Ranking Ph.D.-Granting Human Development and Family Studies Departments via Research and Professional Accomplishments

Alan Reifman, Ph.D.
Texas Tech University

Sylvia Niehuis, Ph.D.
Texas Tech University

Jacki Fitzpatrick, Ph.D.
Texas Tech University

Jennifer K. Chapman, Ph.D.
Eckerd College

C. Rebecca Oldham, Ph.D.
Middle Tennessee State University

Jean Pearson Scott, Ph.D.
Texas Tech University

Dan Fang, Ph.D.
GLG Quantitative Insights, Hong Kong

Shu Yuan, Ph.D.
Texas Tech University

Miriam Lieway, Ph.D.
Texas Tech University

Desiree Walisky, Ph.D.
Texas Tech Health Sciences Center

Leah Gregersen, BS, M.Ed.
Texas Tech University

Family Science Review, Volume 23, Issue 3, 2019

ABSTRACT. Rankings of departments' research productivity and professional accomplishments within an academic discipline can help prospective graduate students who are deciding where to apply, new professionals who are weighing job offers, and department chairs and other administrators who are assessing departments' progress and developing strategic plans. The field of Human Development and Family Studies (HDFS) has long lacked a multifaceted ranking system based on metrics of faculty research productivity and other departmental achievements. We introduce such a system to rank the 50 Ph.D.-granting HDFS departments in North America, using faculty members' publication and citation rates (*h* scores), grants received, and editorial-board memberships, along with departments' success in placing their Ph.D. graduates. These measures correlated well with one another and with previous peer-expert ratings of HDFS departments. However, some departments fared better under quantitative productivity measures than they did under peer ratings, and vice-versa. We discuss pros and cons of various ranking approaches.

Keywords: Faculty; *h*-statistic; Human Development and Family Studies departments; productivity; rankings

Direct correspondence to Alan Reifman, Ph.D., Department of Human Development & Family Studies, College of Human Sciences, Texas Tech University, Lubbock, Texas 79409-1230, United States, Phone: (806) 834-5174 Email: alan.reifman@ttu.edu

Ranking Ph.D.-Granting Human Development and Family Studies Departments via Research and Professional Accomplishments

Human Development and Family Studies (HDFS, also known as Human Development and Family *Sciences*, with dozens of other variations including the terms *family* and/or *development*; Hans, 2014) is a hybrid academic discipline going back more than 80 years. Hamon and Smith (2014) date the emergence of family science to 1880-1920, "when there was growing concern about the problems that families faced as a result of urbanization and industrialization and the need for social reform" (p. 310). Of the two major professional societies to which many HDFS scholars belong, the National Council on Family Relations (NCFR) was founded in 1938 (NCFR, 2019) and the Society for Research in Child Development (SRCD) emerged in stages between 1922 and 1933 (SRCD, 2019). Some of today's leading HDFS departments also date back to a similar era (University of Illinois, approximately 1900; University of North Carolina, Greensboro, phased in from 1920-1946; and Purdue University, 1946), based on historical information reported on these departments' websites. Of relevance to applied family science, HDFS departments provide undergraduate and (at some universities) graduate training to individuals who will work with families in community settings.[1]

Despite the longevity of professional societies and academic departments in HDFS, some in the field question its future robustness (Hamon & Smith, 2014; Hans, 2014). Concerns center on (a) articulating how HDFS is unique and distinct from disciplines such as psychology and sociology; (b) ensuring that HDFS faculty members' collaborations with scholars from other fields do not blur distinctions between disciplines; (c) enhancing awareness and understanding on the part of university administrators, policymakers, and the public of HDFS scholars' work and contributions; and (d) attracting undergraduate and graduate students to the field (Gavazzi, Wilson, Ganong, & Zvonkovic, 2014; Hamon & Smith, 2014; Hans, 2014). Not all HDFS scholars are pessimistic about the field's future, however. Many see interdisciplinary collaboration not as blurring the field's uniqueness but rather as fitting current expectations for how to obtain federal grant funding (Gavazzi et al., 2014). Nevertheless, even if the field is on stronger ground than skeptics contend, shoring up the aforementioned areas of concern would be helpful.

Importance of Research-Productivity Rankings

Systematic evaluation of programs' strengths and weaknesses is part of the process of bolstering HDFS departments. Hamon and Smith (2014) argue that "The need to 'know thyself' and evaluate the effectiveness of our programs in preparing graduates for a variety of employment and graduate school options has never been greater" (p. 312). We operate under the assumption that developing successful research scholars is among the major goals of Ph.D.-granting HDFS departments, along with developing effective teachers and individuals who will

[1] One could argue, based on the historical sequence reviewed above and the centrality of families to individual development, that family science is the core discipline with HDFS being just one type of name used for departments grounded in family science. However, other departments have titles referring only to human development (Cornell University, University of Pennsylvania [Penn], University of Rochester). Hence, we use the term HDFS throughout to reflect this variation in nomenclature.

contribute to welfare of families in other ways. Accordingly, we evaluated Ph.D.-granting HDFS programs' research productivity (e.g., publications, grants) and related professional accomplishments, while acknowledging that *research quality* should not be equated with *overall departmental or program quality*. Faculty members' teaching, mentoring of students, and university/community service, which we do not evaluate systematically in the present article but do address in the Discussion, also contribute to departmental quality with schools assigning different weights to research, teaching, and other contributions according to their missions, histories, and cultures. Departmental research-productivity rankings and the component scores comprising them (such as faculty publication and grant metrics) can be valuable for many purposes. We identify five purposes, arranged from micro to macro levels, where micro factors include interactions between faculty and students and macro include societal and governmental impacts, for example. First, faculty members can cite departmental rankings when advising promising undergraduate students on where to attend graduate school (the purpose of the *Philosophical Gourmet Report*, a ranking of academic philosophy departments; Bruya, 2015) or in advising students receiving Master's degrees about where to go for Ph.D. work. In particular, rankings will allow faculty mentors to best advise students on opportunities for graduate research experience and career placements. Second, faculty job candidates can consult departmental research rankings (in conjunction with universities' locations, departmental teaching loads, candidates' interests and aspirations, and so forth) when deciding where to accept job offers. Third (also at the departmental level) rankings can give chairpersons benchmarks for assessing their programs' research progress over time. Fourth, university administrators can use information contained in research-based rankings (in conjunction with teaching, mentoring, service, etc.) in developing standards for determining faculty members' tenure, promotion, and salary-increases. Fifth, with millions of dollars flowing to universities to support research (along with faculty positions, physical facilities, etc.), "Governments and funding authorities want to know how best to invest their resources and need to be able to differentiate among a large number of institutions" (Altbach, 2006, p. 2).

