

Are There Differences Between Reviewers on the Criteria They Use to Evaluate Research Articles?¹

Michael A. Campion
Purdue University

Have you ever been surprised by how different reviewers seem to focus on different aspects of an article? Some focus on theory, while others focus on methods. Still others focus on data analysis or writing style. This question led to a study on the criteria reviewers use when evaluating a research article. Through a two-part (Delphi-like) process with over 300 reviewers, a comprehensive checklist of criteria for reviewing research articles was developed. This 223-criteria "Article Review Checklist" was published in the September 1993 issue of *Personnel Psychology* (volume 46, number 3). The study also examined similarities and differences between reviewers on the criteria they use. It may be, for example, that there are differences between those employed in academic (e.g., universities) versus applied (e.g., industry, government) settings, or there may be differences related to experience in publishing or reviewing. The present article reports on these differences between reviewers.

Previous Research

There has been a fair amount of previous research on the article review process. Much of it occurring during the 1970s, and much of it was published in the *American Psychologist*. Reviewers from a wide range of different journals and disciplines have been studied, including psychology journals (Brackbill & Korten, 1970; Cicchetti, 1980; Crandall, 1978; Fiske & Fogg, 1990; Gottfredson, 1978; Scarr & Weber, 1978; Scott, 1974; Watkins, 1979; Wolff, 1970; 1973), management journals (Daft, 1985; Kerr, Tolliver, & Petree, 1977; Mitchell, Beach, & Smith, 1985), sociology journals (McCartney, 1973; Smigel & Ross, 1970), and scientists from other disciplines (Chase, 1970).

A substantial proportion of this research focused on the reliability of the review process. Assessed in terms of the correlation between reviewer recommendations, reliability appears to be quite low (e.g., .20 in Fiske & Fogg, 1990; .21 in Hendrick, 1977; .26 in Scott, 1974). However, when assessed in terms of agreement between reviewers, the picture is more positive (e.g., 70% in Crandall, 1978; 78% in Scarr & Weber, 1978; 72% in Smigel & Ross, 1970). More highly-controlled studies with more extensive

¹ Special thanks to the hundreds of reviewers who helped provide the answer to this question. The criteria examined in this study are contained in the Article Review Checklist which is published in *Personnel Psychology* (volume 46, number 3).

measurement scales have obtained higher levels of reliability (e.g., .46 in Gottfredson, 1978, .62 in McReynolds, 1971; .48 in Reilly, Balzer, & Doherty, 1989), although these studies probably overestimate the reliability of the normal article review process. When examining specific reviewer comments rather than ratings of criteria or publication recommendations, the level of agreement has been found to be quite low, however (Fiske & Fogg, 1990).

One of the more provocative examinations of the reliability of the review process was a study that resubmitted previously published, but disguised, studies (Peters & Ceci, 1982). Of the 12 studies, only 3 were detected as resubmissions, and 8 out of the remaining 9 were rejected. Although Peters and Ceci favor a bias explanation for the results created by the author's status and institutional affiliation, extensive peer commentary following the article indicates a notable lack of consensus on the proper interpretation.

More recently, there have been a number of nonempirical writings on the review process. An excellent book devoted entirely to the publication process by Cummings and Frost (1985) contains several chapters discussing issues surrounding reviewing (e.g., Campbell, 1985; Rousseau, 1985; Schwab, 1985). A special issue in the *Academy of Management Review* on theory development also contains articles on criteria for judging conceptual papers (Bacharach, 1989; Whetten, 1989; also see Klimoski, 1991).

There are several differences between the previous research and the present study. First, most of the empirical articles examined editorial recommendations (Crandall, 1978; Peters & Ceci, 1982; Scarr & Weber, 1978; Smigel & Ross, 1970), rather than criteria as is the focus of the present study. Second, previous studies have only examined very limited differences between reviewers such as the journal they reviewed for (Gottfredson, 1978; Mitchell et al., 1985; Kerr et al., 1977; Wolff, 1970) or their scientific discipline (Chase, 1970). The present study examined a wider range of reviewer background variables. Third, the lists of criteria examined in previous studies have been very global (e.g., 10 in Chase, 1970; 11 in Daft, 1985; 14 in Frantz, 1968; 12 in Mitchell et al., 1985; 7 in Scott, 1974; 15 in Wolff, 1970), or the method of ensuring their thoroughness was not described in the report (Gottfredson, 1978; Kerr et al., 1977; Wolff, 1973). (The content analysis of reviewers' comments reported in Fiske and Fogg (1990) is an exception.) The present study used the comprehensive 223-criteria list described above.

