Routledge
Taylor & Francis Group

ARTICLES

Leading North American Programs in Clinical Assessment Research: An Assessment of Productivity and Impact

LESLIE C. MOREY

Department of Psychology, Texas A&M University

To identify doctoral programs with strong concentrations in clinical assessment, I measured productivity and impact of faculty at North American institutions with American Psychological Association accredited clinical programs. Publications, citations, and h-indexes derived from 4 top assessment journals were calculated over a 10-year period (1999–2009). I identified a total of 42 leading programs that collectively accounted for more than half of the publications and citations in these journals. I found a moderate relationship between assessment productivity and both *US News & World Report* program rankings as well as general productivity rankings of clinical programs reported in an earlier study.

Although the field of clinical assessment is one of the oldest in psychology, the past few decades have seen continuing interest in assessment as evidenced by the emergence of new journals such as *Psychological Assessment* and *Assessment* during the 1990s, the development of an assessment-focused organization within the American Psychological Association (Division 12, Section 9), and the publication of numerous specialty assessment book series such as the *Essentials* assessment series by John Wiley & Sons. Research and application of psychological assessment played a central role in the development of the discipline of clinical psychology, and these recent trends suggest a continued and future role for assessment in the discipline.

Given the centrality of assessment within clinical psychology, it is not surprising to find that a number of applicants to graduate training programs express particular interest in assessment training and research. Although accreditation guidelines and disciplinary practices assure that virtually all doctoral trainees in clinical psychology will receive assessment instruction, there are distinct differences among training programs in the degree to which the field of assessment represents a particular area of focus. However, the information available to prospective trainees about such focus is limited. Applicants can rely on reputational rankings of clinical psychology programs, such as those provided annually by *US News & World Report,* but these rankings are fraught with problems (e.g., Winter, Healy, & Svyantek, 1995), and even if accurate, do not convey strengths in particular clinical content areas such as assessment. Researchers such as Stewart, Roberts, and Roy (2007) have attempted to provide a more empirically informed appraisal of programs based on publication productivity. The Stewart et al. article focused on publications by program faculty, which correlated modestly at best with *US News* rankings; it identified University of California San Diego/San Diego State, Pittsburgh, and Penn as among the most productive clinical programs. However, at this point, it is unknown whether productivity across all clinical areas is associated with having faculty publishing work on clinical assessment. The purpose of this article was to provide some information about the top doctoral programs in clinical assessment in terms of research productivity and determine the extent to which such top programs are consistent with those identified in previous studies of clinical program research productivity (e.g., Matson et al., 2005; Stewart et al., 2007) or reputational prestige (e.g., *US News & World Report,* 2008). The former will provide important guidance for the aspiring doctoral trainee in clinical assessment; the latter may provide the field of assessment psychology with information about its role in perceived reputation of clinical doctoral programs as a whole.

METHOD

Although empirically based surveys of research productivity and impact seem considerably more objective than reputation ratings, constructing this empirical base is fraught with complications. These complications involve decisions that will invariably lead to dramatically different conclusions; for example, two such studies of accredited clinical doctoral programs (Matson et al., 2005; Stewart et al., 2007) that have examined data from an overlapping interval reached conclusions that were in many instances appreciably different. As an illustration, one of these studies reported that Louisiana State University's program ranked 6th in research productivity, whereas the other ranked it as 101st. Obviously, there is no "standard" means of defining variables such as productivity, program boundaries, or scientific significance of produced research for the purposes of such studies. In the following paragraphs, I describe the specific approach adopted in this study with a brief rationale for each decision.