Previous Rankings of HDFS Departments

For a quarter century, from the work of Adams, Huston, Braeger, and Goff (1989) to that of Kamp Dush (2014), HDFS and family-science doctoral programs lacked a systematic, publicly available ranking of departments' academic and scholarly productivity. Kamp Dush (2014) produced rankings of Ph.D.-granting HDFS programs using *peer* or *reputational* ratings of departments. After identifying eligible departments (N = 51) through the National Council on Family Relations' (NCFR) degree program guide and consultation with colleagues, Kamp Dush (2014) sent surveys to the departments' graduate program directors. She varied the wording of the key item between respondents (determined randomly), asking them either to "Please evaluate the following *programs* on a scale of 1 Marginal to 5 Outstanding..." or "Please evaluate each *group of faculty..."* on the same scale, with faculty members in each department listed together as a group but without the name of the university. (The latter approach was based on the *Philosophical Gourmet Report*; see Bruya, 2015, for methods.) Means on the two survey versions were highly correlated and combined into an overall rating. Kamp Dush's (2014) rankings, published on her website, attracted great attention in the field. They also raised questions of whether some departments would have been ranked higher or lower had statistical

measures (e.g., publications, grants) been used instead of judges' subjective ratings. Statistical metrics, which Adams et al. (1989) used, may be more objective than judges' ratings. Whereas some judges may do extensive "homework" to investigate the accomplishments of each department's faculty members, others may rely on cues such as name-recognition of departments and the faculty members therein.

Measurement Strategies

Several experts argue for obtaining *multiple* measures of departments' scholarly quality rather than any one metric alone. DuPree and Bean (2018) contend that "because each metric [taken alone] favors certain factors over others…, the use of multiple metrics represents a more balanced evaluation and ranking process" (p. 7). Furthermore, whereas it is difficult to discern reasons for any single rater's judgments (Nisbett & Wilson, 1977), for a body of raters *collectively* one can test the degree to which departments' peer-based ratings correlate with quantitative indicators such as publications, citations, and grants (Ehrenberg & Hurst, 1996). Many different kinds of scholarly productivity metrics exist, ranging from rudimentary counts of a scholar's publications, citation counts, and external grants (Fairweather, 2002; Middaugh, 2001; Webber, 2011), to more complex statistics that incorporate multiple aspects of productivity. One popular multifaceted metric, the *h-index* (Hirsch, 2005), accounts for scholars' numbers of publications and how often these are cited. Ruscio (2016) identifies many strengths of the *h* index, including its being "at least as reliable and valid as its rivals" and "the index most robust to outliers or missing data" (p. 905), and its robustness to possible distortion by self-citations.

Furthermore, as Copes, Cardwell, and Sloan (2012) argue, such benchmarks should be *discipline-specific* and not generalized across different fields. Copes et al. note that "prior research suggests that *h*-index values vary by discipline… and are affected by such considerations as relative size of the discipline, number of publication outlets, and expectations on article production…" (p. 457). In addition, notes Ruscio (2016), *h* values may differ across fields due to "the number of people working in an area, the extent to which that work reaches beyond specialists, the size of research teams, and the typical number of references for publications in a certain field of study" (p. 907). Hence, *h* comparisons should be valid so long as they are made within a single discipline (here, HDFS). Alhough *h*, like any metric, is imperfect, we believe that alternatives to it are even more problematic. Raw publication counts, for example, do not account for articles' quality (as inferred from their citation frequency). Based on research by Ruscio (2016), we adopt *h* as a high-quality accessible metric of scholarly output.

The Present Study

In accordance with the above recommendations (e.g., DuPree & Bean, 2018), we conducted a multifaceted assessment. Our measures included departmental averages on faculty members' *h*-index scores, faculty grants received, faculty memberships on journal editorial boards, Ph.D.-alumni placements in faculty positions, and Kamp Dush's (2014) peer-based ratings of departments. Even though faculty editorial involvement and Ph.D. alumni placements do not measure research performance directly, these professional accomplishments are likely to

relate *indirectly* to research. A scholar's research eminence may be one factor leading to editorial-board invitations. Early publication output almost certainly helps new Ph.D.-holders obtain jobs. Our analytic approach was straightforward. First, we combined the above measures into a composite index via Principal Components Analysis (PCA), then listed the composite score for each department in the sample (with and without controlling for number of faculty in each department). Department size carries substantive and artifactual implications, in our view. Substantively, some observers may see advantages in larger departments, such as graduate students having more options for finding compatible mentors and upper administration appearing willing to fund faculty positions in HDFS. We refined our assessments over a four-year period (see our accompanying Professional Development and Reflection Corner article). Through this process, we pursued our aim of finding out which Ph.D.-granting HDFS departments would receive the highest rankings using new data from a multifaceted ranking system that complemented and broadened Kamp Dush's (2014) peer-rankings. Peer-ratings such as those that Kamp Dush (2014) collected may well correlate positively with departments' numbers of faculty members. However, on some of our statistical measures, especially departments' numbers of Ph.D. alumni placed onto other departments' faculty, having more faculty in one's own department presumably would lead to enrolling larger classes of graduate students. Some might argue that this artificially increases the department's chances of successful outplacement of students.