Method

The development of the Article Review Checklist is described in an editorial in *Personnel Psychology* (1993, volume 46, number 3). The sample consisted of the editorial board and ad hoc reviewers for *Personnel Psychology*, the editorial boards of *Journal of Applied Psychology* and *Academy of Management Journal*, and reviewers from recent scientific

conferences (e.g., American Psychological Association and Society of Industrial and Organizational Psychology). A total of 344 checklists were sent out and 227 were returned, for a 66.0% response rate .

Reviewers were asked to rate how much weight they gave each of the criteria on the following scale:

5. *Very large weight* (fatal flaw if not satisfactory)
4. *Large weight* (contributes importantly to cumulative judgment, and could be a fatal flaw in some instances)
3. *Moderate weight* (contributes to cumulative judgment, but could not be fatal flaw)
2. *Small weight* (does not influence decision)
1. *Very small weight* (not important or depends on too many other factors)

The checklist also contained 13 background items: (1) highest degree (Doctorate, Masters, or other); (2) years of work experience in the field; (3) current primary employer (psychology department, business or management department, research organization, consulting firm or private practice, government, company, or other); (4) primary professional association (Society of Industrial and Organizational Psychology, Academy of Management, American Psychological Association, American Psychological Society, Industrial Relations Research Association, or other); (5) percentage of time engaged in various activities (research, teaching, practice, management or administration, or other); (6) average number of refereed journal publications (on a 4-point scale ranging from "one or more per year" to "not currently active"); (7) average number of convention presentations (same scale); (8) average number of other publications or presentations (same scale); (9) number of years served on journal editorial boards (counting 1 for each year per journal); (10) average number of journals providing ad hoc reviews per year; (11) number of years served as a reviewer for conventions (counting 1 for each year per convention); (12) primary area of specialty (Human Resources/Industrial Psychology; Organizational Behavior/Organizational Psychology, Industrial Relations, or other); and (13) tenure (yes, no, or not applicable; and, if yes, how many years). Differences between editorial boards were not examined because previous research has shown few differences across similar journals (Gottfredson, 1978; Mitchell et al., 1985; but cf. Kerr et al., 1977), and because of overlapping membership among the boards.

Results

The mean rating across the 223 criteria was 3.71, with a range from 3.00 to 4.55. This suggested the reviewers discriminated somewhat among the criteria in terms of the weightings assigned. The average weights for criteria in each category are presented in Table 1, and some variation across categories can be observed.

There also appeared to be substantial differences among the reviewers. The average standard deviation of the criteria was .87, with a range from .65 to 1.20. As an assessment of interrater agreement, an index proposed by James, Demaree, and Wolf (1984) was calculated. To control the level of chance agreement, the measure compares the variance based on the observed distribution of ratings to the variance expected based on a null distribution (i.e., the distribution if there was no true variance in the judgments). The expected null distribution used in the present study was the slightly negatively skewed distribution described by James et al. (1984). This distribution was selected for three reasons. First, positive leniency is common (James et al., 1984). Second, it is likely that reviewers only suggested criteria to be included in the checklist that were at least somewhat important. Third, the null skewed distribution has a mean very close to the observed distribution (3.6 versus 3.7). Although James et al. warn against using the observed distribution to pick a null distribution because much of the skew may represent true variance, it was nevertheless viewed as a conservative test.

The average interrater agreement was .43, with a range from .00 to .69 (and only three below .10), indicating modest overall agreement at the level of the individual criteria. The level of average agreement among individual criteria was similar across categories (Table 1). Although the agreement on individual criteria was modest, the reliability of the composites of all criteria within each category were very high (Table 1).

Differences between Reviewers

For the purposes of analyzing differences between reviewers, composites were created by averaging the criteria within each of the 15 categories. The coefficient alpha reliabilities of these composites were all quite high, except for the category on the importance of the topic (Table 1). Principle components analysis of this category revealed three factors, but internally consistent subcomposites could not be formed. Therefore, the three individual criteria that best reflected these factors were analyzed separately. These criteria were: theoretical importance, practical importance, and appropriateness to the journal and readership. All reviewer differences were examined based on the remaining 14 composites and these three individual criteria. For each background variable, either a canonical correlation analysis (for continuous variables) or multivariate analysis of variance (for categorical variables) was performed to control for experiment-wise error rate, followed by either correlations or *t*-tests, respectively, to examine individual criterion measures. The significant results are presented in Table 2.

