undergraduate universities, the top and bottom halves differed significantly on 7 of 20 (35%) indices. Thus, collapsing over the three university types, the top and bottom halves did not differ significantly, in average rank, on 32 of 63 (51%) of the individual indices. These results showed that for many indices, higher-ranking or better universities did not differ much or at all from those of lower rank, and vice versa.

Part 3. After converting the raw (pre-rank) data to standard scores, we carried out an exploratory discriminant function analysis with the 2000 data, using *Maclean's'* indices as potential predictors and its three-category university classification scheme as the classification variable. Only five (approximately 25%) indices were required to achieve significant discrimination between the university categories, using the criterion of minimization of Wilks' lambda statistic. In decreasing order of discriminating power, these were: size of medical science grants, size of first-year classes in the range of 1–25 students per class, percentage of expenditures devoted to students, size of first-year classes in the range of 251–500 students per class, and number of medical science grants. The discriminant function, composed of the five predictors, correctly classified 93% of the universities into the three categories used by *Maclean's.* The results from the discriminant analysis also showed that most of the *Maclean's* indices discriminated weakly and nonsignificantly among the universities; further, there was some indication in the data that Medical / Doctoral schools were most likely to be associated with the more important and discriminating indices.

Part 4. As in previous studies, we found that the level of (Spearman *rho*) intercorrelation between the indices within each of the six measures was generally low. Moreover, some of the relatively few significant relationships and correlations among the measures, indices, and final overall rankings were negative in direction. This finding means, for example, that better (higher-ranking) universities actually do relatively poorly on certain indicators, and vice versa for worse (lower-ranking) universities. Similarly, some universities rank poorly on various measures or indices, yet score in the average range, or higher, on one or more of the other indices or measures, or in the final ranking. These results showed generally that the patterns of intercorrelations and interrelationships between indices, regardless of the more general measure of which they were a part, were unreliable and

inconsistent. The finding that some of the intercorrelations between indices were negative was, again, a cause of additional difficulty in interpretation and in making logical decisions among specific universities in such cases.

To further assess reliability of the individual indices, we calculated Cronbach's (1951) alpha statistic (a measure based on the average level of intercorrelation between indices) separately for each university type. Alpha for Medical / Doctoral universities was .802; for Comprehensive universities, .581; for Undergraduate universities, .714. Generally, an alpha value of .800 or greater is recommended as an acceptable level of internal consistency or reliability (for example, in the case of psychological test items or similar indices; see Carmines & Zeller, 1983). Clearly, the indices reach a minimally adequate level of internal consistency in the case of Medical / Doctoral universities, but not in the case of the other two types. Although the specific *Maclean's* indices are nominally the same across the three university types, they appear to elicit different patterns of interrelationships and interpretations depending on the category of university under consideration. Such differences tend to be most pronounced when comparing Medical / Doctoral universities with the two other types.

Part 5. Finally, and from a different perspective, we used cluster analysis (Everitt, 1993; Gordon, 1987) to examine patterns of interrelationship among the universities (see Table 1 and Table 2), both within and across the three university types set out by *Maclean's*. This analysis identified unique families or clusters of schools, in which the similarity of each member's corresponding profile was maximized, and intercorrelations among members were high. This procedure thus identified and grouped together schools of similar profiles using the *Maclean's* indices, so that members were considered highly similar within a cluster, but clearly different from members outside that cluster. The primary clusters identified are shown in Table 1.

Using algorithms outlined by Ward (1963), the cluster procedure calculated the squared Euclidian distances (as estimates of distance among schools) for the 47 schools, based on their raw scores for the indices (22 for Medical / Doctoral, 21 for Comprehensive, 20 for Undergraduate universities). These indices, based on proportions where appropriate, were grouped into six main measures by *Maclean's*: (1) Student Body – average entering grade, number (proportion) of out-of-province students in first year, proportion of students with entering grade of 75% or higher, proportion of students who graduate, number of international (graduate) students, and number of student awards; (2) Classes – class sizes in the first and second years, class sizes in the third and fourth years, and percentage of classes taught by tenured faculty; (3) Faculty – proportion of faculty with doctoral degrees, number of awards per full-time faculty member,