Methods

Sample and Procedure

Our study focused entirely on the stature of Ph.D.-granting HDFS departments in the 2014-15 academic year. We started from the departmental faculty lists in Kamp Dush's (2014) final report, which she obtained from departmental websites. Kamp Dush required departments to have a Ph.D. track focused on basic scholarship (i.e., research, teaching), rather than only a clinical/therapeutic track. She initially identified 52 programs, but dropped East Carolina University, which had only a clinical (medical family therapy) program. We later learned that the only HDFS graduate degree the University of Alabama awarded was the Master's and not the Ph.D., so we dropped it, resulting in retention of 50 departments.[2]

Combined, departments in this sample employed 767 tenured/tenure-track faculty members as of 2014-15, an average of 15.34 per department (SD = 7.42, range = 3-34). A major challenge was determining which faculty members listed on departments' 2014-15 websites were actually tenured or tenure-track faculty, our target group. For this purpose (and data-verification more generally), we contacted chairpersons of all 50 departments, receiving responses from 29 (58%). We describe our outreach to department chairs in the Professional Development and

[2] The University of Maryland-College Park has two departments that are potentially relevant to HDFS ranking studies, Family Science (in the School of Public Health) and Human Development and Quantitative Methodology (HDQM; in the College of Education). Kamp Dush (2014) included only the Family Science program, acknowledging that HDQM had been "mistakenly left off." Since 2014 peer-ratings of the Maryland HDQM program were not available, we likewise only included the university's Family Science department. Supplemental analyses on Maryland's HDQM program are available from the authors.

Reflection Corner article. Two considerations guided our decisions of whom to include in the sample. First, we required scholars' primary affiliation to be with a school's main HDFS department as opposed to a different academic unit or campus. Second, we included only those individuals who were expected to publish. Individuals with primary appointments outside HDFS would presumably devote most of their time and effort elsewhere, leaving many with partial or merely nominal affiliations to HDFS. Poling, LoSchiavo, and Shatz (2009) described potential ways in which faculty members' experiences at a branch or regional campus of a large university might differ from those of their main campus counterparts. According to these authors, branch-campus faculty typically would have higher teaching loads, less expectations from administrators to publish, and no opportunities to work with graduate students. The degree to which these statements apply to HDFS departments with branch campuses likely varies university by university. To treat departments as consistently with each other as possible, therefore, we excluded branch-campus faculty. We also excluded lecturers/adjuncts and staff members. Inclusion of those not expected to publish could artificially depress a department's average research productivity. We retained extension faculty so long as there appeared to be expectations for them to publish research (ascertained by seeing if they had research publications and via communication with department chairs).

Measures

Faculty members' publication records (*h*). Using the HDFS faculty lists, we sought individuals' *h* statistics from the online program *Publish or Perish* (PoP; Harzing, 2007). A scholar's *h* value is the largest possible number satisfying the condition that he or she has *h* articles that have each been cited at least *h* times (e.g., an *h* of 10 represents a scholar publishing 10 articles that have each been cited 10 or more times). PoP links into established online databases and computes *h* and other publication statistics. We selected Google Scholar, a free, easily accessed database that also offers better coverage of scholarly publications than do other databases (Ruscio, 2016). Whereas PoP provides quick and (apparently) accurate results for scholars with uncommon last names, it sometimes brings up numerous excess entries beyond one's target for scholars with common surnames. We therefore employed alternative strategies for some scholars (e.g., seeing if they had created their own Google Scholar profile, which reports *h*; using an online vita, if available, to check Google Scholar for citation counts of individual articles and compute *h* by hand). We took extra caution in seeking *h* statistics for scholars with hyphenated or compound last names (e.g., Claire Kamp Dush or Sarah Schoppe-Sullivan), searching under the hyphenated/compound name and each component of the last name alone. Furthermore, for faculty not appearing on our original lists but who were on their department's faculty in 2014-15 (according to department chairs), we obtained *h*-statistics for them in 2018 but needed to prorate them to what they (approximately) were in 2014-15. Based on average career-trajectories of *h* provided in Copes et al. (2012), we decided to reduce assistant professors' 2018 *h*-scores by 1.5 and full professors' by 2, to estimate 2014-15 *h*-scores (there were no associate professors whose scores needed to be adjusted). An age-weighted variation on *h* (AW-Index) accounts for articles' years since publication, potentially evening-out departmental differences in junior/senior faculty distribution. During preliminary stages of this project we examined the correlation between *h* and AW scores and found it to be $r = .94$. We therefore stopped recording AW.

Faculty members' grant receipt.

Our measure of grant receipt focused on U.S. federal funding agencies and large foundations.[3] We devised three categories of grant receipt, entailing funding from (a) the National Institutes of Health (NIH); (b) other U.S. federal research funding, such as the National Science Foundation, U.S. Department of Education, and U.S. Department of Agriculture; and (c) private foundations. In our view, NIH is broad enough to serve as its own category because it subsumes many agencies pertinent to HDFS (National Institute of Aging, National Institute of Alcohol Abuse and Alcoholism, National Institute of Child Health and Human Development, National Institute of Drug Abuse, and National Institute of Mental Health). Our team recorded NIH grant data from the organization's website, whereas we asked department chairs to compile all other forms of funding their faculty members received. We required grants to be worth at least $50,000 per annum during the funding period (comparable to an NIH R03 award, a familiar type of grant in many disciplines) and that the faculty member be the Principal Investigator (PI). The latter requirement would help avoid duplicate counting if, for example, the PI and co-PI were from the same department. We counted grants from 2010-2015 to provide what we thought was a broad enough window to capture faculty members' grant funding from one or more sources. However, we decided against going back further (e.g., 20 or 30 years), as doing so seemed likely to favor arbitrarily departments with greater numbers of senior faculty. Each faculty member could earn up to three points, one for obtaining at least one NIH grant, one for obtaining at least one grant from another federal agency, and one for obtaining a foundation grant. We gave considerable weight to federal research funding as, in the words of Craig and Lombardi (2012), "In the US… the federal research funding competition serves as a major determinant of university research prestige because it is primarily peer reviewed and the competition is conducted on an open nationwide basis" (p. 4). To be inclusive, we included any grant with a research component, even if the project had additional goals (e.g., service-delivery).

