



What are we really evaluating when we rank journals : Comparisons of views

AUTHOR(S)

Michael Polonsky, P Whitelaw

PUBLICATION DATE

01-01-2003

HANDLE

[10536/DRO/DU:30016354](https://hdl.handle.net/10536/DRO/DU:30016354)

Downloaded from Deakin University's Figshare repository

Deakin University CRICOS Provider Code: 00113B

WHAT ARE WE REALLY EVALUATING WHEN WE RANK JOURNALS - COMPARISONS OF VIEWS.

Michael Jay Polonsky and Paul Whitelaw
Victoria University

Track: Marketing Education

Abstract

This paper examines differences in academics perceptions of how journals should be evaluated in terms of their prestige, contribution to theory, contribution to practice and contribution to teaching. Comparisons are made between individual and institutional weightings, regional variations and whether an individual works at an institution offering a PhD/DBA. Some differences were identified, suggesting that that evaluative criteria used to rank journal may be influenced by employment situations.

Introduction

Within the marketing literature there have been numerous attempts to evaluate, rank or rate journals. Hawes and Keiller (2003) identified that between 1980 and 1997 there were at least 10 such attempts published in journals. Their work did not include the various conference papers on this topic. For example, works by Baumgartner and Pieters (2003), Theohartkas and Hirst (2002) or Mort, et al. (2002) were not included.

The ranking of journals has been examined within a range of disciplines and has traditionally taken two difference approaches – Citation based evaluations and Perceptual evaluations. Authors such as Baumgter and Peiters (2003) have used citation analysis to examine the degree to which one journal is referred to in others. The Social Science Citation (SSC) index also uses this approach, although many marketing journals are not included in the SSC (Baumgter and Peiters 2003). The examination of citations is “objective”. It may, however, have some inherent biases, such as; lagged effects (Jobber and Simpson 1988), databases only include limited “selected” journals (Neway and Lancaster 1983), and journals from some regions are under-represented (Nobes 1995).

Perceptual evaluations are the second approach used to examining journal rankings (Brown and Becker 1991, Luke and Doke 1987). Various factors have been found to moderate perceptual rankings. These include; institutional and individual demographics (Hult et al. 1997), the objective/focus of the ranking (Polonsky and Waller 1993), regional variations (Danielsen and Delorme 1976, Theohartkas and Hirst 2002), and the journal’s focus (Danielsen and Delorme 1976, Hawkins et al. 1973).

The rankings generated across marketing studies have been found to differ (Hawes and Keiller 2002). This might be expected, especially if the criteria used for rankings differ (Hawes and Keiller 2002; Polonsky and Waller 1993; Polonsky et al. 1999). From a research perspective, it is *very curious* that most ranking studies rely on single item constructs (perceptions or citations). Baumgartner and Peiters (2003 p125) suggest that single item measures such as perceptions allows researchers to “capture the multi-faceted construct” to rank journals. However, the composition of this “perceptual” construct is never articulated or examined. If the importance of evaluative criteria differs between participants, it would result

in two individuals ranking a given journal or set of journals differently. Thus, understanding the underlying perceptual criteria used by individuals is important.

While journal rankings may be important, the focus of this paper is *not* to examine journal rankings, but rather to examine the underlying criteria that might be used to rank journals. The work also examines how these evaluative criteria vary according to the moderating factors of: whose criteria are used for ranking (individual or their institution); whether the institution is research focused (offers PhD/DBA); and whether there are regional differences (Americas, Europe, Asia-Pacific) in views.

Methodology

The study involved three stages. First, an examination of the multi-disciplinary literature on journal rankings was examined. This identified that there have been few attempts to empirically measure whether rankings differ based on the underlying evaluative criteria (ANBAR 2002; Polonsky and Waller 1993), even though there is a suggestion that journal rankings should in fact differ depending on these criteria (Theohartkas and Hirst 2002; Polonsky et al. 1999). However, this examination suggests that there is a lack of systematic examination of the underlying evaluative criteria used to rank journals.