number of Social Science and Humanities grants, and number of medical and science grants; (4) Finances – size of operating budget, percentage of budget allotted to scholarships and bursaries, and percentage of budget spent on student services; (5) Library – total holdings, holdings per student, number of acquisitions per year, and expenses; (6) Reputation – frequency of alumni support, and results of a reputational survey that *Maclean's* sent to senior university officials, high-school guidance counsellors, chief executive officers of Canadian corporations, corporate recruiters, and heads of organizations.

Even for the cluster analyses within the three *Maclean's* categories, subclusters were identified. For Medical/Doctoral schools, subcluster 1 contained Calgary, Dalhousie, Laval, Manitoba, Saskatchewan, and Sherbrooke; subcluster 2 contained Alberta, McMaster, Montreal, Ottawa, and Western Ontario; subcluster 3 contained British Columbia, McGill, Queen's, and Toronto.

For Comprehensive schools, subcluster 1 contained Waterloo alone; subcluster 2 contained Concordia, Memorial, New Brunswick, and Regina; subcluster 3 contained Guelph, Simon Fraser, Victoria, Windsor, and York.

For Undergraduate schools, subcluster 1 contained Acadia, Bishop's, Mount Allison, Moncton, and St. Francis Xavier; subcluster 2 contained Brandon, University College of Cape Breton, Laurentian, Nippissing,

TABLE 1

Similar Universities in Three Primary Clusters, as Generated by Cluster Analysis

Cluster	*Universities (ordered alphabetically)*
1	Calgary, Carlton, Concordia, Lakehead, Manitoba, Memorial, Moncton, New Brunswick, Ryerson, Saskatchewan, Sherbrooke, Wilfrid Laurier, Windsor
2	Acadia, Bishop's, Brandon, Brock, University College of Cape Breton, Laurentian, Lethbridge, Mount Allison, Mount St. Vincent, Nipissing, Prince Edward Island, Regina, St. Francis Xavier, St. Mary's, Trent, Winnipeg
3	Alberta, British Columbia, Dalhousie, Guelph, Laval, McGill, McMaster, Ottawa, Queen's, Simon Fraser, Toronto, Victoria, Waterloo, Western Ontario, York

TABLE 2

Universities (13) Identified in Cluster Analysis as Dissimilar to Their Groupings by Maclean's

University (ordered alphabetically)	*Cluster Analysis Classification*	*Grouping by* Maclean's
Calgary	Comprehensive	Medical / Doctoral
Guelph	Medical / Doctoral	Comprehensive
Lakehead	Comprehensive	Undergraduate
Laurier	Comprehensive	Undergraduate
Manitoba	Comprehensive	Medical / Doctoral
Regina	Undergraduate	Comprehensive
Ryerson	Comprehensive	Undergraduate
Saskatchewan	Comprehensive	Medical / Doctoral
Sherbrooke	Comprehensive	Medical / Doctoral
Simon Fraser	Medical / Doctoral	Comprehensive
Victoria	Medical / Doctoral	Comprehensive
Waterloo	Medical / Doctoral	Comprehensive
York	Medical / Doctoral	Comprehensive

Prince Edward Island, and Trent; subcluster 3 contained Brock, Lakehead, Lethbridge, Mount Saint Vincent, Northern British Columbia, Ryerson, Saint Mary's, St. Thomas, Wilfrid Laurier, and Winnipeg.