After reviewing each faculty member's grant-receipt, we computed mean points for each department. This system took no note of a single faculty member receiving multiple grants within a category (e.g., NIH) or dollar amounts. Doing so might have made the grants measure more sensitive than what we selected, but also carried the potential for outliers in dollar amounts. Also, limiting how much any one faculty member's (potentially prodigious) grant-receipt could contribute to his or her department's score would do more to reward departments who had multiple faculty members receiving grants. We received full grant receipt data (on federal and foundation grants from department chairs, and NIH data from the web) from only 29 of the 50 departments in our sample. Because we did not want to disadvantage non-responding schools, we derived new estimates for the missing schools. Specifically, we regressed our aggregate of NIH, other federal, and foundation grant statistics (i.e., *full data*) on NIH-only grant data (which we had for all 50 schools) in the 29 departments for which both data-points were available. We then used the resulting regression equation to generate *predicted* full-grant-data for the 21 departments with only NIH scores, which we used as a substitute.

[3] For universities in Canada, department chairs had the opportunity to provide us with lists of their faculty members' grants (typically from Canadian funding sources) when we contacted them to verify our data. The chair of one Canadian department provided this information. Based on details provided (i.e., funding sources, amounts), we assigned scores similarly to how we did for U.S. universities.

Faculty members' editor/editorial-board positions.

Our team initially recorded editorships and editorial-board memberships of HDFS faculty in our sample for eight leading journals, all of which are flagships of leading professional associations: *Journal of Marriage and Family, Journal of Family Theory and Review, Family Relations, Journals of Gerontology, Developmental Psychology, Journal of Family Psychology, Journal of Research on Adolescence,* and *Child Development.* We then attempted to broaden the scope of journals covered (another suggestion from the 2015 NCFR conference), to include cognate fields such as Marriage and Family Therapy, and Family and Consumer Sciences, as well as human-development or family-studies journals beyond those listed above. We gathered data on faculty members' editorial roles beyond the eight original journals via the request letter to department chairs. We categorized each faculty member as having at least one editorship or editorial-board membership or not; doing so avoided outlier values from scholars who were on many boards and rewarded departments with editorial involvement from multiple faculty members. Finally, we computed each department's percentage of faculty with at least one editorial role. Since we received department chairs' editorial-board updates from only 27 of the 50 departments, we conducted a similar regression analysis as described above for grant-receipt data to ensure no missing data on departments' share of faculty members with editorial roles.

Departments' placement of Ph.D. alumni.

Our criterion was each department's number of Ph.D. graduates placed onto the faculty of another department in our sample (i.e., a Ph.D.-granting HDFS department) as of 2014-15.[4] Three clarifications regarding inclusion criteria are in order. First, one can distinguish two types of placement: from a scholar's graduate school to his or her *first* academic position, and from graduate school to a scholar's 2014-15 location (regardless of any earlier positions). Initially, based on departments' websites and faculty members' online vitae, we obtained more data (n = 344) for placements focused on scholars' 2014-15 location than for purely first-job placements (n = 262). Hence, we used placements focused on 2014-15 location. Second, we did not count a department's hiring of its own Ph.D. graduates.[5] Third, colleagues at many schools suggested to us that placements of their alumni in non-HDFS departments (e.g., psychology, education) or in other venues (e.g., community agencies) should reflect merit on the Ph.D.-granting departments. Fourteen departments, in response to our letters to department chairs, provided *broad-placement* lists, detailing alumni employment in Ph.D.-granting and non-Ph.D.-granting HDFS departments, departments of other academic disciplines, community or industrial settings, and

[4] Oprisko, Dobbs, and DiGrazia (2013) proposed a measure of *placement efficiency* representing "the number of tenured and tenure-track placements of Ph.D.-graduates per tenured and tenure-track faculty at a given Ph.D.-granting institution" (p. 3). Our measure, when we control for departments' faculty size, is similar.

[5] Some research shows that individuals who began their faculty careers at the same university at which they received their Ph.D.s are less successful at publishing academic articles than are individuals whose Ph.D.-granting and faculty-hiring institutions are different (Horta, Veloso, & Grediaga, 2010). This finding is consistent with the conjecture that internal hires may be judged less stringently than external ones, although it does not prove it. Further, the widespread use of seemingly pejorative terms for internal hiring in the literature (*academic inbreeding* produced nearly 900 hits on Google Scholar and *faculty inbreeding* produced over 200) suggests that the practice is frowned upon by many observers.

post-doctoral fellowships. One department reported 184 alumni placements. We did not include broad placements in our main analyses due to the small number of departments reporting them, but conducted supplementary analyses with them (available from the authors).

Results

Correlations among Study Variables

We first examined correlations among our major variables pertaining to Ph.D.-granting HDFS departments in 2014-15 (Table 1). With the exception of editorial-board memberships/editorships, our measures of departmental productivity correlated well (*r*'s from .38-.70 for correlations of average faculty *h* with grant-receipt, Ph.D. alumni-placement, and Kamp Dush [2014] peer-rating scores; and Kamp Dush scores with grant-receipt and alumni-placement). These correlations support the convergent validity of these measures as reflecting a common construct of departmental productivity. Number of faculty in a department showed strong positive correlations with *h*, alumni-placements, and Kamp Dush (2014) scores, warranting additional analyses to control for department size. Indeed, correlations between *h* and Kamp Dush scores, placements and Kamp Dush scores, and *h* and placements were all reduced appreciably when partialling out numbers of faculty. However, the two correlations involving Kamp Dush scores remained above .50. The correlation between grant-receipt and Kamp Dush scores actually went up slightly when number of faculty was controlled.