Nearly all reviewers (93.5%) had Ph.D. degrees, thus analyses of differences based on degree were not needed. Years of work experience in the field averaged 13.7 years ($SD = 8.84$). More experienced reviewers gave

slightly less weight to theoretical importance, and more weight to the procedures, discussion and conclusions, and presentation.

Primary employers were as follows: psychology department (27.6%), business or management department (47.5%), research organization (5.5%), consulting firm or private practice (6.0%), government (1.8%), company (6.5%), and other (5.1%). There was an adequate sample to examine two divisions of the data. Those in academic versus nonacademic positions gave more weight to theoretical importance, less weight to practical importance, less weight to the discussion and conclusions, and less weight to the presentation. Those in psychology versus business departments gave less weight to theoretical importance, more weight to the appropriateness of the topic to the journal, and more weight to the presentation.

Primary professional associations were as follows: Society of Industrial and Organizational Psychology (53.8%), Academy of Management (35.7%), and all others (10.5%). No differences were observed in the weightings based on professional association.

Across reviewers, an average of 41.6% ($SD = 22.5$) of time was spent conducting research, 25.4% ($SD = 17.0$) teaching, 13.0% ($SD = 20.1$) in practice, 15.1% ($SD = 23.1$) in management or administration, and 3.8% ($SD = 12.3$) in other activities. Those spending more time doing research gave more weight to theoretical importance, and less weight to practical importance, sample and setting, procedure, discussion and conclusions, and presentation. Similarly, those who spent more time teaching gave less weight to practical importance, discussion and conclusions, and presentation. The opposite trend was observed for those who spent more time in practice or in management or administration.

The reviewers were very active with respect to publishing, with 68.2% averaging one or more refereed journal publications per year, 78.0% averaging one or more convention presentations per year, and 63.1% averaging one or more other publications or presentations per year. Those publishing and presenting more gave more weight to theoretical importance and less weight to practical importance.

The reviewers were also very active with respect to reviewing, with the average respondent serving 6.3 years ($SD = 9.61$) on journal editorial boards, providing ad hoc reviews for 3.3 ($SD = 2.07$) journals per year, and serving 6.4 years ($SD = 6.28$) as a reviewer for conventions. Similar to publication activity, those more involved in reviewing gave less weight to practical importance and more weight to theoretical importance and contribution.

Primary specialty areas were 58.0% human resources/industrial psychology, 28.6% organizational behavior/organizational psychology, 2.9% industrial relations, and 10.5% other areas. Those in human resources versus organizational behavior gave more weight to practical importance.

Finally, 73.0% of the academic reviewers had tenure, and they had tenure for an average of 8.3 years ($SD = 6.67$). Neither the possession of tenure nor years of tenure related to the criterion weights, however.

Discussion

Summary and Conclusions

Perhaps the most striking observation is the relative lack of systematic reviewer differences in the weightings they assigned to the criteria. Despite the discriminations reviewers made among criteria, the differences between reviewers based on their backgrounds were comparatively few in number and small in magnitude. After examining over a dozen background variables, the most notable trend could be summarized simply as follows. Those who are in nonacademic positions, who spend more time in practice and management activities, or who are less involved in publishing and reviewing, tend to weight the practical importance of research articles slightly higher and the theoretical importance slightly lower than those with the opposite backgrounds. They also give slightly more weight to the discussion, conclusions, and presentation. These criteria, like practical importance, could be considered the more immediately useful aspects in terms of applying the findings of a research article. The differences were not large on these criteria, and few differences were found on most of the categories of criteria or on many of the background variables.

These differences tend to confirm commonly held assumptions that more academic-oriented scientists are somewhat more interested in theoretical implications of research, while practice-oriented scientists are naturally somewhat more interested in practical implications. In terms of the review process, these differences may support the value of using both academic and nonacademic reviewers.