In essence, the cluster analysis, using the universities' pattern of rank scores on the various indices, showed that the universities could be grouped into different sets, each containing members with profiles similar to those of others in the same set and clearly different from those of members in other sets. As outlined above, inspection of the cluster pattern showed that in several cases (shown in Table 2) universities were seen to belong to clusters inconsistent with their membership in one of the three *Maclean's* university categories. In these cases, then, one may distinguish between a priori or conceptual similarity, as defined by *Maclean's*, and empirical similarity, as defined by statistical analysis of the ranking data themselves. In most cases, also, the pattern of relationships within and between clusters was not clearly reflective of rank differences (as defined by *Maclean's*) between higher- and lower-ranking universities within the

three university types. Both cluster and discriminant analyses also showed that the indices most strongly differentiating membership in the three university classifications and clusters were often those concerning parameters of debatable relevance to the matter of university choice as a pragmatic issue for students. For example, although indices concerning student services may be relevant as part of this issue, we would question whether students newly graduated from high school are in a position to judge or even be aware of other indices such as library holdings, or faculty grants, as supposed measures of qualitative variation among universities. Again, the presence of several universities within each of the identified clusters is a reminder that the schools within any one cluster show considerable variation in terms of their final rank as assigned by *Maclean's*. Moreover, their pattern of rankings on the various measures and indices show them to be far less different than the final rankings by *Maclean's* imply – thus rendering even more difficult the task of choosing among them.

DISCUSSION

The published ranking data in *Maclean's* show weak and inconsistent interrelationships both among and between individual measures and indices as well as between these parameters and final rankings. These data thus provide an unreliable guide for students attempting to use them as indicators or means by which to choose a university. We note from these results that a measure or indicator, though perhaps bearing face validity, can be conceptually misguided and possibly completely invalid – yet still be accurately measured and calibrated, and painstakingly reported. That is, such an indicator may indeed have high reliability in terms of accuracy of measurement, but little validity either conceptually or empirically. Interpretation of the 2000 rankings is thus subject to numerous pitfalls, including: low percentages of significant correlations, lack of differentiation among universities on the six measures and component indices, low correlations between indices and final ranking, and vulnerability to known problems in making reliable comparisons or quantitative statements with statistical data reaching the level only of an ordinal – that is, rank-order, scale (Siegel, 1959). In many cases, therefore, the magazine's assumptions about components (indices or measures) that should logically or theoretically be correlated, or that should appear together in defining a good (or bad) university, were seldom supported by the data. The dilemma for students is worsened by the fact that, for a given university, the ranking parameters clearly present them with several reasons to attend it – but also not to attend it. Moreover, a given university will be either above or below others on one or more, usually many, parameters.

From another perspective, concerning predictive or criterion validity, we have found in our previous studies that in selecting a university, students seldom use either published ranking data or the specific *Maclean's* indices or measures underlying the published ranking data. Students instead cite personal factors, such as location, their own finances, and where their friends are attending university, as the most important in university selection (e.g., Alexitch, 1994, 1997; Alexitch & Page, in press; Stanga, 2000).

But, apart from concerns relating to the statistical reliability, validity, and consistency of ranking exercises such as those of *Maclean's*, many other issues remain that cannot be resolved or embraced by hard data or statistical analysis. One of many possible results of the rankings is their unintended effects upon the quality of a university's academic and intellectual spirit as these are experienced and perceived by students. This raises the possibility that the rankings may help to generate yet another form of the educational self-fulfilling prophecy at the postsecondary level (Page & Rosenthal, 1990; Rosenthal & Jacobson, 1968), this time one that will affect some students adversely and others positively. In this sense, the financially driven publication and annual national circulation of rankings may not have as major priorities the academic or intellectual welfare of many university participants. Apart from its status as a commodity or product to be advertised and marketed, students' university affiliation is part of their self-definition, self-esteem, identity, and general sense of well being. Accordingly, the academic performance and general orientation of students attending lower-ranking universities may be affected negatively by continued exposure and references to ranking exercises and data. Each year, for example, these students are publicly informed by *Maclean's* (and here using the terminology of *Maclean's* in each case) that their universities are considered less innovative, worse (as opposed to better) overall, of lower overall quality, less likely to generate "leaders of tomorrow" ("Measuring Excellence," 2000, p. 67) and so on. Moreover, *Maclean's* has in the past referred to the issue of the "last-chance" university.