Generation of Composite Research-Productivity Scores

We conducted Principal Components Analysis (PCA), a similar technique to factor analysis, to group all (or most) of the productivity metrics into a single scale or dimension. Further, PCA would allow metrics most strongly correlated with each other to have greater weight in constructing a departmental productivity/achievement index. A one-component solution seemed most appropriate, with the first component exhibiting an eigenvalue of 2.53 and accounting for 50.6% of the variance in the measures (the second component had an eigenvalue of 0.93 and accounted for 18.6% of the variance). As shown in the left-hand column of Table 2, all five measures (*h*, grant-receipt, editorial roles, alumni-placement, and 2014 Kamp Dush scores) loaded strongly on the one component, all reaching .45 or higher and three of the loadings attaining values of .74 or higher. We retained a composite index based on these weightings. We repeated the PCA on the same five measures, but purged of shared variance with number of faculty in a department. (Each of the five measures was regressed on department size, with residual scores saved.) The PCA on the five purged indicators yielded similar loadings (Table 2) to those in the original analysis, with the eigenvalue and percent of variance accounted for only slightly lower.

Table 1
Means (with Standard Deviations) and Correlations of Departmental Characteristics

	Mean (SD)	1	2	3	4	5	6
1. No. of faculty	15.34 (7.42)	---					
2. Mean *h*	14.36 (6.75)	.43**	---				
3. Mean grant receipt[a]	.41 (.25)	.06	.42**	---			
		(n/a)	(.44***)				
4. Editorial positions[b]	.47 (.13)	.08	.22	.25	---		
		(n/a)	(.20)	(.24)			
5. No. of alumni placed	5.76 (6.54)	.54***	.38**	.19	.20	---	
		(n/a)	(.19)	(.19)	(.19)		
6. Kamp Dush (2014) rating[c]	3.10 (.67)	.64***	.63***	.39**	.26	.70***	---
		(n/a)	(.51***)	(.46***)	(.27^{+})	(.55***)	

Parenthesized correlations are with departments' numbers of faculty members partialled out. $N = 50$. $^{+}p < .06$, $*p < .05$, $**p < .01$, $***p \leq .001$. [a]The maximum score a faculty member could receive was 3. [b]Proportion of faculty members holding at least one editorial role. [c]The maximum score a department could receive was 5.

Table 2
Weights for Principal-Component Scores Used to Generate Single Composite (Value Adjusted for Department Size in Parentheses)

Mean *h*	.78 (.71)
Grant receipt	.61 (.70)
Editorial positions	.46 (.49)
Alumni placements	.74 (.62)
Kamp Dush (2014) score	.89 (.86)
Eigenvalue/% of Variance	2.52/50.5% (2.35/47.0%)

Final Rankings

Finally, we ranked departments on the five-indicator composite scores (both unadjusted and adjusted for department size). Table 3 presents departments' original and adjusted rankings.[6] The Pennsylvania State University (Penn State) topped rankings under both methods. Using unadjusted PCA scores, Penn State's composite score (3.14) exceeded runner-up Cornell's score (1.86) by a sizable margin. Cornell ranked second and fourth in our unadjusted and adjusted analyses, respectively. Other than at the very top of the list, many of the differences between departments involved small fractions of a point, so readers should not overinterpret instances of one school ranking slightly higher than another school. In other words, there are groups of departments whose faculty productivity and other academic markers are similar, so that prospective graduate students should not have to worry about going to great lengths to enroll in a program that ranks one place above another. Schools' composite scores are on a *z*-score metric,[7] which may aid interpretation of the magnitude of differences between departments. For example, Syracuse and Connecticut each had composite scores at or near 0.00, placing them almost exactly in the middle of the distribution of departments. Hypothetically, if these schools could increase their composite scores to 0.50 in a future ranking study, they would increase their percentile by nearly 20 percentage points (from the 50th percentile to the 69th). For schools with a composite score of 1.00 (e.g., Auburn), increasing the score to 1.50 would move them from the 84th percentile to the 93rd (i.e., moving ahead of roughly five schools they were currently behind). Controlling for departments' numbers of faculty had fairly dramatic effects in some instances. Arizona State, with 28 faculty members, dropped from third place with no adjustment for department size to 16th with adjustment. Conversely, the University of Arizona, with nine faculty members, jumped from seventh to second.

[6] Each HDFS department's average or percentage scores on separate component metrics (*h*, grant-receipt, editorial-board memberships, etc.) are available from the authors so long as the original data were in the public domain (e.g., grant data from federal-government funding agencies' websites).

[7] The PCA procedure first standardizes each of the separate metrics (*h*-scores, grant-receipt, editorial-board memberships, alumni placement, and Kamp Dush ratings), then computes a weighted composite based on the Component Score Coefficient Matrix (IBM Knowledge Center, not dated).

Table 3
Rankings of Ph.D.-Granting Human Development and Family Studies Programs in 2014-15, Using Multiple-Measure Composite (Schools Listed from Highest to Lowest Ranking on Unadjusted Composite, with Actual Scores Shown in Parentheses; Schools with Similar Scores Clumped Together)

Department (No. Faculty)	Rank (Score)	Rank Adjusted for Size
Penn State (34)	1 (3.14)	1
Cornell (25)	2 (1.86)	4
Arizona State (28)	3 (1.48)	16
UNC-Greensboro (21)	4 (1.23)	10
Minnesota (20)	5 (1.18)	8
Illinois (20)	6 (1.075)	9
Arizona (9)	7 (1.074)	2
Tufts (13)	8 (1.06)	3
Auburn (17)	9 (1.01)	7
Oregon State (15)	10 (.97)	5
Virginia Tech (18)	11 (.95)	12
California-Davis (14)	12 (.86)	6
Purdue (19)	13 (.85)	15
Northwestern (29)	14 (.81)	27
Georgia (15)	15 (.75)	11
Colorado State (16)	16 (.485)	19
Texas-Austin (13)	17 (.483)	13
Maryland (Family Studies; 15)	18 (.43)	18
Missouri (18)	19 (.33)	23
Wisconsin-Madison (11)	20 (.22)	17
Michigan State (22)	21 (.20)	29
Syracuse (10)	22 (.00)	20
Connecticut (22)	23 (-.01)	36
Oklahoma State (23)	24 (-.07)	39
Pennsylvania (4)	25 (-.09)	14
Florida State (11)	26 (-.10)	21
Iowa State (23)	27 (-.26)	43
Texas Tech (16)	28 (-.277)	33
Ohio State (12)	29 (-.280)	25