Nevertheless, the lack of larger and more consistent differences may be the most noteworthy finding. It may suggest that reviewers are not overly influenced by their own backgrounds and situations in the criteria they apply to research articles. This conclusion is limited, of course, by the methodology used in this study and the range of background variables examined. Substantial variation existed between reviewers on the weightings they assigned the criteria that cannot be explained by this study. In terms of the review process, this variation may suggest that more reviewers should be used to judge each manuscript (e.g., use 3 rather than the more typical 2).

Apparent versus Real Reviewer Differences

There are several reasons apparent differences between reviewers may be much larger than real differences. First, interrater reliability may be a poor way to measure reviewer similarity because covariation is adversely influenced by range restriction and response tendencies. The journal review

process has a restricted range, as evidenced by the low acceptance rates, and most reviewers are correspondingly severe in their judgments. Interrater agreement may be a better measure of similarity because it is less sensitive to such distributional properties, and it reflects consensus rather than covariation among raters which seems more consistent with the decisional nature of the review process. As noted previously, research has tended to show that the reliability of the review process is low, while agreement is moderate to high. The findings of the present study confirm these observations.

Second, the amount of similarity in the comments reviewers make to the author depends upon the level of aggregation. Similarity may appear low if the focus is on specific comments (Fiske & Fogg, 1990), but similarity may appear higher if the comments are aggregated to a more global level. For example, reviewers may make a high proportion of comments on conceptual issues and few on analytic issues on one manuscript, but give the opposite pattern on another manuscript. In other words, they may not make the exact same comments, but there may be good similarity in terms of the overall nature (i.e., categories) of their concerns. Likewise, it should also be noted that the comments reviewers make to authors may not always reflect their recommendations to the editor. In other words, there may be differences in specific suggestions for improvement to the author, but similarity in the acceptance recommendations made to the editor.

Third, agreement is much higher if it is defined as the majority of opinion. That is, agreement appears more positive if it is based on similarity between two out of the three reviewers. With the unexplained differences between reviewers documented above, such a definition of agreement may be more realistic.

Even though apparent differences among reviewers may be larger than real differences, there still are some real differences built in by design. Reviewers are often picked for different perspectives or areas of expertise (Campbell, 1985). The use of both academic and practitioner perspectives as noted above is an example. Other examples might include the use of reviewers who are known to hold strong (and often opposing) views, or the use of both theory-oriented and methods-oriented reviewers on the same article.

Finally, the editor's role is to help maintain measurement equivalence across articles submitted to the review process. This is accomplished by subjectively standardizing the reviewers' recommendations (i.e., correct for leniency or severity based on the history of previous judgments from each reviewer), by only using inexperienced reviewers when they can be paired with experienced reviewers, by acting as an additional reviewer of each paper, and by seeking to maintain a consistent level of rigor over time.

Future Research

First, future research could examine actual reviewer behavior. One clear limitation of this study is the self-report nature of the methodology. Reviewers were asked what criteria they use, but this may not be the same as what they actually use. In this sense, the study examined prescriptive norms, or idealized reviewer behavior, rather than descriptive norms, or summaries of actual reviewer behavior (Gottfredson, 1978).

Second, future research could validate the criteria against external measures of scientific quality or contribution. Although research using retrospective self-report data suggests that scientific impact may be related to some criteria (Daft, Griffin, & Yates, 1987), research using citation indices reveals only modest relationships with these types of judgmental criteria (Gottfredson, 1978).

Third, future research could further examine the disagreement among reviewers that does exist. It is possible that such differences are systematic, but require a detailed analysis of differences in scientific training to explain. It is also possible that such differences represent nonsystematic variation or error, and several previous authors have suggested that the review process should be more highly standardized to reduce this error (e.g., Brackbill & Korten, 1970; Wolff, 1973). It is noteworthy that several reviewers who participated in this study spontaneously recommended against such an effort by arguing that judging science was too complex to be standardized and scientific creativity itself would be stifled if subjected to such standardized review.

Fourth, future research could examine potential biases that might operate in the review process. For example, previous research on the review process focused substantial attention on the influence of the author's reputation and affiliation (Kerr et al., 1977; Peters & Ceci, 1982; Rowney & Zenisek, 1980). In contrast, none of the reviewers participating in this study mentioned the author as a relevant consideration. It should also be noted that many journals utilize a "double-blind" review process wherein both the reviewers and authors are unaware of each other's identities. Another criterion shown to be influential in previous research, but not suggested for inclusion in the present study, was whether the study reported statistically significant results (e.g., Atkinson, Furlong, & Wampold, 1982; Kerr et al., 1977). These, and other more direct forms of bias (e.g., support for one's preferred theoretical perspective; Mahoney, 1977), may still potentially influence reviewer judgments, and thus might offer a useful direction for future research.