We note too that in the 2000 data the overall rankings were significantly correlated (at $p < .05$) with both of the reputational indices, that is, the universities' level of alumni support and their rank in *Maclean's* reputational survey. Medical/Doctoral universities generally were ranked as having more positive reputations than Comprehensive and Undergraduate universities. Also, although the seriousness of the problem cannot be determined, the *Maclean's* 2000 reputation survey elicited poor response (return) rates – that is, an average of 9.80% for persons other than university officials (i.e., guidance counsellors, CEOs, corporate recruiters, and heads of organizations). The return rate for university officials was much higher, at 44.51%. This is unsurprising, especially because in the past,

Maclean's has imposed ranking penalties upon institutions not cooperating or returning requested data for its indices.

With regard to data in the *Maclean's* 2001 rankings (see the November 19, 2001 issue), there are no substantive differences compared to the conclusions reported herein. Using Spearman *rho* (rank-based) correlations, we find, for example, that again most of the indices are not significantly or reliably related either to each other or to final rankings. For Undergraduate schools, for example, 11% of all possible intercorrelations between indices were significant, at $p < .05$. For Undergraduate schools, considering only the indices belonging to the same measure (that is, within-measure correlations), 8 of 31 (26%) were significant. The data show again that, in general, all universities feature both positive and negative indicators, regardless of final ranking. Higher ranking schools thus frequently have mediocre or poor scores (ranks) on several or many indices, and vice versa for lower ranking schools. For Comprehensive schools, for example, 5 of the 22 comparison indices used (22%) were significantly related to final rankings.

We find also that the bottom versus top halves of universities within a given category do not show signifcant diferences in many indices used as indicators of university quality. For example, considering the 22 indices upon which Canadian Comprehensive universities were compared, those schools with final ranks in the upper half, versus those in the lower half, were significantly different on only 4 out of 22 indices (18%). Pooling the three university types, 23 of 66 indices (34%) in these comparisons showed significant differences. The data show again that most universities, even contrasting those in the top versus bottom halves, are not particularly different. For prospective students, this means again that there are about as many reasons (represented by the indicators or indices) to not attend a given university as there are to attend it.

CONCLUSION

One may ask why *Maclean's* continues to emphasize these parameters when their interpretation and interrelationships may be shown by simple statistics to be unreliable if not totally misleading. As one answer, many colleagues inform us that publications might engage in ranking exercises as a form of advertising or marketing. Such an answer constitutes a type of editorial hypothesis that we do not pretend to have evaluated herein. This view effectively claims that consumers have become willing to pay current prices to obtain what is marketed as necessary information, as consistent with the popularized goals of cost-effectiveness and objective decision making. In this light, counselling assistance to students appears to be in

danger of being dominated by linear and data-driven information such as input-based, albeit easily quantifiable, performance "indicators." So far, output-based and less easily quantifiable measures that would have enhanced the content validity and generalizability of the evaluative indicators selected, such as measures of student satisfaction, students' assessments and statements about their university, or measures of morale, sense of well-being, or level of post-university achievement, have not been used. Moreover, other measures assessing the unique missions, different populations, and different educational situations of particular universities have played little role in the *Maclean's* ranking exercises. As a single example, Carleton features one of Canada's best schools of journalism, yet the university's overall rank in the 2000 data was low – eighth out of 11 Comprehensive universities. We are left with an important question for future research and debate: How is the quality, as distinguished from the quantitative aspects, of a university to be unequivocally determined? Some further aspects of this issue have been informatively described and elaborated by Bruneau and Savage (2001).

