




Prominent but Less Productive: The Impact of Interdisciplinarity on Scientists' Research^{*}

Administrative Science Quarterly
2017, Vol. 62(1)105–139
© The Author(s) 2016
Reprints and permissions:
sagepub.com/
journalsPermissions.nav
DOI: 10.1177/0001839216665364
journals.sagepub.com/home/asq


Erin Leahey,¹ Christine M. Beckman,² and
Taryn L. Stanko³

Abstract

Federal agencies and universities in the U.S. promote interdisciplinary research because it presumably spurs transformative, innovative science. Using data on almost 900 research-center-based scientists and their 32,000 published articles, along with a set of unpublished papers, we assess whether such research is indeed beneficial and whether costs accompany the potential benefits. Existing research highlights this tension: whereas the innovation literature suggests that spanning disciplines is beneficial because it allows scientists to see connections across fields, the categories literature suggests that spanning disciplines is penalized because the resulting research may be lower quality or confusing to place. To investigate this, we empirically distinguish production and reception effects and highlight a new production penalty: lower productivity, which may be attributable to cognitive and collaborative challenges associated with interdisciplinary research and/or hurdles in the review process. Using an innovative measure of interdisciplinary research that considers the similarity of the disciplines spanned, we document both penalties (fewer papers published) and benefits (increased citations) associated with it and show that it is a high-risk, high-reward endeavor, one that partly depends on field-level interdisciplinarity.

Keywords: organizational ecology, organizational innovation, scientific careers, creativity, interdisciplinary research

Because of its expected benefits to science and society (Sanz, Bordons, and Zulueta, 2001; Rhoten and Parker, 2006), scholars increasingly engage in an

^{*} The first two authors are equal contributors.

¹ School of Sociology, University of Arizona

² Robert H. Smith School of Business, University of Maryland

³ Orfalea College of Business, California Polytechnic State University

interdisciplinary mode of research, which “integrates perspectives, information, data, techniques, tools, concepts, and/or theories from two or more disciplines” (National Academies of Science, National Academy of Engineering, and Institute of Medicine, 2005: 188). Though the practice of interdisciplinarity is not new (Abbott, 2001), it is increasingly prevalent in the natural sciences (Rhoten and Pfirman, 2007) and social sciences (Brint, 2005; Jacobs and Frickel, 2009). Universities are reorganizing to facilitate interdisciplinary research by developing cross-disciplinary problem-focused centers and funding cross-department and cross-college research initiatives (Pray, 2002; Biancani, McFarland, and Dahlander, 2014). And since the mid-1980s, the National Science Foundation (NSF) has supported cross-cutting funding opportunities and interdisciplinary research centers. Scientists laud interdisciplinary research as a progressive hot topic that is running rampant; arguably one must be interdisciplinary to be world-class (Pray, 2002, quoting Irwin Feller). But evidence in support of this contention is sparse, and there is “relatively little research on many of the underlying issues” (Jacobs and Frickel, 2009: 44). In particular, systematic investigation of interdisciplinarity’s effects on scientific careers has been neglected.

Two strands of work in organizational theory are useful in understanding the professional costs and benefits of engaging in interdisciplinary research. The first documents the benefits of bringing together distinct ideas across domains, an idea developed from research on recombinant innovation (Fleming, 2001; Schilling, 2005; Singh and Fleming, 2010): pooling non-redundant information from disparate sources is the foundation from which novel ideas spring (Weitzman, 1998; Hargadon, 2002). This work suggests that bridging disconnected knowledge spaces results in better ideas that will be rewarded in the marketplace (Lo and Kennedy, 2015). The second documents penalties associated with category spanning (Zuckerman, 1999), showing that rather than being perceived as innovative, offerings spanning multiple domains have an ambiguous identity that is difficult for audiences to understand and are thus devalued (Hsu, Hannan, and Kocak, 2009).

Most empirical work emphasizes the negative perceptions of the audience or market, what we may call reception-side penalties. But the category literature has also theorized production-side penalties, suggesting that category-spanning products are more difficult to produce: investing in multiple categories limits mastery and dilutes quality, resulting in a “Jack of all trades” who is master of none (Hannan and Freeman, 1989; Hsu, 2006b). Recent empirical work has found that category-spanning products are lower quality, likely because the production process itself is wanting (Kovács and Johnson, 2014). This contributes to but does not completely explain the reception-side devaluation of category-spanning products (Negro and Leung, 2013; Kovács and Johnson, 2014). From this we gather that production- and reception-side effects can operate simultaneously.

On the production side, penalties that accrue to category-spanning offerings can result not only from limited mastery and inferior quality, which are highlighted in the categories literature (Hsu, Hannan, and Kocak, 2009; Negro and Leung, 2013), but also from the cognitive and collaborative challenges of category-spanning work, which are highlighted in the scholarship on science (Fleming and Sorenson, 2004; Cummings and Kiesler, 2005; Schilling and Green, 2011; Wagner et al., 2011). Grasping ideas and perspectives from

another field is cognitively taxing and time-consuming, working with diverse collaborators from multiple disciplines can produce frustration and conflict, and reviewers may have difficulty digesting and evaluating cross-disciplinary products (Boix Mansilla, Lamont, and Sato, 2016). All of these challenges may lengthen the time to publication and thereby depress scholars' productivity.

On the reception side, work that draws on disparate intellectual domains should have broad appeal and achieve greater scholarly visibility. Atypical and novel combinations of ideas have greater impact (Schilling and Green, 2011; Uzzi et al., 2013; Larivière, Haustein, and Börner, 2015), so scientists—like Pontikes' (2012) market-makers—may be drawn toward, rather than confused by, multi-category offerings like interdisciplinary research. Yet experimenting with new combinations is also uncertain and risky (Fleming, 2001; Foster, Rzhetsky, and Evans, 2015): it may sustain benefits overall but also experience a higher variance in reception. Interdisciplinary research should boost the overall visibility of scholarly work—the mean levels of citations to it—but also increase the variance of citations that a scientist's body of work garners.

Finally, the penalties and benefits of interdisciplinary work likely depend on the nature and intellectual life cycle of one's field. The difficulties associated with category spanning may be reduced, and the benefits may be accentuated, when category spanning is popular (Lo and Kennedy, 2015) or growing in popularity. In fields with a tradition of interdisciplinary research, training may reduce the difficulties of producing it. In fields that are becoming more interdisciplinary, audiences may particularly value such work. Thus the level of and trend toward interdisciplinarity at the field level may modify the main effects of interdisciplinary research on productivity and visibility.

Our paper makes contributions to organizational theory and refinements to the measurement of interdisciplinarity. First, we distinguish between, and assess the simultaneous effects of, production side and reception side processes. Second, we highlight production penalties that result from cognitive and collaborative challenges (Cummings and Kiesler, 2005; Boix Mansilla, Lamont, and Sato, 2016) and disentangle potential mechanisms that manifest in different stages of the production process. Third, we begin to understand the normalization of category spanning by examining an important, contextual-level moderator of such effects: field-level interdisciplinarity. And last, we move away from a binary conceptualization of "spanning" (Fleming, Mingo, and Chen, 2007; Hsu, Hannan, and Kocak, 2009) toward a continuous conceptualization. Following Leahey and Moody (2014) and other work on categories (Kovács and Johnson, 2014; Kovács and Hannan, 2015), we account for the cognitive distance between the categories (here, disciplines). Below, we test our ideas using primary data collected on almost 900 scientists and their 32,000 scholarly papers.