Kansas State (10)	30 (-.30)	24
Utah State (23)	31 (-.31)	44
Tennessee (13)	32 (-.40)	28
Brigham Young (33)	33 (-.42)	49
Rochester (13)	34 (-.63)	37
Brandeis (4)	35 (-.64)	22
Guelph (11)	36 (-.82)	38
Manitoba (10)	37 (-.84)	35
North Dakota State (13)	38 (-.88)	41
Alberta (5)	39 (-.91)	26
Kentucky (11)	40 (-.97)	40
Montclair State (16)	41 (-.99)	46
West Virginia (6)	42 (-1.02)	31
Louisiana State (5)	43 (-1.05)	30
Nebraska (17)	44 (-1.13)	47
North Texas (5)	45 (-1.20)	34
New Mexico (3)	46 (-1.25)	32
Mississippi State (7)	47 (-1.32)	42
Texas Woman's (22)	48 (-1.40)	50
Delaware (18)	49 (-1.438)	48
Loma Linda (9)	50 (-1.441)	45

Discussion

The present study sought to expand the scope of HDFS departmental research-productivity rankings, from Kamp Dush's (2014) peer-ratings to multifaceted statistical indices, to see which Ph.D.-granting HDFS departments from 2014-15 would rank atop the field. In the following sections we summarize our findings and examine implications for various actors in the academic enterprise: administrators, faculty, and graduate students.

Summary of Findings

Many schools, such as Penn State and Cornell, ranked similarly regardless of whether analyses controlled for departments' number of faculty members. Other schools' rankings shifted after controlling for faculty size (e.g., Arizona State and Arizona). Comparing our rankings to Kamp Dush's (2014) peer-ratings, one program fared better under quantitative metrics: Tufts, an elite private university. In contrast, programs that ranked higher via peer than quantitative metrics largely were at Midwestern land-grant universities. For example, the University of Illinois ranked sixth and ninth in the present analyses, but were number two in Kamp Dush

(2014). Some observers may consider these Midwestern schools more traditional homes for HDFS programs, so these departments may register more strongly in evaluators' minds.

Implications for Evaluating Faculty and Departmental Productivity

One question emerging from our research is whether it is better to evaluate departments' research productivity via quantitative metrics (e.g., based on publications, grants) or via peer-expert ratings. One could argue that peer-ratings such as those used by Kamp Dush (2014), while not duplicating metrics such as *h*-scores and grant-receipt, still have enough statistical overlap with them (i.e., correlations around .60 and .70) to make the peer ratings quite useful (Sullivan, 2015). These high correlations suggest that the judges in Kamp Dush's (2014) study were well attuned to faculty members' research output across the set of HDFS departments. In fact, given the hefty time commitment involved in searching for *h* statistics, grant-funding records, and so forth, one may conclude that the information obtainable from just one method (i.e., peer-rating surveys) needs no augmentation from additional assessments. However, since we observed some discrepancies between departments' rankings arising from different methods (see above), researchers interested in departmental research productivity and achievement will need to determine what investigative methods to use in light of their own goals, time, and resources.

Although having more faculty members may artificially enhance some markers of departmental research productivity and achievement, this effect generally did not appear large. Departments that accomplish a great deal with relatively few faculty members did well in analyses controlling for numbers of faculty. Whereas we commend these departments for their industriousness, it may still be better for students to seek graduate training in larger departments to increase their chances of finding compatible mentors. Of course, prospective students also need to match their interest areas (e.g., childhood, adolescence, adulthood, marriage, parenting) to their departments' areas of strength.

Implications for Departments and Career Development of Faculty and Graduate Students

Our data and analyses have considerable applied value for academic administration, training, and development. University department heads (and other administrators) can use the information herein to identify specific, objective strengths and weaknesses of their respective units' research productivity. This information, which is not available with a reputation-based ranking alone, can be useful as part of strategic planning to improve departmental quality and ranking. To increase the output of faculty members who are deemed not to be publishing at a suitable level, administrators can offer them advising, training, and resources (e.g., workshops and software pertaining to new research methods and statistics). A large multidisciplinary study identified several direct and indirect predictors of faculty members' publication and presentation volume over the past three years (Hardré, Beesley, Miller, & Pace, 2011). These predictors included departmental support and self-efficacy as a researcher (positively) and teaching load (negatively).

Administrators can also use the kinds of faculty-productivity measures we have described to allocate sabbaticals and even revamp departments' course curricula when it seems appropriate

to reducing faculty members' teaching loads. Some departments may need or wish to undertake deeper, more systemic initiatives to improve their research cultures (Holosko & Barner, 2016). According to Holosko and Barner, elements conducive to building and strengthening a research culture include the presence of "senior leaders and professors who mentor students and early career scholars toward achieving ongoing careers in scholarship and research activities" (p. 282), infrastructure, resources, and faculty buy-in. We did not analyze our results separately by faculty rank, but future research could do so to identify senior faculty in each department who publish and seek grants with junior faculty colleagues and graduate students, thus helping to enhance their department's research culture.