Database

I obtained information regarding publications and citations using Thomson Scientific's Institute for Scientific Information

Web of Science, and I collected it in December 2009. I selected the top four journals, in terms of journal 5-year Impact Factor in the *Journal Citation Reports Social Sciences Edition* (Thomson Reuters, 2009), which included the word *assessment* in their title, to be surveyed under the assumption that these journals (unlike most others) would be exclusive in a focus on assessment research. These journals included *Psychological Assessment* (5-year impact factor = 3.63), *Assessment* (2.35), *Journal of Personality Assessment* (2.08), and *Journal of Psychopathology and Behavioral Assessment* (1.61). Selecting the top four seemed like a natural dividing line as the fifth journal, *European Journal of Psychological Assessment,* primarily publishes papers from universities outside of North America, which is the focus of this survey. It is important to note that the largest category of articles in these four journals deal with the assessment of personality and psychopathology, with smaller proportions involving other assessment areas such as psychometrics, intellectual or neuropsychological, marital and family, vocational, or other types of assessment. I included all articles published in the previous 10 years—1999 to 2009, inclusive—in these four journals in the survey. A total of 2,225 articles were indexed in these journals during this 10-year period.

Doctoral Universities/"Programs"

I searched the database described previously using the Organization key in the Web of Science data using terms representing the university or school name for all doctoral clinical psychology programs accredited by the American Psychological Association (total number of programs = 234). Previous studies have been criticized on the basis of inclusiveness or exclusiveness of particular faculty within a "clinical program" (e.g., Andersen et al., 2008; Heesacker & Elliott, 2007), and this issue is not easily resolved given the multitude of joint, adjunct, and cross-appointments that exist in most academic settings. In this study, my focus was on the organization unit rather than on individual faculty members, with the intent of more comprehensively representing the breadth of research and training opportunities available to prospective assessment trainees. Thus, a program was credited with a publication regardless of whether the author was primarily appointed in a clinical program, another program in psychology, or another department in the same university in the belief that any such person would represent an important potential resource for a trainee. Unlike the Stewart et al. (2007) study, accredited programs awarding the Doctor of Psychology (PsyD) were also included in this survey.

Attempting to represent a broad representation of potential training opportunities necessitated some decisions about the boundaries of a university. I made the decision to combine information from all programs (even if separately accredited) located on the same campus or (in the case of medical schools) located in the same metropolitan area. Thus, for example, separately accredited programs at the University of Kansas and at Vanderbilt University were treated as a single program to fully represent the broader set of training opportunities potentially available to trainees in either program. As another example, the representation of the program at Harvard University included work conducted at units associated with Harvard Medical School (such as Massachusetts General Hospital), as the latter represents an important potential assessment resource available to trainees in the Harvard doctoral program. As a contrasting example, the

doctoral program at University of Texas–Austin did not include work conducted at the University of Texas Health Sciences Center at Houston (a separate metropolitan area), as the distance between the two would likely preclude the latter from serving as a training resource for the former.

Authorship

For the purposes of this study, I gave equal credit for a single author, first author, or a later author of any paper published in the target journals. Although other surveys have attempted to distinguish between contributions based on authorship order (e.g., Gibby, Reeve, Grauer, Mohr, & Zickar, 2002), this can be problematic due to different conventions such as the practice within medical school settings for senior authors of a laboratory to be listed last. Because the unit of the search was based on organization rather than individual faculty members, when multiple authors from the same institution contributed to a particular paper, this paper only counted once for that institution. However, if coauthors were from different institutions as defined previously, I credited all contributing distinct institutions with the paper.

Productivity and Impact Indicators

To represent both the amount and the significance of the assessment research conducted in the training programs, I examined three different measures. *Total publications* indicated the sum of all publications in the target journals with at least one author from a particular university. *Citations* tracked the number of times that these publications had been cited since they were published, with more influential papers being cited more frequently. Finally, the *h-index* (Hirsch, 2005) for a given university was obtained as a combined representation of both production and impact. The h-index is the number of papers with citation count $\geq h$; thus, a program that produces four papers that have been cited at least four times has an h-index of 4.

RESULTS

The 2,225 journal articles published in the four study journals included authors affiliated with 177 distinct universities with American Psychological Association accredited programs. Given that there are 234 such accredited programs (some in the same university), this number suggests that trainees in most clinical doctoral programs have access to an active assessment researcher; and this appears to underscore the importance accorded to assessment coverage in clinical doctoral training. These 2,225 articles received a total of 19,539 citations, meaning that the average article in this time span was cited a total of 8.78 times during this 10-year period.