PENALTIES AND BENEFITS OF INTERDISCIPLINARY RESEARCH

Production Penalties

Even in this era of broad enthusiasm for interdisciplinary research (NAS, 2005), disciplinary fields remain an entrenched way in which scholarly activity is organized (Sá, 2008) and evaluated (Lamont, 2009), making field-spanning offerings unexpected. Like other categories, fields are not arbitrarily constructed but

rather reflect the distinct environments that disciplines face, the formal and tacit skill sets that members acquire (Zuckerman et al., 2003), and the distinct claims of jurisdiction that occasionally come into dispute (Abbott, 2001). Disciplinary fields are categories that help academics parse and digest a vast intellectual terrain in order to conduct and evaluate scholarship.

From the scholarship on categories, we know that when categories are not strictly adhered to, penalties ensue. Zuckerman et al. (2003) suggested that penalties emanate from both the production side (e.g., it's harder to produce, and quality may suffer) and the reception side (e.g., it is confusing for audiences to place a category-spanning offering). Recent empirical tests have focused mostly on the reception penalties: audiences tend to overlook, devalue, or outright reject offerings that span categories because they are difficult to comprehend and do not fit existing schemas (Hsu, Hannan, and Kocak, 2009). Category-spanning actors and offerings tend to be poorly received: feature-film actors who take on too many diverse roles have difficulty obtaining work (Zuckerman et al., 2003); films spanning multiple genres are less appealing to audiences (Hsu, 2006a); and eBay sellers who try to market their product in multiple categories are less successful in the auction (Hsu, Hannan, and Kocak, 2009). Audiences penalize offerings that are difficult to classify. Following Leahey and Moody (2014), we extend this scholarship to the realm of science.

In science, less-examined production penalties may result from the cognitive and collaborative challenges that slow the development of mastery. Because cognitive resources are finite, the argument goes, mastery may be more difficult to attain when time, energy, and effort are distributed across many categories (Freeman and Hannan, 1983; Carroll, 1985; Hannan and Freeman, 1989; Hannan, Pólos, and Carroll, 2007). This early work in population ecology theorized but did not empirically examine the problems of diffused resources and focus that may disadvantage generalists in production, including greater competition and the likelihood of mortality. Recent work does not examine mastery per se but suggests that it accounts for lower-quality products (Negro and Leung, 2013; Kovács and Johnson, 2014). In applying these ideas to the realm of science, we hew closer to the concept of mastery itself and argue for technical and logistical hurdles when spanning domains (Zuckerman, 2005; Lingo and O'Mahony, 2010). For example, research on patents suggests that category spanning increases technological uncertainty and creates cognitive limitations (Fleming, 2001; Lo and Kennedy, 2015). Lamont, Mallard, and Guetzkow (2006) lent support to the idea that category-spanning ideas are harder to produce: it is challenging for scholars to accommodate the concepts of multiple fields and to produce output that is standard in form and content (see also Wagner et al., 2011; Kaplan, Milde, and Cowan, 2016).

In addition, because it is typically collaborative, interdisciplinary research may produce epistemological or methodological conflict (Murray, 2010; van Rijnsvoever and Hessels, 2011) between members of different fields. The literature on scientific collaboration suggests that this lack of consensus increases coordination costs (Cummings and Kiesler, 2005, 2007; Shrum, Genuth, and Chompalov, 2007), conflict and tension (Owen-Smith and Powell, 2003; Murray, 2010), and role strain (Boardman and Bozeman, 2007), which may outweigh the general benefits of collaborative work (Gans and Murray, 2015); for a recent review see Leahey (2016). Because commonality and consensus speed

the implementation of ideas (Beckman, 2006; Boix Mansilla, Lamont, and Sato, 2016), such interdisciplinary divergences and tensions may slow progress toward publication.

These initial production penalties, experienced as scientists plan, conduct, and write up their research, likely follow the paper into a second stage of production: the review stage. In the peer-review process, experts from different fields often disagree on the merits of a paper and evaluate them differently (Boix Mansilla, 2006; Lamont, 2009). Birnbaum (1981) found that research that does not fit neatly within the substantive bounds of “normal science” instills irritation, confusion, and misunderstanding among reviewers and editors. This makes the road to publication challenging, and the final product may take longer to produce. A similar phenomenon occurs in the patent context: approval times are longer when inventions span (or in their language, blend) categories (Lo and Kennedy, 2015). Thus the review process may be slower and more challenging for interdisciplinary work, contributing to lower rates of productivity.

Even among highly successful interdisciplinary scholars, we find some initial evidence that production penalties result from cognitive challenges. In interviews the first author conducted with recipients of the Mellon Foundation’s New Directions fellowship (who, post-tenure in one field, received formal training in another field), we find evidence that interdisciplinary research takes additional time, commitment, and effort (McBee and Leahey, 2016). It’s mentally more taxing: “I’ll be reading epistemology and political philosophy . . . and then do research and mathematical things. It’s just a bit of a stretch of the brain to do such things.” It takes longer: “I think part of me taking longer to go from conception to a finished published article has to do with trying to think through two separate disciplinary concerns.” In the end, it can take a toll on one’s productivity: “I just don’t feel like I’ve made great strides of improvement in being as productive as I could be.” Interviews with these scholars, who almost all hail from the humanities where sole-authored work is still the norm, suggest that scholars who are more engaged with interdisciplinary work—alone or in collaboration with others—experience hurdles that dampen productivity. Thus we hypothesize a production penalty:

Hypothesis 1 (H1): Interdisciplinary research is associated with lower productivity.

Reception Benefits

Although category-spanning products are typically discounted or dismissed, we do not expect interdisciplinary research to incur a reception penalty—that is, to be received poorly in the academic marketplace (i.e., rarely cited). As Pontikes (2012) demonstrated, research documenting penalties has largely been conducted on “market-takers”—audiences that are looking to identify, place, and evaluate a product. To market-takers, multiple categories—and ambiguous classification more generally—may produce confusion and prevent a timely and effective search. In contrast, “market-makers” are looking to identify novel, original, and groundbreaking research that redefines the (academic) market structure (Pontikes, 2012). This distinction helps reconcile the reception penalties noted in the categories literature with the reception benefits noted in the innovation literature, such as Fleming’s research documenting a citation bonus to patents that span patent classes and subclasses (Fleming, 2001; Singh and

Fleming, 2010). Market-makers see ambiguous classification not as confusing but as appealing, and perhaps as an indication of innovative work that has brought insights from one arena to illuminate another.

The scholarship on scientific innovation documents the benefits that accrue to joining domains of knowledge. Research on innovation suggests that information pooled from disparate sources provides a foundation—perhaps *the* foundation—from which new ideas spring (Hargadon, 2002; Fleming and Waguespack, 2007). This link is so tight as to warrant use of the term “recombinant innovation” (Weitzman, 1998), meaning that bridging knowledge domains and developing new combinations serves as the foundation for innovation (Schumpeter, 1912; Weick, 1979) and will be recognized as such. Singh and Fleming (2010) found that a research team’s experience diversity and network size predict breakthrough patents. Research on the impact of category spanning in science has shown that atypical, category-spanning offerings have higher impact (Shi et al., 2009; Schilling and Green, 2011; Uzzi et al., 2013; Leahey and Moody, 2014; Larivière, Haustein, and Börner, 2015; Lo and Kennedy, 2015). Interdisciplinary publications, as a form of atypical, domain-spanning publications, likely experience these same benefits.