A second qualitative stage sought to better understand what criteria were important for evaluating journals. Semi-structured interviews were undertaken with 6 full-professors in the marketing discipline (11 in the US and Australia were invited to participate). Individuals were purposefully selected to ensure they were a) active researchers, b) had global academic experience (US - 3; UK -3; Australia - 6); and c) were involved in staff evaluation. Interviews were summarised and checked by participants.

Interview discussions focused on identifying the criteria that could be used to rank journals. Discussions also discussed methodological issues associated with the study (for example, what criteria should be used, should journals be rated or ranked, and what scales should be used). The most salient outcomes were that four criteria were identified for examination - prestige, contribution to theory, contribution to practice and contribution to teaching, and that the rating scale should be used with A,B,C type “intervals”, which are frequently used in journal evaluations. While respondents agreed that these criteria could be used to evaluate journals, they did not necessarily agree on the importance or relevance of each (which is the focus of this study). A draft survey was then developed and distributed to the 11 professors for feedback (one retired and was replaced) and modifications were made to the final instrument.

The final staged involved an on-line survey. This survey asked respondents to evaluate journals with which they were *familiar* from a set of 94 journals. This evaluation involved four criteria using the recommended seven-point scale (A+, A, B+, B, C+, C, D and N/A). However, it should be noted that this data is not examined within this paper. In addition to evaluating the journals, the respondents were also asked a series of questions such as; whether they had published or submitted work to these journals, as well as a range of individual and institutional demographic questions. They were also provided with two, 100-point summed scaled questions to identify the relative importance of the four evaluation criteria. The first scale was for the individual’s weighting and the second was for the individual’s perception of their institution’s weighting). It is this latter data, which is examined in this paper.

The ultimate aim of the broader research project will be to develop weighted ratings for journals where a weighting index is determined. However, it is essential to understand differences in weighting criteria before these complex ratings can be evaluated. An announcement for the survey was distributed via Elmar. There is no demographic data available on list members. Peter Palij, the founder and past moderator of the list has indicated, "that at the end of 2002, there were 3,250 subscribers with an approximate; 60/40 split for US/non-US, about 2500 faculty, 600 doctoral students and 150" corporate members. In terms of institutions, there was something in the order of 500 to 550 institutions represented worldwide." Elmar was selected for the distribution of the survey announcement because of its global and inclusive nature. The initial response after three weeks was 117 respondents and supplementary requests for participants were sought from two academic associations (1-Europe-based and 1 Aus/NZ-based) via emails to their members. This generated ten additional respondents for a total of 127.

In some cases technology constraints resulted in some people being unable to complete the web-based survey. Those indicating they had a problem, which was incorporated into the call for participation and survey, completed a "hard copy" of the instrument. It was not possible to identify how many people had this problem or how many started the survey but did not complete it.

Data was analysed in a variety of ways. The weightings used to evaluate journals by individuals were compared to those used by their institutions - as perceived by respondents. Any difference in weightings would indicate that the specific rationale or focus of the rating would result in different evaluations. In turn, this would impact on how performance is evaluated (Polonsky and Mankelov 2000), which is one reason for rating journals (Bakir et al. 2000). The weightings, individual and institutional, were then compared across regions (the Americas, Europe, and Asia-Pacific). Authors such as Theoharakis and Hirst (2002) found that evaluations of journals differed globally and thus it is important to examine whether global variations in criteria exist. Lastly, differences in whether the respondent's institution offered a research degree (PhD/DBA) were examined. Hult et al. (1997) suggested that this impacts on how academics view journals and should affect the underlying dimensions for rating journals.

Sample

The research was undertaken to examine the perceptions of the global marketing academic community, not just those at "leading institutions" (Theoharakis and Hirst 2002), active researchers, or Dean/Heads of School (Browne and Becker 1991, Mort et al. 2002). While Elmar is global and inclusive, a general announcement for participation in a web-survey might have resulted in a low response rate. It was hoped that the salience of the topic (Herberlein and Baumgartner 1978) and incentive (Kanuk and Berenson 1975) would have improved the level of e-participation (Chittenden and Rettie 2003). The small sample size (n=127) is possibly a limitation of the study.