Our results would likewise be useful as one piece of input for individuals having to choose among HDFS departments, namely prospective graduate students and faculty job candidates. Prospective graduate students, particularly those seeking appointments in academia, can research departments' academic placement records for their graduates, along with faculty publication and grant activity. The latter pieces of information would likely suggest which graduate departments offered the greatest opportunities for students to publish with faculty members and perhaps to gain experience in grant preparation. As shown in the Table 1 correlations, the three variables most associated with departments' successful outplacement of Ph.D. recipients were number of faculty in the department, faculty members' publication productivity (*h*), and reputational ratings (Kamp Dush, 2014). Therefore, beyond quantitative metrics such as publications and grants, coming from large, well-regarded departments appears likely to help Ph.D. graduates' job placement. Mentors can incorporate the above considerations within their advice to students seeking graduate training (Kamp Dush, 2015; Lefkowitz, 2018). New Ph.D. recipients (and their more senior colleagues considering a career move) would also be able to evaluate faculty job offers from a given department in terms of expectations they would be likely to face regarding publication, grant activity, and journal editorial board memberships.

Limitations and Suggestions for Further Research

Several potentially complicating features suggest a need for further, more refined analysis in advancing the study of HDFS rankings. Different departments have family studies versus human development emphases, so judges' academic specialty could affect their ratings. Some HDFS departments also include Marriage and Family Therapy (MFT); others do not. We included MFT faculty in our analyses if listed within HDFS departments, but future studies should devise methods to maximize comparability in faculty training across departments.[8] Also, departments go through cycles of being top-heavy (i.e., mostly senior faculty) or bottom-heavy, so rankings in any one year (or even for a few years) may not represent a department's long-term quality. Our findings are limited to the 2014-15 academic year. However, to the extent departments have retained most of the same faculty for an extended period (and possibly added highly accomplished junior faculty along the way), top-ranking departments in the present article would still be high-level. Regarding the metrics we used, external grant receipt is a commonly

[8] A few departments include other Human Science fields such as Textiles and Financial Planning within HDFS departments. We included data from only the HDFS faculty members in these departments.

used marker of faculty productivity (Middaugh, 2001), but not all departments prioritize federal and other external grants, which harms them in our rankings.

Our scope of faculty performance was also somewhat narrow. Whereas faculty publishing (especially with students) is one part of strong graduate training, other features not measured here – such as excellent faculty teaching and mentoring, student research and teaching experience, and professional socialization – are also important. Future research could systematically investigate faculty performance in Ph.D.-granting HDFS departments with regard to teaching, mentoring, and professional socialization. Indeed, examples for assessing faculty members in these roles exist in the literature, albeit in fields outside HDFS. Ishiyama, Miles, and Balarezo (2010) reviewed catalogue information online for 122 Ph.D.-granting political sciences programs in the U.S., assessing whether a given program offered a formal course or seminar on teaching political science and, if offered, whether the class was required or an elective. As with the need for multifaceted measures of departmental research productivity discussed above, it would also be desirable also to have multiple measures of teaching preparation. The present authors' university (Texas Tech) offers the Teaching Effectiveness And Career enhancement (TEACH) program. This consists of workshops and teaching observations and is available on a competitive basis to Ph.D. students across all disciplines who aspire to college teaching careers. Universities with HDFS programs could be assessed on whether they offer focused teacher preparation similar to the TEACH program. Another measure of HDFS departments' teaching effectiveness would be whether faculty members had won teaching awards, either at their home universities or in the field at large. In the area of mentoring, Berk, Berg, Mortimer, Walton-Moss, and Yeo (2005) introduced a set of survey measures, asking students to rate the extent to which their mentor "was accessible," "was supportive and encouraging," and "challenged me to extend my abilities (e.g., risk taking, try a new professional activity, draft a section of an article)." Surveying doctoral students in all HDFS graduate programs would be a more extensive undertaking than relying solely on data available online (e.g., presence of courses on teaching, faculty teaching awards) but would yield rich information. Should there be interest in doing so, researchers could compile multiple measures of departments' teaching and mentoring quality, parallel to what we developed for departments' research productivity.

Finally, the 58% response rate among department chairs is a possible concern more specific to our methods than are the above limitations. Two arguments rebut this concern. First, some non-responding chairs may have believed that the information we had already gathered for their department was accurate and needed no updating. Second, we adjusted for chairpersons' non-response as best we could via regression estimates. Therefore, the chairs' response rate may be less of a concern than it initially seems.

Conclusion

In our view, the breadth of our measures was very strong even though it did not encompass every conceivable element of departmental productivity and graduate training. Metrics such as faculty publications, citations, and grant activity have face validity as measures of academic productivity, whereas those schools that ranked highly under our methods (Penn State, Cornell) also add face validity. Furthermore, different measures of academic productivity and achievement correlated well, lending convergent validity. Metrics for teaching and

mentoring could be expanded and used for future studies of HDFS departmental quality. Although our focus was on graduate programs, a system analogous to that of *U.S. News and World Report's* for ranking undergraduate programs, consisting of a weighted composite of peer and statistical indicators, may also be most appropriate for HDFS.

Alan Reifman is a Professor in the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Sylvia Niehuis is an Associate Professor in the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Jacki Fitzpatrick is an Associate Professor in the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Jennifer K. Chapman is an Assistant Professor in the Department of Human Development at Eckerd College, St. Petersburg, FL 33711.

C. Rebecca Oldham is an Assistant Professor in the Department of Human Sciences at Middle Tennessee State University, Murfreesboro, TN 37132.

Jean Pearson Scott is a Professor (Retired) in the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Dan Fang is a Ph.D. recipient from the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Shu Yuan is a Ph.D. recipient from the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Miriam Lieway is a Ph.D. candidate in the Department of Human Development and Family Studies at Texas Tech University, Lubbock, TX 79409-1230.

Desiree N. Walisky, L.C.S.W., is the Director of Research, Policy, and Development and a Child and Family Therapist at the Center for Superheroes, Texas Tech University Health Sciences Center, Lubbock, TX 79413, and Ph.D. candidate in Educational Psychology & Leadership at Texas Tech University, Lubbock, TX 79409-1071.

Leah Gregersen is an educator and businessperson in the Lubbock, Texas area.

References

Adams, G. R., Huston, D., Braeger, T., & Goff, I-A. K. (1989). Family science: A study of departmental and faculty eminence. *Family Science Review, 2*, 249-266.