In academe, positive reception is typically gauged via citations (Merton, 1977; Evans, 2008; Lynn, 2014), although there is variation in what a citation signals (van Dalen and Henkens, 2005). Citations may reflect disciplinary alliances and mutually reinforcing citation practices, instrumental attempts to flatter potential reviewers, or even the controversial nature of an article. Authors may cite previous work in a casual way or rely on it heavily. They may think of it highly or dismiss it as flawed (Ferber, 1986; Latour, 1987; Najman and Hewitt, 2003; Lynn, 2014). Yet citations give us a sense of how a paper is received by the scientific community. Among other things, reference to a scholarly work indicates its usefulness and influence because it contributed, in some way, to a subsequent work. A scientist’s citation count gives a good indication of his or her visibility in the scientific community, and for this reason it is factored into promotion and tenure decisions and has been shown to influence earnings (Diamond, 1986; Sauer, 1988; Leahey, 2007). One goal of this paper is to empirically test for such a reception benefit: whether scholarly visibility (as indicated by citation counts) accrues to published interdisciplinary work, even after controlling for the size of the prospective audience, which increases when multiple audiences are targeted. We predict:

Hypothesis 2 (H2): Interdisciplinary research is associated with higher visibility.

We expect to find empirical support for both a production penalty (H1) and a reception benefit (H2), and it is likely that interdisciplinary research is a high-risk, high-reward activity that sometimes gains great prominence but often does not. Even if interdisciplinary research positively affects visibility (H2), it likely increases the probability of both breakthroughs (like achieving scientific prominence) and failures (like being ignored). This is what Hsu, Negro, and Perretti (2012) found in their study of film: films that span genres have a higher likelihood of exceptional success compared with films that do not. In the more relevant realm of science, Foster, Rzhetsky, and Evans (2015: 892) found that biomedical articles that make new connections among chemicals have greater “uncertainty of reward,” and Singh and Fleming (2010) documented a similar

effect among patents. Similarly, in the social sciences, “unfamiliar or atypical combinations of knowledge yield novel outcomes with greater variance in performance” (Schilling and Green, 2011: 1322). Thus, although on average interdisciplinary research will be well received, we also expect interdisciplinary work to have greater variability in citations:

Hypothesis 3 (H3): Interdisciplinary research is associated with greater variability in visibility.

Finally, we are distinguishing between spanning categories and spanning distant categories. Most investigations simply assess whether boundary spanning has occurred (Clemens et al., 1995; Fleming, Mingo, and Chen, 2007; Hsu, Hannan, and Kocak, 2009) without considering the relationship between the spanned entities. The category literature is just beginning to consider distances between categories (Rosenkopf and Almeida, 2003; Schilling and Green, 2011; Kovács and Johnson, 2014; Kovács and Hannan, 2015), as is the scholarship on domain spanning in science (Wagner et al., 2011; Uzzi et al., 2013; Leahey and Moody, 2014; Larivière, Haustein, and Börner, 2015; Aharonson and Schilling, 2016). Are the entities cognitively similar, like civil and chemical engineering? Or are they cognitively distant, like geography and optics? Interdisciplinary research can be more or less novel, depending on the relationship between the spanned fields (Carnabuci and Bruggeman, 2009: 608), and spanning two closely related categories is hardly different from not spanning at all. In terms of penalties, the cognitive and coordination challenges associated with category spanning likely increase as the distance between domains increases. In terms of benefits, the utility and value of category-spanning research also increases with distance: the most fertile creative products are “drawn from domains that are far apart” (Poincaré, 1952), and the best conceptual metaphors are those that create ties across great distances (Knorr Cetina, 1980). We thus hypothesize that distance matters:

Hypothesis 4 (H4): An alternate measure of interdisciplinary research that does not incorporate distance will have weaker effects on productivity and visibility.

Intellectual Context and Interdisciplinary Research

We expect the penalties and benefits associated with interdisciplinary research to be shaped by each scientist’s intellectual context, that is, the interdisciplinary nature and dynamism of his or her primary field. Interdisciplinarity varies across fields. For example, we know from previous research that some fields, like the life sciences, are more interdisciplinary than other fields, such as electrical engineering. Analyses of the Survey of Earned Doctorates (Millar and Dillman, 2012) reveal that the life sciences have the highest proportion of dissertations self-classified as interdisciplinary (36 percent), especially compared with engineering and math (26 and 21 percent, respectively). Using a continuous measure of interdisciplinary research that incorporates cognitive dissimilarity (distance), Porter and Rafols (2009) and Porter et al. (2007) found that biotechnology (mean = .654, on a scale of 0 to 1) and medicine (mean = .664) are more interdisciplinary than electrical engineering (mean = .53). From this, we suggest that field-level characteristics should influence how a piece of

scholarship is received. Actors tend to value what they are accustomed to and what is familiar (McPherson, Smith-Lovin, and Cook, 2001; Phillips, 2011; Trapido, 2015). We see this in science: Lo and Kennedy (2015) found that audience members and patent examiners are more favorable when category spanning is more familiar and typical. On the reception side, this suggests that highly interdisciplinary fields, like the life sciences, will expect and appreciate interdisciplinary work, whereas less interdisciplinary fields will view it as unnecessary or unusual. On the production side, scholars in highly interdisciplinary fields may experience fewer challenges, as there are models for guidance, potentially better training in interdisciplinary research, and a more amenable and productive review process for it, whereas scholars in less interdisciplinary fields may experience greater difficulties producing it.

Yet we also know that fields change over time and at different rates (Jacobs and Frickel, 2009). According to Whitley (2000), not all fields are equally entrenched in disciplinary thought, and fields may change at different rates as a result of various internal and external factors. In fields moving toward interdisciplinary research, actors should be more receptive to this work because it is on the leading edge of a new trend. Thus scholars in increasingly interdisciplinary fields should experience greater reception benefits for their interdisciplinary work. Challenges associated with interdisciplinary research may also be tempered when a field is changing. For example, Ruef and Patterson (2009) examined the emergence of the credit rating system and found fewer penalties associated with category spanning when the system is in flux. If one's disciplinary environment is conducive to (and increasingly pursuing) interdisciplinary research, then the benefits of it might be stronger and the penalties less stiff:

Hypothesis 5a (H5a): In highly interdisciplinary and increasingly interdisciplinary fields, interdisciplinary research's positive association with visibility will be stronger than in less interdisciplinary fields or more static fields.

Hypothesis 5b (H5b): In highly interdisciplinary and increasingly interdisciplinary fields, interdisciplinary research's negative association with productivity will be weaker than in less interdisciplinary fields or more static fields.

METHODS

Sample

To test our hypotheses, we identified all scientists associated with 52 industry/university-cooperative research centers (IUCRCs) funded by the National Science Foundation. These centers are housed at universities and conduct research that is of interest to (and partly supported by) industry in areas as diverse as biosurface physics, civil and environmental engineering, and architectural science. Research centers are an integral part of universities' and federal agencies' efforts to promote interdisciplinary research (Sá, 2008; Biancani, McFarland, and Dahlander, 2014), so they are ideal sites for our investigation of its effects. Center affiliation is relatively common: almost one-third (32 percent) of faculty at Research I universities nationwide are affiliated with a research center (Boardman and Corley, 2008). But because previous research suggests that scientists affiliated with research centers tend to be more active in research (Corley and Gaughan, 2005), including collaborative (Gaughan and

Ponomariov, 2008) and interdisciplinary research (Ponomariov and Boardman, 2010), we empirically assessed the representativeness of our sample by comparing it with the population of scientists (captured in NSF's Scientists and Engineers Statistical Data System, SESTAT) and the population of papers (using the Web of Science corpus). To preview results from these robustness checks, we find that our sample is representative of scientists working in Research I universities, except in terms of gender, which may be expected given that the IUCRCs we sampled are based predominantly in male-dominated fields. And their work is no more (or less) novel than the population of published papers.