Given the lack of information on demographic composition of Elmar and the limited information on the demographic profile of the marketing discipline within countries (Baker and Erdogan, 2000; Danaher and Starr, 1998; Hetzel 2000; Sinkovics and Schlegelmilch 2000), it is not possible to evaluate the representativeness of the sample to either Elmar or the constituent regions. However, some demographic characteristics of respondents were as follows – location of employment (Americas 59%; Europe 20%; Asia-Pacific 18%; not stated

4%); PhD/DBA granting (yes 56%, no 44%); average age (41.16 years); mean years in academia (11.01); mean years in current position (5.14 years); gender (female 29%, male 71%) and level (Professor 19%; Associate/Senior Lecturer 24%; Assistant/Lecturer 41%; Tutor/instructor 3%; and Other 13%).

Analysis

Overall, the respondents felt the importance of criteria (expressed as a percentage) were ordered from contribution to theory (35%), prestige (26%), practice (23%) and teaching (14%) and that for their perceptions of their institution, prestige (56%) was generally more important than contribution to theory (19%). This in itself might have implications if individuals select publishing opportunities on criteria that are different to their institutions.

T-tests were used to examine whether individuals' personal attribution to the levels of importance of the evaluative criteria differed from their perceptions of their institution. Column 3, of Table 1 shows that there are statistically significant differences ($p < .05$) between each of the four criteria pairs of comparisons. Individuals believed that a journal's Contribution to Theory, Practice and Teaching is more important than does their institution. On the other hand, individuals believed that prestige is less important than does their institution.

Table 1: Comparisons of Individual (Ind) - Institutional (Inst) weightings and Regional variations

(Paired comparisons ** $p < .05$; * $p < .10$, ns=not significant)

C1 Importance of the Journals'	C2 Group	C3 Total Mean (std) N=127	C4 Asia-Pacific Mean (std) N=23	C5 Europe Mean (std) N=25	C6 Americas Mean (std) N= 74	C7 ANOVA F [Sig]
Prestige	Ind.	25.87 (16.11) **	27.39 (17.39) **	18.16 (12.01) **	28.05 (16.82)**	3.672 [.028] **
	Inst.	55.59 (30.82)	56.09 (34.38)	58 (27.27)	57.70 (29.37) **	0.031 [0.97]
Contribution to theory	Ind.	34.90 (13.98) **	34.35 (14.56) **	42.04 (15.92) **	32.85 (12.73) **	4.194 [0.02] **
	Inst.	19.22 (16.87)	23.26 (22.09)	22.60 (19.53)	17.72 (13.63)	1.410 [0.248]
Contribution to practice	Ind.	23.15 (10.22) **	21.96 (8.22) **	21.44 (9.99) **	23.98 (11.08) **	0.724 [0.49]
	Inst.	10.69 (10.83)	9.35 (10.37)	9.6 (10.17)	12.07 (11.26)	0.826 [0.44]
Contribution to teaching	Ind.	13.72 (9.12) **	16.30 (8.56) **	15.36 (7.60) **	12.41 (9.59) **	1.758[0.18]
	Inst.	6.46 (9.77)	6.52 (8.85)	5.00 (7.47)	7.24 (10.91)	0.479 [0.62]

A t-test comparison of differences in individual and institutional evaluations within each region (Columns 4, 5 and 6) identifies a similar pattern, suggesting that there are globally consistent patterns. A MANOVA across all variables ($F=1.67$; $p < .10$) identified that there were differences across the set of variables. The individual ANOVA comparison across the three regions (Column 7) indicates that there are few statistical differences. In terms of individual differences, European respondents seem to perceive prestige to be least important and theory development as most important. This might suggest that European and Asia-Pacific academics are less concerned with a journal's prestige and more concerned its ability to develop theory. This could explain why Theotarkis and Hirst (2002) found regional

differences in the evaluations of the same journal. There were no statistically significant institutional weighting differences across the regions.