Berk, R. A., Berg, J., Mortimer, R., Walton-Moss, B., & Yeo, T. P. (2005). Measuring the effectiveness of faculty mentoring relationships. *Academic Medicine, 80*, 66–71.

Bruya, B. (2015). Appearance and reality in the Philosophical Gourmet Report: Why the discrepancy matters to the profession of philosophy. *Metaphilosophy, 46*, 657–690. doi: 10.1111/meta.12161/full

Copes, H., Cardwell, S., & Sloan, III, J. J. (2012): *h*-index and m-quotient benchmarks of scholarly impact in criminology and criminal justice: A preliminary note. *Journal of Criminal Justice Education*, 23, 441-461. doi: 10.1080/10511253.2012.680896

Craig, D. D., & Lombardi, J. V. (2012). Measuring research performance: National and international perspectives (pp. 3-13). In J. V. Lombardi, E. D. Phillips, C. W. Abbey, & D. D. Craig (Eds.), *The top American research universities: 2012 annual report*. Tempe, Arizona: The Center for Measuring University Performance. Retrieved from https://mup.asu.edu/sites/default/files/mup-pdf/MUP-Publication-2012-Measuring-Research-Performance-National-and-International-Perspectives.pdf

DuPree, W. J., & Bean, R. A. (2018). Evaluating scholarship productivity in COAMFTE-accredited doctoral programs: An update. *Journal of Marital and Family Therapy.* Published online first at: doi: 10.1111/jmft.12318

Ehrenberg, R. G. & Hurst, P. J. (1996). *The 1995 NRC ratings of doctoral programs: A hedonic model* (NBER Working Paper Series No. 5523) [Electronic version]. Cambridge, MA: National Bureau of Economic Research.

Fairweather, J. S. (2002). The mythologies of faculty productivity: Implications for institutional policy and decision making. *The Journal of Higher Education, 73*, 26-48. doi: 10.1080/00221546.2002.11777129

Gavazzi, S. M., Wilson, S. M., Ganong, L., & Zvonkovic, A. (2014). Furthering the conversation on the discipline of Family Science: Comments on the articles by Hamon & Smith and Hans. *Family Relations, 63*, 333–342. doi: 10.1111/fare.12076

Hamon, R. R., & Smith, S. R. (2014). The discipline of Family Science and the continuing need for innovation. *Family Relations 63*, 309–322. doi: 10.1111/fare.12068

Hans, J. D. (2014). Whither famology? Department name trends across four decades. *Family Relations, 63*, 323–332. doi: 10.1111/fare.12075

Hardré, P. L., Beesley, A. D., Miller, R. L., & Pace, T. M. (2011). Faculty motivation to do research: Across disciplines in research-extensive universities. *The Journal of the Professoriate, 5*, 35-69.

Harzing, A.W. (2007) *Publish or perish* [software]. Available from http://www.harzing.com/pop.htm

Hirsch, J. E. (2005, 15 November). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science, 102*, 16569–16572. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1283832/

Holosko, M. J., & Barner, J. R. (2016). Research productivity in top-ranked schools in psychology and
social work: Research cultures do matter! *Research on Social Work Practice, 26*, 278-285.
doi: 10.1177/1049731514549815

Horta, H., Veloso, F. M., & Grediaga, R. (2010). Navel gazing: Academic inbreeding and scientific productivity. *Management Science, 56*, 414-429. doi: 10.1287/mnsc.1090.1109

IBM Knowledge Center (not dated). *Component Score Coefficient Matrix (SPSS).* Armonk, New York: IBM. Retrieved from https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/tutorials/fac_cars_compsccoeffmtrx_01.html

Ishiyama, J., Miles, T., & Balarezo, C. (2010). Training the next generation of teaching professors: A comparative study of Ph.D. programs in political science. *PS: Political Science & Politics, 43*, 515-522. doi: 10.1017/S1049096510000752

Kamp Dush, C. (2014). *The HDFS Report 2014: Claire Kamp Dush's ranking of HDFS programs in North America.* Retrieved from http://u.osu.edu/adventuresinhdfs/the-hdfs-report-2014/

Kamp Dush, C. (2015, March 11). Where should I go to graduate school? *Adventures in Human Development and Family Science.* Retrieved from https://u.osu.edu/adventuresinhdfs/2015/03/11/where-should-i-go-to-graduate-school/

Lefkowitz, E. (2018, October 16). How to choose a graduate program? *Developmental aspects of sexual health* [Blog]. Retrieved from https://www.evalefkowitz.com/blog/how-to-choose-a-graduate-program

Middaugh, M.F. (2001). *Understanding faculty productivity: Standards and benchmarks for colleges and universities*. San Francisco: Jossey-Bass.

National Council on Family Relations (2019). *NCFR history*. Retrieved from https://www.ncfr.org/about/history

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259. doi: 10.1037/0033-295X.84.3.231

Oprisko, R. L., Dobbs, K. L., & DiGrazia, J. (2013, September 15). Placement efficiency: An alternative ranking metric for graduate schools. *Georgetown Public Policy Review*. Retrieved from http://gppreview.com/2013/09/15/placement-efficiency-an-alternative-ranking-metric-for-graduate-schools/

Poling, D. A., LoSchiavo, F. M., & Shatz, M. A. (2009). Regional campus success: Strategies for psychology faculty. *Teaching of Psychology, 36,* 194–196. doi: 10.1080/00986280902959788

Ruscio, J. (2016). Taking advantage of citation measures of scholarly impact: Hip hip *h* index! *Perspectives on Psychological Science, 11*, 905-908. doi: 10.1177/1745691616664436

Society for Research in Child Development. (2019). *The roots of SRCD.* Retrieved from https://www.srcd.org/about-us

Sullivan, J. (2015). *Discussant remarks at session "Ranking HDFS departments: A look at publication statistics."* Annual conference of the National Council on Family Relations, Vancouver, British Columbia.

Webber, K. L. (2011). Measuring faculty productivity. In J. Shin, R. Toutkoushian, & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 105-121). New York, NY: Springer Science & Business.