For each scientist, including doctoral students, post-docs, and faculty at all career stages, we collected archival publication data as of 2005. Our analysis is limited to the subset of 854 Ph.D.-level scientists with a publication record as of 2005 because interdisciplinary research, productivity, and visibility cannot be computed for scientists who have not published. Over 80 percent of the scientists are in science or engineering fields, with the other 20 percent coming from math, computer science, or the social sciences. Information on approximately 32,000 articles published by these center affiliates, such as number of authors, institutions represented, journal of publication, and number of citations, came from Thomson Reuters' Web of Science (WoS). We obtained data on all the articles referenced by these 32,000 focal articles, notably their WoS subject categories (SCs), which we used to measure the extent to which each scholar's research is interdisciplinary. To these data we added field-, university-, and individual-level data from various sources. Our main analyses involve these archival data, but for supplementary analyses, we relied on publicly available data on journal turnaround times and a primary survey of these center-based scientists. Because our sample is, like most quantitative studies of science, restricted to the public record (i.e., published articles), we collected a sample of unpublished working papers to assess the generalizability of our data and results.

Dependent Variables

Productivity. To capture each scholar's productivity, we relied on the total number of articles published in WoS journals from the beginning of a scholar's career—the year he or she first published—until 2005. This is a conservative measure of productivity that excludes book chapters and articles published in less internationally recognized journals; it is likely more accurate than self-reported productivity used by studies that rely on survey data, such as the Survey of Earned Doctorates. Using article counts as a measure of productivity is standard in the literature on scientific productivity. Article counts remain standard even as collaboration increases, because this measure is highly correlated with coauthor-weighted publication counts (Cole and Zuckerman, 1984; Wagner-Dobler, 1997). We did, however, assess the robustness of our results to this alternative measure. We dismissed journal impact factor (JIF) weighted productivity measures because they confound quantity with quality—the two distinct outcomes that we expect interdisciplinary research to affect differentially. We examined the number of articles at the person and person-year level in our analyses.

Visibility. We measured visibility by collecting the (forward) citations that had accrued to each published article indexed in WoS as of 2010. Citations to an individual paper are a more precise measure of a paper's impact than the prestige of its journal of publication; however, we controlled for journal impact factor in all models. In addition to this paper-level measure, we aggregated by taking the mean to obtain person- and person-year-level measures of visibility for certain analyses. To test hypothesis 3 about variability in citations, we calculated the standard deviation of each scientist's citations.

Independent Variables

Interdisciplinary research (IDR). Our measure of interdisciplinary research is borrowed from Porter and colleagues (2007: 134). We measured it at the paper level, aggregating up to the person or person-year level when relevant by taking the mean IDR score across each scientist's set of papers. Compared with other measures of diversity, Porter's measure of "integration" incorporates not only the variety (i.e., number) of categories and their balance (i.e., the evenness of the distribution) but also their similarity (i.e., their cognitive distance) into one index (Rafols and Meyer, 2010). The categories of interest are the 244 WoS subject categories (examples include sociology, chemical engineering, and organic chemistry); the Web of Science assigns 1–6 SCs to each indexed journal, which we extended to each constituent reference. As input for the measure of interdisciplinary research, we pooled all SCs from the focal paper's set of references, rather than SCs of the focal paper itself or of papers that cite the focal paper, an approach that best gauges knowledge integration (Porter et al., 2007: 127). This captures the breadth of research referenced and presumably integrated in a paper. Although the variety of SCs and their balance is specific to each focal paper's bibliography, their similarity (s_{ij} , from a SC \times SC co-citation matrix, which we convert to cosines) is derived from the population of all WoS-indexed articles and thus shared by all focal papers.¹ Porter's measure is a particular parameterization of the Sterling Index:

$$1 - \sum_{ij} s_{ij} p_i p_j$$

where s_{ij} is the similarity between SCs i and j , p_i is the proportion of referenced papers in subject category i , and p_j is the proportion of referenced papers in subject category j (Rafols and Meyer, 2010: 267–268). We demonstrate the interdisciplinary research calculation for three hypothetical articles in Online Appendix A (<http://asq.sagepub.com/supplemental>). Intuitively, a paper's interdisciplinary research score increases as it references more relatively unrelated SCs (Porter et al., 2007: 277).² The interdisciplinary research score ranges from 0 to 1, with scores closer to 1 indicating greater interdisciplinarity. For paper and person-year analyses, the IDR score is zero for 551 papers and 99 person-

¹ Off-diagonal elements in a SC \times SC matrix for the year 2007 would indicate the number of papers that cited both SCs (i and j) between 2002 and 2006. To construct our SC \times SC matrix, we summed three square matrices for the years 1987, 1997, and 2007 for the person-level analyses. We used the matrix closest to the year of publication for the person-year-level analyses.

² The interdisciplinary research score is only mildly ($r = .17$) related to the number of papers referenced. The reported results are unchanged with and without this control.

years. These values occur when the person has a very low productivity for that year or has a single paper that references only one subject category. Given that these low productivity years obscure the production penalty that we hypothesize for interdisciplinary research papers, we restricted all analyses to exclude observations with zero scores. The visibility analyses are robust to including these observations; the productivity effects are curvilinear with this inclusion because interdisciplinary research scores of zero occur only in low productivity years, and high scores are also associated with low productivity (as we hypothesize). For the person-year analysis, we used the mean interdisciplinary research score of papers published in that year to predict productivity.

Though a number of other measures of interdisciplinary research have been proposed and used in the literature, only a few recent papers consider the relatedness of the categories that are brought together (Larivière, Haustein, and Börner, 2015; Yegros-Yegros, Rafols, and D'Este, 2015). Most institution-level measures, such as the number of disciplines that members represent (Birnbaum, 1981), person-level measures, such as whether an individual faculty member has a joint appointment (Jacobs and Frickel, 2009), and publication-level measures fail to consider the relatedness of fields. For example, Clemens and colleagues (1995: 454) assessed whether the paper is cited in a discipline other than the discipline the author(s) represented. Larivière and Gingras (2010) assessed the percentage of a paper's cited references made to journals from other fields without considering the relatedness of those fields—for example, whether a sociology paper references a cognate field like anthropology or a more distant field like geology. The same limitation characterizes other research on knowledge products that span categories, including Hsu and colleagues' (2009) research on films and Fleming and colleagues' (2007) research on patents. In contrast, we relied on Porter's measure of interdisciplinary research, which incorporates distance or dissimilarity between every pair of fields, and—to assess the extent to which distance matters above and beyond sheer variety (H4)—we compared it with the effects of a less-refined measure: the total number of unique subject categories (SCs) appearing in a scholar's pooled set of references.

We ascertained the validity of this measure in various ways. First, by showing how the interdisciplinary research score increases as a scientist references not only more, but more-unrelated fields, Online Appendix A provides face validity. Second, the interdisciplinary research score, which is based on the variety, evenness, and distance of referenced SCs, is also positively correlated ($r = .18$) with the number of SCs that characterize the focal paper itself (recall that a given paper is assigned 1–6 SCs). This is to be expected if a focal paper's SCs adequately capture the content of a paper that is more thoroughly gleaned from an analysis of the SCs it references, and it demonstrates convergent validity. Third, as we should expect, the mean IDR scores of articles published in the interdisciplinary journals *Science*, *Nature*, and *PNAS* (.71, .71, and .70, respectively) are significantly higher than the average IDR score in our sample (.62)—more than half a standard deviation above. Thus we are reassured that the measure we used adequately captures the concept of interest to us.