Table 2 reports the analysis of comparisons based on whether the respondent works at a PhD/DBA granting institution. The overall patterns found in Table 1, relating to differences in individual and institutional weightings, are found to hold (Column 3 & 4). A MANOVA comparison between respondents employed at PhD/DBA granting and non-granting institutions identifies that there are statistically significant differences in weighting criteria ($F=3.23$; $p<.01$). An examination of the individual ANOVAs identified three statistically significant differences. Individuals at PhD/DBA granting institutions seem to value a journal's ability to develop theory more than individuals at non-granting institutions. On the other hand, individuals at non-PhD/DBA granting institutions seem to value a journal's contribution to practice more than those at granting institution. This would seem to support the suggestion that there are in fact differences in the foci of individuals in these institutions (Hult et al. 1997). This could possibly impact on how individuals within non-PhD/DBA institutions are evaluated, especially if they choose to target practically focused outlets.

Table 2: Differences in Weighting Based on PhD/ DBA Granting
(** $p<.01$; * $p<.10$, ns=not significant)

C1 Importance of the Journals'	C2 Group	C3 PhD or DBA Granting Mean (std) N=71	C4 Non PhD or DBA Granting Mean (std) N=56	C5 Comparisons Between Granting and Non-Granting
Prestige	Individual	23.85 (15.17)**	28.4 (17.18) **	NS
	Institution	57.39 (30.09)	53.55 (32.08)	NS
Contribution to theory	Individual	36.89 (15.48)**	32.24 (11.51) **	$F=3.32$ $P<.10$
	Institution	22.01 (18.83)	15.6 (13.45)	$F=4.54$ $P<.05$
Contribution to practice	Individual	21.60 (10.29)**	25.31 (9.87) **	$F=3.81$ $P<.10$
	Institution	10.14 (11.14)	11.24 (10.51)	NS
Contribution to teaching	Individual	14.85 (8.98) **	12.24 (9.35) **	NS
	Institution	5.94 (8.89)	6.89 (10.78)	NS

Conclusions, Implications and Future Research

The results seem to suggest that when evaluating journals there are factors that impact on how journals are evaluated. For example, whether an individual or the institution is evaluating the journal will have a significant effect on the importance of the criteria used. In addition, there are some regional differences in the perceived importance of criteria, which may impact on global evaluations of journals. Lastly, as has been suggested previously, the type of institution one works at does seem to affect how individuals view the evaluative criteria for rating journals. However, these differences seem to be more salient for individuals than institutions, although non-PhD/DBA granting institutions seem to value less a journal's contribution to theory.

The implications of these results are that the criteria used for evaluating a journal does seem to vary by; region, where the person teaches (PhD/DBA granting or not), the motivation for ranking. A given journal's ranking *might* therefore vary because of differences in the specific evaluative criteria used or because of the weightings themselves. *If* this were the case, it would be important that those individuals and institutions using rankings understand the criteria that are used and their relative importance. Communicating "global" ranking of journals may thus be overly simplistic and ignore the existence of the underlying complexity

of this construct. As a consequence, different sets of criteria might result in different rankings and therefore there is a strong case for better conceptualisation and enunciation of the various rating and ranking criteria.

References

ANBAR (2003) 'Literati Awards Archive' (<http://taddeo.emerald-library.com/vl=6697476/cl=11/nw=1/rpsv/literaticlub/whatsnew/archive.htm>) accessed 14/6/03.

Baker, M.J. and B.Z. Erdogan (2000) "Who We Are and What We Do-2000," *Journal of Marketing Management*. 16(7): 679-698.

Bakir, B., S.J. Vitell and G.M. Rose (2000) "Publications in major marketing journals: An analysis of scholars and marketing departments," *Journal of Marketing Education* 22(2): 99-107.