Finally, future research might examine whether studies like this are a potentially useful developmental experience for reviewers, editors, and editorial boards. Reflecting on their reviewing practices and sharing opinions with others may have helped refine their skills. This was the opinion of a large number of reviewers participating in the present study. Perhaps editorial

boards should conduct such organizational development studies when they are initially formed in order to identify the criteria most important to the journal and enhance agreement among reviewers.

References

- Atkinson, D. R., Furlong, M. J., & Wampold, B. E. (1982). Statistical significance, reviewer evaluations, and the scientific process: Is there a (statistically) significant relationship? *Journal of Counseling Psychology, 29*, 189-194.
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review, 14*, 496-515.
- Brackbill, Y., & Korten, F. (1970). Journal reviewing practices: Authors' and APA members' suggestions for revision. *American Psychologist, 25*, 937-940.
- Campbell, J. P. (1985). Editorial: Some remarks from the outgoing editor. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 321-333). Homewood, IL: Irwin.
- Chase, J. M. (1970). Normative criteria for scientific publication. *American Sociologist, 5*, 262-273.
- Crandall, R. (1978). Interrater agreement on manuscripts is not so bad! *American Sociologist, 33*, 623-624.
- Cummings, L. L., & Frost, P. J. (Eds.). (1985). *Publishing in the organizational sciences*. Homewood, IL: Irwin.
- Daft, R. L. (1985). Why I recommended that your manuscript be rejected and what you can do about it. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 193-209). Homewood, IL: Irwin.
- Daft, R. L., Griffin, R. W., & Yates, V. (1987). Retrospective accounts of research factors associated with significant and not-so-significant research outcomes. *Academy of Management Journal, 30*, 763-785.
- Fiske, D. W., & Fogg, L. (1990). But the reviewers are making different criticisms of my paper! Diversity and uniqueness in reviewer comments. *American Psychologist, 45*, 591-598.
- Frantz, T. T. (1968). Criteria for publishable manuscripts. *Personnel and Guidance Journal, 47*, 384-386.
- Gottfredson, S. D. (1978). Evaluating psychological research reports: Dimensions, reliability, and correlates of quality judgments. *American Psychologist, 33*, 920-934.
- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85-98.
- Kerr, S., Tolliver, J., & Petree, D. (1977). Manuscript characteristics which influence acceptance for management and social science journal. *Academy of Management Journal, 20*, 132-141.
- Klimoski, R. (1991). Theory presentation in human resource management. *Human Resource Management Review, 1*, 253-271.
- Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*, 161-175.
- McCartney, J. L. (1973). Manuscript reviewing. *Sociological Quarterly, 14*, 290, 440-444.
- McReynolds, P. (1971). Reliability of ratings of research papers. *American Psychologist, 26*, 400-401.
- Mitchell, T. R., & Beach, L. R., & Smith, K. G. (1985). Some data on publishing from the authors' and reviewers' perspectives. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 248-264). Homewood, IL: Irwin.
- Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences, 5*, 187-255.
- Reilly, C. E., Balzer, W. K., & Doherty, M. E. (1989, April). *Journal reviewers judgments of external validity: A policy capturing approach*. Paper presented at the meeting of the

Society for Industrial and Organizational Psychology, Boston, MA.

Rousseau, D. M. (1985). Publishing from a reviewer's perspective. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 182-192). Homewood, IL: Irwin.

Scarr, S., & Weber, B. L. R. (1978). The reliability of reviews for the American Psychologist. *American Psychologist*, 33, 935.

Schwab, D. P. (1985). Reviewing empirically based manuscripts: Perspectives on process. In L. L. Cummings & P. J. Frost (Eds.), *Publishing in the organizational sciences* (pp. 171-181). Homewood, IL: Irwin.

Scott, W. A. (1974). Interreferee agreement on some characteristics of manuscripts submitted to the Journal of Personality and Social Psychology. *American Psychologist*, 29, 698-702.