Field IDR. We used both continuous and categorical measures of field-level interdisciplinary research. Both rely on current field classification as determined

by scholars' CVs and department faculty listings; the nine broad fields include physical sciences; life sciences; math; most types of engineering—chemical, electrical, material, and other; computer science; and social science. Results are robust to classifying based on field of doctoral study. Continuous measures include the (1) *average IDR score* for each field and (2) *trend in IDR* for each field, calculated as the slope from a field-specific regression of IDR on publication year. The categorical measures indicate whether the scholar is in the life sciences (the highest IDR field in our sample) and whether the scholar is in electrical engineering (one of the lowest IDR fields in our sample). Data from the Survey of Earned Doctorates, using self-reports on the interdisciplinary nature of one's dissertation, confirm these as high (life sciences) and low (electrical engineering) interdisciplinary fields. Our data further corroborate this: 22 percent of the life science journals represented in our sample are interdisciplinary—i.e., their subject category is “multidisciplinary”—compared with only 7 percent of the electrical engineering journals. We assessed field differences in the effects (H5) in two ways. First, we assessed whether IDR's effects are modified by field-level trends in IDR by including an interaction term in the person-level model. Second, we compared two subsets of scholars: those working in a high-IDR field (life sciences) and those working in a low-IDR field (electrical engineering).

Control Variables

The scientific papers we analyzed are nested within persons, so when comparing publications across persons, we controlled for characteristics at this level, including gender, professional age, and status. Gender and professional age have been shown to influence engagement with IDR (Boix Mansilla, 2006; van Rijnsoever and Hessels, 2011; Dahlander and Frederiksen, 2012), as well as productivity and visibility (Leahey, 2006; Maliniaka, Powers, and Walter, 2013). Status is important to control for because it has been shown to influence how a scientist's work is received (Azoulay, Stuart, and Wang, 2014). The gender of each scientist was derived from analysis of first names, as well as information (pictures and pronouns) used on the scientists' websites. Professional age was calculated by subtracting the year of Ph.D. receipt (obtained from CVs and Proquest Dissertation Abstracts) from the publication year of each article; in person-level models we used professional age as of 2005. To proxy status of current institution, we used the number of faculty who are members of the National Academy of Sciences, collected from The Center for Measuring University Performance (<https://mup.asu.edu>). To measure the quality of an individual scholar, we included the ranking of one's Ph.D.-granting institution, obtained from the Academic Ranking of World Universities (<http://www.arwu.org>) and reverse-coded so that higher values indicate higher quality. Although certainly a rough measure of individual quality, it helps to rule out concerns about “smarter” scholars being more (or less) likely to engage in IDR research.

The papers we analyzed are also nested within journals and fields, so we incorporated controls at these levels. Because IDR has been linked with the prestige of the journal in which it appears (Rinia, van Leeuwen, and van Raan, 2002), we controlled for the mean impact factor of the journals each scholar has published in. When modeling productivity, we controlled for average turnaround time for journals in the field, obtained from Bjork and Solomon (2013).

When modeling visibility, we controlled for average citations per paper in the field, obtained from Thomson Reuters. Because the relationship between interdisciplinary research and citation counts depends on the field (Hamilton, Narin, and Olivastro, 2005; Zitt, Ramanana-Rahary, and Bassecoulard, 2005; Larivière and Gingras, 2010), and because we wanted to rule out increased audience size as an alternative explanation for heightened visibility, we controlled for field size: the number of Ph.D.s produced in a recent year, obtained from the NSF's Survey of Earned Doctorates. In addition, we controlled for the potential reach of each paper (Fleming and Sorenson, 2004: 919), captured by the number of SCs that classify the focal paper itself, rather than its references. We presume that papers classified in multiple fields will be brought to the attention of a larger body of scholars, potentially boosting citations; thus this is another control for audience size. Although our sample is center-based, we did not include center-level controls because none reached statistical significance and the intra-class correlation coefficient (.11) suggested minimal clustering by center.

In paper-level and person-year analyses, we controlled for additional variables that might influence visibility and/or productivity. Teams produce more highly cited papers, on average, than sole authors (Wuchty, Jones, and Uzzi, 2007; Montpetit, Blais, and Foucault, 2008), and authors who have worked together previously may face fewer production penalties than new collaborative teams. Thus we included the number of authors on the paper and a binary variable indicating whether this combination of authors had published together before; in yearly analyses, these were measured as the average number of authors and the proportion of papers with repeat collaborators, respectively. We also controlled for both lagged cumulative publication experience and lagged cumulative citations, logged to remedy skewness. We controlled for publication year in paper-level analyses, as older publications have more time to accrue citations than more recent publications. Finally, we controlled for number of references in models of visibility because papers with a longer reference list may garner more citations.

Statistical Approach

Interdisciplinarity is measured at the paper level, but some scientists' research outcomes, such as productivity, cannot be measured at that level, so we present models at the paper, person-year, and person level of analysis. This allows us to model outcomes at the most appropriate level, incorporate level-specific control variables, and assess the robustness of our results.

Our first analysis takes place at the person level ($N = 854$ scientists). We relied on path analysis, a type of structural equation model (SEM) that contains only observed, but no latent, variables (Bollen, 1989). This technique has been used in previous studies of scholars' productivity and visibility (e.g., Leahey, 2007). Here, path-analytic techniques are ideal for two reasons. First, we have two main outcomes of interest—productivity and visibility—and path analysis allows us to model these endogenous outcomes simultaneously, in the same model. We do not need to run separate models with the two outcomes of interest. In fact, in a single model we can assess the determinants of these outcomes, as well as the relationship between the two outcomes themselves. Second, path models allow us to examine direct as well as indirect effects—of, for example, $IDR \rightarrow productivity \rightarrow visibility$ —and thus explicitly model the

influence of intervening variables, unlike a regression-based mediation analysis. For ease of interpretation of SEM output, we partitioned the table of results into separate sections, which show the effects on each endogenous outcome separately (IDR, productivity, and visibility). Like regression, model specification is driven by theory, and different variables can influence the different outcomes. Because we compared alternative outcomes and subsamples, we did not focus on comparing the fit of different model specifications (e.g., full versus restricted models) but present Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics for this purpose. The structural equation modeling (*sem*) package in Stata 13, which we used to estimate path models, is ideal for handling missing data. Rather than deleting data in a listwise fashion, which is the default strategy in most statistical packages, this package relies on a full information maximum likelihood estimation procedure. This strategy permits the inclusion of all available data (Anderson, 1957) and bypasses the need to impute data.

We also took advantage of our data structure—panel data over the career of individual scientists—to estimate models at the paper and person-year levels of analysis. This allowed us to assess the robustness of the person-level results, to better assess causal direction and examine the effects of paper-level variables. A Hausman test suggests a fixed-effects model is more appropriate than a random effects model ($\chi^2 = 405.57$; $\chi^2 = 382.90$; $p < .0001$ for productivity and visibility, respectively). We estimated fixed-effects regression models at the person-year level ($N = 8,525$)—because a person's productivity cannot be captured at the paper level—but also at the paper level ($N = 29,775$) when modeling visibility. The fixed-effects models allowed us to look at within-person variation on our dependent variables and control for unmeasured person-level characteristics and thereby provided the most conservative test of our hypotheses. We used multiple imputation because of missing data for some variables, but results are the same without it.

RESULTS

Table 1 describes our data and measures for the sample of 854 published scientists. Table 2 provides descriptive statistics for all variables used in the person-year-level models. There is ample variation on the key outcomes, visibility and productivity, both of which are right-skewed when aggregated to the person level, so we took the natural log, as previous researchers have done (Allison and Long, 1987; Prpic, 2002; McBrier, 2003), or used negative binomial models when appropriate (tests indicate negative binomial is preferred to Poisson given overdispersion). The key predictor, IDR, also displays variation—ranging from .08 to .88 on a scale of 0 to 1. To detect problems of multicollinearity, we calculated the variance inflation factors (VIF). In all models the VIF scores were below six, well below the recommended cutoff value of ten (Neter, Wasserman, and Kutner, 1985). In the section on robustness checks below, we demonstrate the representativeness of our sample by comparing our sample of papers with the population of WoS papers analyzed by Uzzi et al. (2013) and comparing our sample of scientists with the population of Ph.D.-level scientists from the NSF's surveys.