Baumgartner, H. and R. Pieters (2003) "The Structural Influence of Marketing Journals: A Citation Analysis of the Discipline and its Subareas Over Time," *Journal of Marketing* 67(1): 123-139.

Browne, W.G. and B.W. Becker (1991) "A Longitudinal Study of Marketing Journal Familiarity and Quality," Summer AMA Proceedings 702-710.

Chittenden, L. and R. Rettie (2003) "An Evaluation of E-mail Marketing and Factors Affecting Response," *Journal of Targeting, Measurement and Analysis for Marketing*, 11(3): 203-217.

Danaher, P. and R. Starr (1998) "A Study of Marketing Departments in Australasian Universities." *European Journal of Marketing*, 32 (9/10): 771-784.

Danielsen, A. and C.D. Delorme Jr. (1976), "Some Empirical Evidence on the Variables Associated with the Ranking of Economics Journals," *Southern Economic Journal*, 43(2):1149-1160.

Elmar (2003) 'Elmar Digest' <http://www.columbia.edu/~pbp1/elmar/toc.html> accessed 2003

Hawes, J.M. and B. Keiller (2002) "Assessing Marketing Journals: A Mission-Based Approach," *Journal of the Academy of Business Education*, 3(2): 70-86.

Herberlein, T.A. and R. Baumgartner (1978), "Factors Affecting Response rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature," *American Sociological Review*, 43(4): 447-462.

Hetzl, P. (2000) "Where are we Going? Perceptions of French Marketing Academics," *Journal of Marketing Management*. 16(7): 717-744.

Hawkins, R.G., L.S. Ritter and I. Walter (1973), "What Economists Think of Their Journals."

Journal of Political Economy, 81(July-Aug):1017-1032.

Hult, G.T.M., W.T. Neese, and R.E. Bashaw (1997), "Faculty Perceptions of Marketing Journals." *Journal of Marketing Education*, 19(Spring):37-52.

Jobber, D. and P. Simpson (1988), "A Citation Analysis of Selected Marketing Journals," *International Journal of Research in Marketing*, 5:137-142.

Kanuk, L. and C. Berenson (1975), "Mail Survey and Response Rates: A Literature Review," *Journal of Marketing Research*, 12 (4): 440-453.

Luke, R. H. and E.R. Doke (1987), "Marketing Journal Hierarchies: Faculty Perceptions 1986-1987," *Journal of the Academy of Marketing Science*, 15(1):74-78.

Mort, G.S., J.R. Mcoll-kennedy, G. Keil and G. Soutar (2002) "Senior Faculty Perceptions of the Quality of Marketing Journals: An Australian and New Zealand Perspective, *ANZMAC* 2431-2437.

Neway, J.M. and F.W. Lancaster (1983), "The Correlation Between Pertinence and Rate of Circulation Duplication in Multidatabase Searches," *Journal of the American Society for Information Science*, 34(4):292-293.

Nobes, C.W. (1985), "International Vacations in Perceptions of Accounting Journals," *Accounting Review*, 60(4):702-705.

Polonsky, M.J. and G. Mankelow (2000) "Where Are We Going? Perceptions of US Marketing Academics." *Journal of Marketing Management*. 16 (7): 717-743

Polonsky, M., G. Jones and M. Kearsley (1999) "Accessibility, an Alternative Method of Ranking Marketing Journals?" *Journal of Marketing Education*. 21(4): 181-193.

Polonsky, M.J. and D. Waller (1993) "Marketing Journals and Asia Pacific Marketing Academics: An Exploratory Study" *Asia - Australia Marketing Journal*, 1(1): 61 - 70.

Sinkovics, R.R and B.B. Schlegelmilch (2000). "Marketing Academics in Austria, Germany and Switzerland: Hulbolt's Ideas Give Way to Performance Pressure," *Journal of Marketing Management*. 16(7): 745-760.

Theoharakis, V. and A. Hirst (2002) "Perceptual Differences of Marketing Journals: A Worldwide Perspective," *Marketing Letters* 13(4): 398-402.