Smigel, E. O., & Ross, H. L. (1970). Factors in the editorial decision. *American Sociologist*, 5, 19-21.

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14, 490-495.

Wolff, W. M. (1973). Publication problems in psychology and an explicit evaluation schema for manuscripts. *American Psychologist*, 28, 257-261.

Table 1. Descriptive Statistics

Category	Number of Criteria	M	SD	Coefficient Alpha	Average Criterion Agreement	Composite Agreement
Importance of topic	8	3.62	.43	.50	.41	.85
Literature review	16	3.61	.49	.85	.45	.93
Conceptual development	18	3.78	.53	.89	.40	.92
Criteria for reviews and conceptual papers	12	3.80	.55	.83	.37	.88
Sample and setting	13	3.82	.57	.88	.42	.90
Measurement	20	3.74	.53	.91	.41	.93
Design—experimental	19	3.93	.55	.90	.45	.94
Design—nonexperimental	12	3.71	.56	.88	.47	.91
Design—meta-analysis	11	4.06	.50	.84	.52	.92
Design—qualitative	15	3.92	.62	.92	.46	.93
Procedure	10	3.71	.61	.86	.37	.85
Data analysis and results	21	4.07	.50	.92	.50	.96
Discussion and conclusions	20	3.64	.55	.93	.48	.95
Presentation	14	3.49	.62	.92	.40	.90
Contribution	14	3.57	.54	.81	.35	.88

Note. $n = 215$ to 220 .

Table 2. Differences Between Reviewers

Years of work experience

- Those with more experience (Wilks Lambda = .81, $F = 2.35$) gave:
- * less weight to theoretical importance ($r = -.21$)
 - * more weight to procedures ($r = .20$)
 - * more weight to discussion and conclusions ($r = .20$)
 - * more weight to presentation ($r = .26$)

Primary employer

- Those in academic, compared to nonacademic, positions (Lambda = .72, $F = 3.68$) gave:
- * more weight to theoretical importance ($t = 4.76$)
 - * less weight to practical importance ($t = -3.32$)
 - * less weight to discussion and conclusions ($t = -2.06$)
 - * less weight to presentation ($t = -2.37$)

- Those in psychology, compared to business, departments (Lambda = .72, $F = 2.82$) gave:

- * less weight to theoretical importance ($t = -2.44$)
- * more weight to appropriateness of the topic to the journal ($t = 2.04$)
- * more weight to presentation ($t = 2.18$)

Primary professional association (Lambda = .84, $F = 1.62$, not significant)

Work Activities (Lambda = .58, $F = 1.39$)

Those spending more time doing research gave:

- * more weight to theoretical importance ($r = .22$)
- * less weight to practical importance ($r = -.26$)
- * less weight to sample and setting ($r = -.14$)
- * less weight to procedure ($r = -.17$)
- * less weight to discussion and conclusions ($r = -.18$)
- * less weight to presentation ($r = -.20$)

Those spending more time teaching gave:

- * less weight to practical importance ($r = -.19$)
- * less weight to discussion and conclusions ($r = -.15$)
- * less weight to presentation ($r = -.16$)

Those spending more time in practice gave:

- * less weight to theoretical importance ($r = -.18$)
- * more weight to practical importance ($r = .23$)

Those spending more time in management or administration gave:

- * less weight to theoretical importance ($r = -.16$)
- * more weight to practical importance ($r = .17$)
- * more weight to discussion and conclusions ($r = .17$)
- * more weight to presentation ($r = .17$)

Publishing

Those publishing and presenting more (Lambda = .55, $F = 2.15$) gave:

- * more weight to theoretical importance ($r = .27$ and $.25$)
- * less weight to practical importance ($r = -.16$ and $-.21$)

Those doing more reviewing (Lambda = .65, $F = 1.51$) gave:

- * more weight to theoretical importance ($r = .19$)
- * less weight to practical importance ($r = -.16, -.21, \text{ and } -.16$)
- * more weight to contribution ($r = .16$)

Primary specialty area

Those in human resources, compared to organizational behavior (Lambda = .77, $F = 2.47$), gave:

- * more weight to practical importance ($t = 3.24$)

Tenure (Lambda = .84, $F = 1.41$, not significant)

Years of tenure (Lambda = .81, $F = 1.05$, not significant)

Note: All statistics significant at $p < .05$ (two-tailed).