Table 1 provides the intercorrelations (Spearman's rho) among the rank ordering of all accredited programs according to the three impact and productivity indicators as well as the correlation of program rank order with two other "program quality" measures: the total number of program publications provided by Stewart et al. (2007) and the 2008 program ratings as listed in *US News & World Report* (2008). As might be expected, the three productivity indicators I examined in this article correlated substantially, with all correlations in excess of .90. The three indicators demonstrated moderate correlations with both the publication count ratings from Stewart et al. as well as with

TABLE 1.—Rank order correlation of productivity and impact indicators.

Indicator	Publications	Citations	h-Index	Stewart Roberts, & Roy (2007)	US News and World Report (2008)
Publications	1.00				
Citations	.92	1.00			
h-index	.96	.95	1.00		
Stewart et al.	.49	.50	.50	1.00	
US News	.51	.53	.56	.62	1.00

Note. N = 234; all correlations significant, $p < .001$.

the *US News* rankings. Given that there is overlap in year coverage with the Stewart et al. survey (which sampled articles from 2000–2005), it is interesting that the correlation with their estimates is no larger than with the *US News* rankings, indicating

that the most "productive" doctoral programs are not necessarily the most productive assessment programs. Consistent with numbers reported by Stewart et al., their productivity estimates demonstrated a significant correlation (.62) with the most recent *US News* rankings. To determine whether assessment-specific productivity contributes to reputational rank independently of general productivity, I calculated the partial correlation between total assessment publication ranking and *US News* rankings, controlling for Stewart et al.'s total publication rank, and found it to be .31 ($p < .001$), suggesting that a significant portion of the association of assessment productivity and program reputation is independent of general productivity.

Table 2 provides a rank of the top assessment programs listed by total publications but including only those programs ranking in the top 25 on at least one of the three productivity indicators. As a means of comparison, also listed are the rank orderings of these programs according to total publications as calculated by

TABLE 2.—Doctoral programs ranked by total publications in four top assessment journals.

Publication Rank	Citation Rank	h-Index Rank	Stewart et al. Total Pubication Rank	US News and World Report (2008)	School	Assessment Publications	Citations	h-Index
1	1	2	42	4	University of Minnesota	54	939	15
2	4	3	58	57	University of South Florida	53	506	13
3	5	3	n/a	n/a	Harvard University	52	453	13
4	8	7	7	83	Kent State	50	382	11
5	2	1	21	33	University of Kentucky	48	855	16
6	28	24	34	71	Texas A&M University	43	244	8
7	3	7	3	9	University of Pennsylvania	41	599	11
7	33	24	86	71	University of Tennessee	41	214	8
9	7	10	8	33	Boston University	37	436	10
10	10	15	20	43	Florida State University	36	352	9
11	40	36	n/a	131	Virginia Consortium: ODU/EVMS/W&M	34	197	7
12	25	15	35	43	SUNY-Buffalo	33	248	9
13	56	36	141	92	Texas Tech University	32	138	7
14	6	5	122	99	University of North Texas	31	443	12
15	9	5	2	9	University of Pittsburgh	28	373	12
15	30	10	54	6	Yale University	28	225	10
15	14	15	1	25	San Diego State/UCSD	28	319	9
18	19	10	33	9	University of Iowa	27	290	10
19	37	15	30	25	University of Missouri–Columbia	26	200	9
19	36	24	94	n/a	Adelphi University	26	203	8
21	14	15	6	1	University of California–Los Angeles	25	319	9
21	27	15	101	62	Louisiana State University	25	245	9
21	64	50	74	131	University of Toledo	25	124	6
24	18	10	106	16	University of Virginia	24	292	10
24	69	50	83	n/a	Fairleigh Dickinson University	24	105	6
26	11	7	79	92	University of Hawaii	23	341	11
26	12	10	46	n/a	University of British Columbia	23	335	10
26	23	15	88	116	University of Tulsa	23	256	9
26	31	24	37	9	Pennsylvania State University	23	221	8
26	38	24	126	110	University of Arkansas	23	199	8
26	43	24	62	33	University of Georgia	23	178	8
33	20	24	90	16	University of Texas–Austin	22	279	8
36	22	36	32	9	Duke University	20	268	7
36	25	24	41	6	University of North Carolina–Chapel Hill	20	248	8
36	40	24	103	n/a	Simon Fraser University	20	197	8
41	24	24	26	25	Emory University	18	255	8
44	13	15	13	1	University of Wisconsin–Madison	17	326	9
48	17	36	12	62	University of Alabama–Tuscaloosa	15	302	7
50	21	36	61	25	University of Miami	14	270	7
55	14	24	11	8	Indiana University	13	319	8
55	35	24	22	n/a	York University	13	213	8
55	39	15	36	9	Stony Brook–SUNY	13	198	9