Some aspects of *Maclean's*' recent coverage of universities we view as useful and informative; for example, in its "Annual Guides" to Canadian universities, factual and descriptive information, necessarily unique to each school, is summarized separately from the matter of comparative rankings. Yet, if ranking exercises become even further entrenched, future researchers will need to find out what happens to students known to have attended universities publicly labelled as better and those publicly labelled as worse. Such research would seem similar to that concerning the variable of social class as a determinant of educational expectations about students' success and intellectual potential (Rhem, 1999). Further, because rankings are related to the comparative level of financial and other resources across different universities, we would predict that the basic pattern of rankings is unlikely to change significantly over time. This factor may itself represent a type of self-fulfilling mechanism. The Ontario government recently decided to base future funding to universities on indicators such as graduation rates and students' employment success after graduation. Such imposed criteria encourage universities to compete as if they were minor league baseball teams, with license to define themselves as educationally effective if they can portray themselves as scoring well on the government's indicators. It is unfortunate that the ranking data, fuelled by *Maclean's*' editorial metaphor of competition among schools, continue to show low levels of reliable and meaningful discrimination among universities, and tend to use parameters neither familiar nor, in our view, realistically helpful to students. We believe that the findings presented here should be of concern to educators, parents, and students, at all educational levels.

NOTES

1. In our previous research, and in analysis of the *Maclean's* 2000 data reported here, our goal was not to test hypotheses or any predictive model, nor to take a position on the potential value or nature of university ranking exercises that may be developed in the future. Our goal here is to examine the 2000 data in light of previous analyses, and to comment only to the extent we believed warranted by the data. We use short names for universities, generally following the names used by *Maclean's*.

   To maintain readability and conserve space, the exact statistical *p* values from the large number of statistical tests computed in the present study are not included here. Additional details or information are available upon request from the authors, in the Department of Psychology, University of Windsor.

2. For the 1999 *Maclean's* ranking data (see Page, Cramer, & Page, 2001), the proportion of significantly related indices was .56. For the 1998 ranking data, this figure was .26.

3. *P (z)* refers to the z statistic generated by the Wilcoxon test.

REFERENCES

Alexitch, L. (1994). *Undergraduate expectations and perceptions of a university education in the 1990s.* Unpublished doctoral dissertation, University of Windsor, Ontario.

Alexitch, L. (1997). Students' educational orientation and preferences of advising from university professors. *Journal of College Student Development, 38,* 333–343.

Alexitch, L., & Page, S. (2001). Students' academic choices and expectations: Influence of gender, educational orientation, and life values. *Guidance and Counselling, 17,* 8–16.

Bertin, Oliver. (2001, January 22). Ivey MBA tops in Canada. *Toronto Globe and Mail,* p. B1.

Bruneau, W., & Savage, D. (2001). *Counting out the scholars.* Halifax, NS: James Lorimer & Company.

Carmines, E., & Zeller, R. (1983). *Reliability and validity assessment.* Beverley Hills, CA: Sage Publications.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Everitt, B. S. (1993). *Cluster analysis* (3rd ed.). New York: Wiley.

Gordon, A. D. (1987). A review of hierarchical classification. *Journal of the Royal Statistical Society, 150,* 119–137.

Measuring Excellence [Special issue]. (2000, November 20). *Maclean's, 113*(46).

Page, S. (1998). Rankings of Canadian universities: Statistical contrivance versus help to students. *Canadian Journal of Education, 23,* 452–460.

Page, S. (2000). Rankings of Canadian universities: A new marketing tool. *Journal of Marketing for Higher Education, 10,* 59–69.

Page, S., & Cramer, K. (2001). *Maclean's* rankings of health care indices in Canadian communities, 2000: Comparisons and statistical contrivance. *Canadian Journal of Public Health, 92,* 295–300.

Page, S., Cramer, K., & Page, L. (2001). Rankings of Canadian universities: A magazine's marketing tool. *Guidance and Counselling, 16,* 51–57.

Page, S., & Rosenthal, R. (1990). Sex and expectations of teachers and sex and race of students as determinants of teaching behavior and student performance. *Journal of School Psychology, 28,* 119–131.

Rhem, J. (1999). Pygmalion in the classroom. *National Teaching and Learning Forum, 8*(2), 1–2. Retrieved October 3, 2002 from http://www.ntlf.com/html/pi/9902/pygm_1.htm

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart, & Winston.

Siegel, S. (1959). *Nonparametric statistics.* New York: McGraw Hill.

Stanga, R. (2000). *Who chooses Windsor? A profile of student selection.* Unpublished master's thesis, University of Windsor, Ontario.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236–244.