As expected (H1), IDR depresses scholarly productivity (see table 3, model 1); this effect is statistically significant at the 5-percent level. Recalling that the

Table 1. Correlations and Descriptive Statistics for Published Scientists (N = 854)

Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11
1. Visibility (total # citations)*	930	1779											
2. Productivity (total # articles)†	46	60	.83										
3. Interdisciplinarity (mean IDR score)†	.63	.10	.04	−.05									
4. Mean journal impact factor (JIF)	1.68	1.21	.37	.21	.25								
5. Professional age (years since Ph.D.)	24	11	.22	.40	−.17	−.12							
6. Ranking of Ph.D.-granting institution	2.84	1.70	.12	.10	.05	.09	.09						
7. Status (# NAS members at center’s university)	34	39	.09	.07	−.03	.05	.06	.33					
8. Female (= 1, male = 0)	.13	.34	−.08	−.14	.06	.03	−.22	.00	−.02				
9. Trend in field IDR	.0000	.0008	.02	.03	.13	−.02	.04	.06	−.04	−.04			
10. Field size (# of Ph.D.s)	26,091	15,658	.06	.06	−.07	.19	.03	.01	−.04	−.01	−.31		
11. Average citations by field	10.77	7.12	.11	.11	.17	.32	.04	.02	.05	.06	−.25	.41	
12. Journal turnaround time by field	4.91	1.93	.11	.15	−.01	.09	.16	.06	.02	.00	.39	.25	.33

* As of 2010.

† As of 2005.

Table 2. Correlations and Descriptive Statistics for Person-years (N = 9,780)*

Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9
1. Visibility (total # citations)	81	180									
2. Productivity (total # articles)	3	3	.58								
3. Interdisciplinarity (mean IDR score)	.62	.15	.06	.02							
4. Number of authors (avg. per year)	7.50	40.01	.12	.16	.04						
5. Mean journal impact factor (JIF)	1.69	1.80	.31	.14	.14	.19					
6. Repeat collaborations (avg. per year)	.15	.28	.01	.07	−.04	.10	−.06				
7. Potential reach (N focal paper SCs)	1.60	.75	−.07	−.07	.23	−.05	−.11	−.02			
8. Visibility (cumulative, lagged)	482	959	.29	.43	.07	.07	.19	−.03	−.04		
9. Productivity (cumulative, lagged)	23.95	35.88	.30	.57	.03	.05	.13	−.02	−.04	.62	
10. Number of references	11.77	13.98	.15	.09	.17	.08	.29	−.02	−.02	.21	.11

* Correlations greater than .02 are significant at the .05 level.

outcome variable is logged, the coefficient of −.96 suggests that an increase of .10 in the IDR score reduces productivity by 9.6 percent over one’s career, controlling for professional age and other factors.³ This effect suggests that interdisciplinary scientists do indeed experience lower productivity. This productivity

³ When the dependent variable has been log-transformed and the predictors have not, a one-unit increase in the independent variable produces a 100 × (coefficient) percent change in the dependent variable.

penalty holds, and even gets stronger ($b = -.99$, $S.E. = .39$), when we weight the article count by number of coauthors, such that a paper with two authors contributes only .5 to a scholar's productivity. We also examined IDR's effect on productivity at the person-year level, at which IDR scores are averaged and publications are summed. Using a negative binomial model with indicator variables for each individual (table 4, model 1) allows us to rule out differences in individual propensities to engage in IDR. Here, too, IDR has a significant and negative effect ($b = -.148$) on productivity. In those years when scholars do more interdisciplinary work, they publish fewer articles. Thus we find consistent support for a production penalty across both specifications, consistent with H1.

Once published, however, IDR shines: in support of H2, we find that IDR increases scholarly visibility. The coefficient for IDR (.62 in table 3, model 1) suggests that an increase of .10 in the IDR score increases a scholar's citations, on average, by 6.2 percent. In table 4, model 2—a paper-level model with person and year fixed effects—we see that a paper's IDR score has a positive and significant effect on visibility. This effect also holds when we exclude group authors, whose papers tend to be highly cited, and when we consider a narrow window of only 2001–2002 publications, a period beyond the five years useful for predicting long-term citations (Wang, 2013; results not shown). Because articles in multidisciplinary, high-impact journals like *Science*, *Nature*, and *PNAS* could be driving the reception benefit, we omitted these 171 papers (.05 percent of total) from all measures. All hypothesized results are also robust to this change.⁴

Although IDR is more visible, on average, we also find support for H3: having a record of interdisciplinary scholarship increases the overall variance of a scientist's papers (H3). In table 3, model 2, we see that scientists who publish more IDR are more likely to produce both frequently cited and rarely cited works—they experience more hits and more flops than their mono-disciplinary counterparts. We modeled the standard deviation in citations rather than mean citations and also controlled for the standard deviation rather than the mean of journal impact factor. To investigate this finding further, we examined whether it is indeed the high-IDR papers that display more variability in citations, rather than highly interdisciplinary scientists having a few disciplinary papers that are not well cited. Using the median IDR score, we distinguished low-IDR and high-IDR papers. Then, for each scientist who had at least two papers of each type ($N = 647$), we calculated the standard deviation of the citations received by their low-IDR papers and did the same for their high-IDR papers. As theorized, the variance of scholars' high-IDR papers (mean = 30.4) is higher than the variance of their low-IDR papers (mean = 24.2), and this difference is statistically significant ($p < .0015$).

The distance (or cognitive dissimilarity) between fields contributes to the reception benefit and determines the productivity penalty, providing support for H4. We assessed this by removing distance from the IDR measure: we simply calculated the total number of unique SCs referenced by a scientist across all his or her papers with no regard for their similarity (range 1–107, mean = 32).

⁴ These journals are all classified into the "multidisciplinary" subject category (SC). Another way to identify interdisciplinary journals is to examine those with multiple SCs assigned (the maximum is 6). At the journal level, the number of SCs is negatively related to journal impact factor (–.15), reassuring us that a correlation between interdisciplinary journals and impact is not driving the reception benefit we document.

Table 3. Path Analytic Models (Unstandardized Coefficients and Standard Errors)