Note. Stewart et al. = Stewart, Roberts, and Roy (2007); n/a = not applicable; ODU = Old Dominion University; EVMS = East Virginia Medical School; W&M = William and Mary; SUNY = State University of New York; UCSD = University of California San Diego.

Stewart et al. (2007) and as ranked by *US News* (2008). Note that the Virginia Consortium in Clinical Psychology was not ranked by Stewart et al. because it is a PsyD program. The list of top programs also includes a number of programs not ranked by *US News,* although in some instances, this involves Canadian universities that are not included in their survey. The Harvard program was not included in either survey, perhaps because it was only accredited by American Psychological Association in 2008.

DISCUSSION

This study provides useful information about the distribution and density of clinical assessment researchers across various institutions. The results make it apparent that there are many different doctoral programs with significant intellectual resources for psychological assessment research and training. The 42 schools listed in Table 2 combined for 1,194 articles (some overlapping) and 13,106 independent citations, accounting for 53.6% of the publications and 67.1% of the citations to work in four top assessment journals. The data also underscore previous cautions against an overreliance on "reputation" as an indication of program quality. In many instances, some of the strongest assessment programs in the country received markedly discrepant *US News* rankings (e.g., South Florida, Kent State, Tennessee, Texas A&M, North Texas, or the Virginia Consortium). Nonetheless, Table 2 also includes most of the programs ranked at the top of the *US News* listing, supporting the contention that many of the most reputationally prestigious clinical doctoral programs continue to earn that prestige, at least with respect to their contributions to assessment research.

The comparison with the data reported in Stewart et al. (2007) is also interesting in that there is only a moderate association between total publications represented in the two databases. There are a number of reasons why these numbers would be discrepant. First and most obvious, a program may include many productive clinical faculty but have few who do research in clinical assessment. However, differences in study methodology also play an important role, with this study's organizational-level focus (as opposed to the individual faculty emphasis of Stewart et al., 2007) attempting to include all research on a given campus. This approach thus seeks to recognize important contributions to assessment made by nonclinical (e.g., quantitative, cognitive) psychology faculty as well as contributions made by faculty in other colleges such as Education or Medicine.

It is noteworthy that the associations between productivity indicators in this study correlated as well with the *US News* rankings as they did with Stewart et al.'s (2007) more general productivity data. Although factors contributing to reputational rankings are often quite complex, the finding that assessment-related productivity predicted reputational rankings of clinical doctoral programs above and beyond general research productivity may suggest that expertise in assessment continues to be viewed as an important part of a quality doctoral program, at least by the department chairs that complete the *US News* survey. It would be interesting to replicate this study with similar investigations of the other major content areas within clinical psychology, such as clinical geropsychology (American Psychological Association Division 12, Section 2), to obtain a more specific picture of the impact of various content areas within clinical psychology on perceived program quality.