Variable	Model 1	Model 2 S.D. of citations*	Model 3 N of SCs [†]	Model 4 Trend in field IDR	Model 5 Electrical engineers	Model 6 Life scientists
<i>Effects on visibility (logged)</i>						
Interdisciplinary research (IDR)	.62** (.25)	18.19*** (6.49)	.01*** (.00)	.66*** (.25)	-.00 (.38)	1.68* (.91)
Productivity	1.27*** (.03)	6.65*** (.66)	1.13*** (.03)	1.27*** (.03)	1.27*** (.06)	1.16*** (.06)
Mean journal impact factor (JIF)	.33*** (.02)	7.50*** (.45)	.32*** (.02)	.33*** (.02)	.35*** (.06)	.23*** (.03)
Professional age (years since Ph.D.)	-.01*** (.00)	-.04 (.07)	-.01*** (.00)	-.01*** (.00)	-.02*** (.00)	-.01 (.01)
Ranking of Ph.D.-granting institution	.02 (.01)	.50 (.40)	.01 (.01)	.02 (.01)	.02 (.03)	.07 (.04)
Status (# NAS members at center's university)	.00** (.00)	.06*** (.02)	.00** (.00)	.00** (.00)	.00*** (.00)	-.00 (.00)
Female (yes = 1)	.07 (.07)	.37 (1.91)	.05 (.07)	.07 (.07)	.09 (.16)	.11 (.14)
Trend in field IDR	-31.13 (30.60)	-1691.68** (808.43)	-51.20* (30.41)	-34.76 (30.70)	-	-
IDR × Trend in field IDR	-	-	-	360.33 (295.76)	-	-
Field size (number of Ph.D.s)	-.00*** (.00)	-.00 (.00)	-.00*** (.00)	-.00*** (.00)	-	-
Field avg. citations	.00 (.00)	-.09 (.11)	-.00 (.00)	.00 (.00)	-	-
Intercept	1.05*** (.18)	-18.17*** (4.88)	1.54*** (.10)	1.43*** (.10)	1.18*** (.28)	.55 (.67)
R ²	.83	.4	.83	.83	.82	.86
<i>Effects on productivity (logged)</i>						
Interdisciplinary research (IDR)	-.96*** (.33)	-.78** (.32)	.04*** (.00)	-.84** (.33)	-.55 (.52)	-1.61 (1.55)
Mean journal impact factor (JIF)	.22*** (.03)	.18*** (.02)	.01 (.02)	.22*** (.03)	.20** (.08)	.05 (.06)
Professional age (years since Ph.D.)	.05*** (.00)	.05*** (.00)	.04*** (.00)	.05*** (.00)	.04*** (.01)	.04*** (.01)
Ranking of Ph.D.-granting institution	.07*** (.02)	.06*** (.02)	.02 (.02)	.07*** (.02)	.01 (.04)	.19*** (.07)
Status (# NAS members at center's university)	-.00 (.00)	-.00 (.00)	-.00 (.00)	-.00 (.00)	.00** (.00)	-.00 (.00)
Female (yes = 1)	-.34*** (.10)	-.35*** (.10)	-.29*** (.07)	-.35*** (.10)	-.08 (.21)	-.34 (.23)
Trend in field IDR	6.26 (47.18)	-.95 (47.06)	-46.94 (34.76)	-4.42 (47.23)	-	-
IDR × Trend in field IDR	-	-	-	949.86** (395.42)	-	-
Field size (number of Ph.D.s)	.00 (.00)	.00 (.00)	.00 (.00)	.00 (.00)	-	-
Field avg. journal turnaround time	.04* (.02)	.04** (.02)	-.01 (.02)	.04* (.02)	-	-
Intercept	1.96*** (.25)	1.97*** (.25)	1.05*** (.10)	1.34*** (.13)	2.16*** (.34)	2.85** (1.12)
R ²	.32	.32	.62	.32	.31	.30
<i>Effects on interdisciplinary research</i>						
Mean journal impact factor (JIF)	.02*** (.00)	.02*** (.00)	5.36*** (.51)	.02*** (.00)	.05*** (.01)	-.01 (.00)

(continued)

Table 3. (continued)

Variable	Model 1	Model 2 S.D. of citations*	Model 3 N of SCs [†]	Model 4 Trend in field IDR	Model 5 Electrical engineers	Model 6 Life scientists
Professional age (years since Ph.D.)	-.00*** (.00)	-.00*** (.00)	.41*** (.06)	-.00*** (.00)	-.00 (.00)	.00 (.00)
Ranking of Ph.D.-granting institution	.00 (.00)	.00 (.00)	1.31*** (.38)	.00 (.00)	.00 (.01)	.01 (.00)
Status (# NAS members at center's university)	-.00*** (.00)	-.00*** (.00)	-.00 (.05)	-.00*** (.00)	-.00* (.00)	-.00*** (.00)
Status ²	.00*** (.00)	.00*** (.00)	-.00 (.00)	.00*** (.00)	.00* (.00)	.00*** (.00)
Female (yes = 1)	.01 (.01)	.01 (.01)	-1.35 (1.84)	.01 (.01)	-.02 (.03)	-.01 (.01)
Trend in field IDR	13.72*** (3.94)	13.71*** (3.94)	2478.96*** (720.74)	13.71*** (3.94)	—	—
Intercept	.64*** (.01)	.64*** (.01)	12.08*** (2.17)	.02 (.01)	.54*** (.03)	.73*** (.03)
R ²	.11	.11	.17	.10	.14	.13
N individuals	854	854	854	854	177	102
AIC	52,347	60,161	60,705	38,687	7,458	4,017
BIC	52,808	60,679	61,166	39,210	7,623	4,154

* $p < .10$; ** $p < .05$; *** $p < .01$; two-tailed tests.

* For this model we used the standard deviation of journal impact factor rather than its mean in the visibility equation.

† For this model we used the number of subject categories instead of IDR.

Substituting this measure for the IDR measure (see table 3, model 3), the positive effect on citations holds, suggesting that even spanning related fields improves citations (if only slightly), presumably by broadening one's prospective audience. The negative effect on productivity does not hold, however. In fact, this alternative measure of IDR positively affects productivity, perhaps because drawing on multiple disciplines expands the number of possible journal outlets. Simply drawing on more SCs doesn't hinder productivity, unless those SCs are cognitively dissimilar. This suggests that it is more difficult to produce and successfully publish scholarship that spans unrelated fields, such as chemical engineering and anthropology, than related fields, such as chemical engineering and civil engineering. In supplementary analyses (not reported), we confirmed these results at the paper and person levels of analysis by controlling for the number of SCs referenced in each paper, and we confirmed that our IDR measure, which incorporates distance, is the only significant predictor. This offers support for the mechanism behind the production penalty we theorize: interdisciplinary research is cognitively difficult and slow to produce when it blends disparate fields.

Lastly, we examined how the effects of IDR depend on the nature of the field. As expected (H5a), we found that in highly interdisciplinary fields like the life sciences, IDR's positive impact on visibility is stronger than in less interdisciplinary fields like electrical engineering, in which IDR fails to reach statistical significance (table 3, models 5 and 6). These subanalyses fail to support H5b

Table 4. Fixed Effects Panel Regression Models of Productivity and Visibility on IDR*

Variable	Model 1 FE productivity [†]	Model 2 FE visibility [‡]
IDR	-.148** (.067)	.341*** (.066)
Productivity (cumulative, lagged)	-.001 (.000)	-.000 (.000)
Visibility (cumulative, logged and lagged)	.193*** (.009)	-.058*** (.011)
Number of authors	.002*** (.000)	.000 (.000)
Journal impact factor	-.001 (.005)	.153*** (.008)
Repeat collaborations	.342*** (.026)	-.041* (.022)
Number of references		.011*** (.002)
Potential reach (# focal paper SCs)		.038*** (.015)
Intercept	2.940*** (.253)	2.008*** (.251)
Year dummy variables	No	Yes
Person fixed effects	Yes	Yes
N obs.	8,525	29,775
N individuals	842	854
F		39.87
Wald chi-square	12,035.87	

• $p < .10$; ** $p < .05$; *** $p < .01$; two-tailed tests.

* All models exclude observations with IDR = 0.

† Negative binomial regression at the person-year level with indicator variables for all individuals (Allison and Waterman, 2002).

‡ GLS regression at the paper level with person and year fixed effects, robust standard errors, and multiple imputation.

(field differences in the productivity penalty), perhaps due to small subsample sizes and low statistical power.⁵ To examine the dynamics of field-level interdisciplinarity, we interacted the individual IDR measure with the trend in field IDR, which captures fields that are increasing in IDR relative to those that are more stable over time. Results in table 3, model 4 show that this interaction term has a significant and positive effect on productivity, supporting H5b. Combined with the negative main effects of IDR and trend in field-level IDR, this positive interaction term suggests that fields trending toward IDR invoke fewer penalties for producing this type of work. Perhaps such fields provide better training in how to manage the cognitive challenges and/or are more amenable to IDR in the peer review process. The interaction term fails to

⁵ In a comparable multiple group analysis (MGA, a sub-type of SEM), we found the same results. MGA allows us to compare a model that permits the coefficients to vary between the groups (life science and electrical engineering) and a model that constrains the coefficients to be equivalent between the groups. Results support the findings reported in the text: it is only IDR's positive effect on visibility that is stronger in interdisciplinary fields like the life sciences.

reach significance for visibility (H5a). Overall, results from the field-specific subanalyses (comparing life sciences and electrical engineering) and the interaction of IDR and trend in field-level IDR provide partial support for H5a and H5b. The static differences in field IDR (high vs. low) shape the reception of IDR work, but the trends in field IDR moderate the production of that work.