As with any study of program productivity, this study is limited in a number of respects. For example, there is a considerable amount of high-impact assessment work published in other journals than the four included here; and for some specialty areas (e.g., neuropsychological assessment), these other journals may be particularly important outlets. However, unlike the four target journals, these other journals are typically not exclusively assessment in their focus. Thus, this study should be considered as a sampling of the assessment literature rather than an exhaustive coverage. As another limitation related to coverage, certain important works in the assessment field (such as books, test manuals, and technical papers) are not indexed in the Thomson Scientific database and thus are not represented in these data. In addition, the selection of the organizational unit as the focus of analysis, rather than individual faculty members as in Stewart et al. (2007), has certain disadvantages: It tends to favor organizations of greater size, it does not capture differences between researchers who are core faculty in a program and those more peripheral to that program, and it obviates "per capita" analyses as reported in earlier surveys that could characterize the breadth of different researchers in a program. However, it does provide a more comprehensive coverage of research on a given campus, and it eliminates the need to make ambiguous and potentially arbitrary distinctions between "core," "adjunct," "joint," "clinical," "visiting," and other forms of faculty status.

Finally, it should also be recognized that there is considerably more to a quality clinical doctoral program than just assessment research productivity or high publication counts (Heesacker & Elliott, 2007). For example, such research productivity does not necessarily translate into practical training in assessment, course offerings, or even program philosophy. It is important to underscore that these data are not intended to evaluate such programs as a whole—simply to help identify programs with particularly strong assessment concentrations. Furthermore, it is important to recognize that this study is a characterization of programs during a given period of time—but programs evolve and change, and a program with a strong assessment concentration at one point may look quite different 5 to 10 years later. Inspection of the list of leading programs in Table 2 is illustrative. A school at the top of the list, the University of Minnesota, has been recognized as a premier program in assessment since the 1930s and the era of Strong, Hathaway, and Meehl. However, other programs with strong traditions, such as the Stanford of Terman and Merrill, or Illinois during the Cattell era, have evolved toward different emphases and do not appear on this list, and many of the programs listed have developed a strong assessment focus relatively recently. Recognizing such limitations, these data are useful nonetheless as a characterization of the assessment community. The clinical assessment field is one clearly rooted in data, and this study at least provides the prospective scholar with some basis to select a program with strong opportunities in assessment research and training. Fortunately for the trainee and for the discipline, it appears that there are a number of impressive programs from which to choose.

REFERENCES

Andersen, B. L., Beck, S. J., Bornstein, R. A., Emery, C. F., Fristad, M. A., Kiecolt-Glaser, J. K., . . . Yeates, K. O. (2008). Problematic methods in the assessment of scholarly productivity in clinical PhD Programs. *Clinical Psychology Science and Practice, 15,* 102–104.

Gibby, R.E., Reeve, C.L., Grauer, E., Mohr, D., & Zickar, M. J. (2002). The Top I-O psychology doctoral programs of North America. *The Industrial-Organizational Psychologist, 39,* 17–25.

Heesacker, M., & Elliott, T. R. (2007). My dog's better than your dog: Publication counts and quality of clinical psychology training. *Clinical Psychology Science and Practice, 14,* 175–178.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *PNAS, 102,* 16569–16572.

Matson, J. L., Malone, C. J., Gonzalez, M. L., McClure, D. R., Laud, R. B., & Minshawi, N. F. (2005). Clinical psychology PhD program rankings: Evaluating eminence on faculty publications and citations. *Research in Developmental Disabilities, 26,* 503–513.

Stewart, P. K., Roberts, M. C., & Roy, K. M. (2007). Scholarly productivity in Clinical Psychology PhD programs: A normative assessment of publication rates. *Clinical Psychology: Science and Practice, 14,* 157–171.

Thomson Reuters. (2009). *2008 Journal Citation Reports, Social Science Edition*. Philadelphia, PA, and London, England: Author.

US News & World Report. (2008). *Health disciplines: Clinical Psychology (doctorate)*. Retrieved from http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-clinical-psychology-schools/rankings/.aq

Winter, J. L., Healy, M. C., & Svyantek, D. J. (1995). North America's top I-O psychology doctoral programs: *U.S. News and World Report* revisited. *The Industrial-Organizational Psychologist, 33,* 54–58.