IDR's main effects on productivity and visibility hold even in the face of important controls. In the person-level path analytic models, we controlled for precursors of IDR that have been identified by others, including gender, professional age, and status. Perhaps most surprising is the effect of gender on IDR: contrary to widely held perceptions and some previous empirical research (Rhoten and Pфирman, 2007), we find that women are not more likely than men to engage in IDR. Given that our sample underrepresents women, this may simply mean our sample is not well-suited for studying gender. We also find support for Phillips and Zuckerman's (2001) middle status conformity finding. Using the number of National Academy of Science (NAS) members at one's institution as a measure of status, we find the expected inverted U-shaped relationship between status and IDR: as model 1 (table 3) shows, the main effect of status is negative, and the squared term is positive. This suggests that scientists at both low- and high-status universities engage in IDR; scientists at middle-status universities are in a precarious position in which "the prospect of classification as a full-fledged player and the threat of delegitimation both loom large" (Phillips and Zuckerman, 2001: 384) and thus opt to conform to a disciplinary tradition. At the journal and field level, we controlled for variables that likely affect productivity and/or visibility—including field size, average citations per paper in the field, and average turnaround time at journals in the field—as well as lagged productivity (when modeling productivity as well as visibility) and lagged visibility (when modeling visibility as well as productivity). Their inclusion does not alter or render insignificant the main findings.

Drivers of the Productivity Penalty

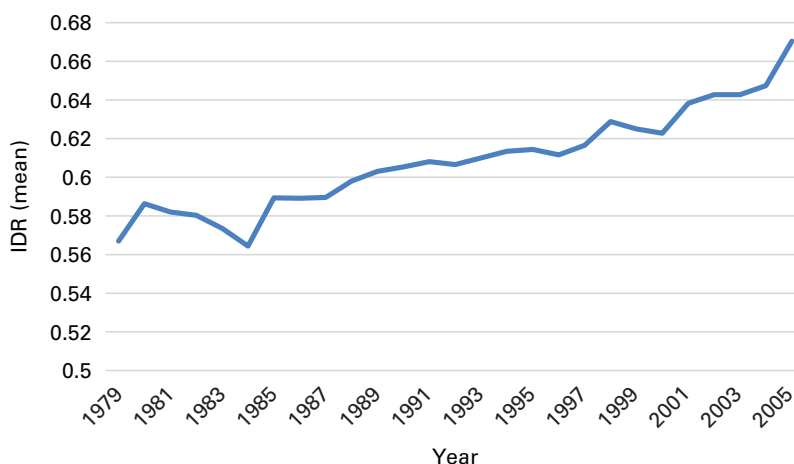
Cognitive and collaborative challenges. We theorize that IDR projects typically experience a steep learning curve as well as communication challenges among diverse collaborators. Supplemental data sources and analyses lead us to conclude that IDR projects indeed face these hurdles. First, we examined the interaction between repeat collaboration and IDR to see whether collaborators learn to work together better over time. Indeed, working repeatedly with a similar set of collaborators reduces the productivity penalty ($b = 1.51$, $S.E. = .077$). Second, we capitalized on survey data that we collected from a subset of scholars in the archival data. In 2005, we sent the survey to all IUCRC-affiliated scholars ($N = 375$) for whom we could find e-mail addresses and received responses from 147. Scholars were asked about the nature of the collaboration on their most recent coauthored paper. For 68 of these, we were able to match these responses to a specific paper in our sample, for which we had calculated IDR. The IDR scores of papers by these scholars are not significantly different from the rest of our sample. Despite the small subsample size, a series of t-tests revealed marginally significant challenges associated with the

production of IDR. When we compared those who responded “excellent” or “good” to key survey questions with those who responded “neutral,” “poor,” or “very poor,” we found that interdisciplinary collaborations experience communication difficulties: communication is reportedly less clear ($p = .065$), more difficult ($p = .083$), and of lower quality ($p = .106$). Moreover, interdisciplinary teams have more difficulty generating ideas ($p = .109$).

And last, we created an alternative measure that captures multidisciplinary rather than interdisciplinarity. Our theory suggests that it is challenging to incorporate distant ideas in a single paper. An alternative is that a scientist writes papers across different disciplines but each paper is mono-disciplinary; this scientist is multi- but not interdisciplinary. Although multidisciplinary may require additional expertise and a diverse network of collaborators, it does not require integration of diverse fields or coordination of diverse collaborators. To capture multidisciplinary, we pooled all of the subject categories across a person’s publication record and then calculated IDR. This differs from a scientist’s IDR score, which is first calculated at the paper level and then averaged across his or her set of papers. Consistent with our theorizing, scientists conducting multidisciplinary research are more productive ($b = .58$, $S.E. = .31$); only scientists publishing interdisciplinary research are less productive.

Challenging review process. Another possibility is that IDR work is penalized by reviewers and editors in the review process, but we find little support for this explanation. To examine this possibility, we collected data on the length of the review process—the turnaround time—for two journals represented in our data from WoS: one publishing articles with an above-average IDR score (.64) and another publishing articles with a below-average IDR score (.59). For the 711 articles published in these journals, written by 145 of our sampled authors, the median time under review was 85 days. But there is no significant correlation ($r = .04$) between time under review and the paper’s IDR score, and the IDR score does not predict time under review in a paper-level model that controls for journal impact, number of authors, repeat collaboration, and five-year time interval with person fixed effects (results not shown).

Confidence in this finding is buttressed by an analysis of data on unpublished working papers that we collected from www.arXiv.org. We searched for working papers written by the 854 authors in our sample and were able to locate 220 papers written by 63 authors (see Online Appendix B). Stripping all references from these working papers, identifying each referenced journal, and matching the journals with WoS SCs allowed us to calculate IDR scores for each working paper. Comparisons of the 220 working papers with the 3,983 published articles by the same authors revealed that working papers are indeed more interdisciplinary. In a truer matched sample that includes only the most recent working paper and the most recent published article for the 64 scholars who have both, however, we find no difference in IDR (see Online Appendix B, Panel A). So perhaps working papers are more interdisciplinary simply because they are more recent: their average posting date on www.arXiv.org is 2009, compared with an average publication date of 1997 for the published articles, reflecting the upward trend documented by Porter and Rafols (2009) and in our data, graphed in figure 1. This is confirmed by a regression predicting IDR: once publication year and number of authors are added as controls, working papers are not more interdisciplinary than

Figure 1. Trend in interdisciplinary research over time.

published articles (results not reported). As a more stringent test, we took a closer look at our sample of working papers (also reported in Online Appendix B, Panel B) and identified which were subsequently published, using authors' updates on www.arXiv.org and online searches for each paper. We compared working papers that eventually got published with those that did not. The eventually published papers ($N = 122$) are actually more interdisciplinary than the still unpublished papers ($N = 115$), a difference significant at $p = .054$ (two-tailed). When we limit this comparison to authors with one unpublished working paper and one eventually published working paper in www.arXiv.org (24 scholars and 48 papers), we find no difference in IDR. Taken together, these results suggest that IDR papers are not hindered in the review process.

These supplementary data and analyses support our theorizing about cognitive and collaborative challenges associated with interdisciplinary research. The second stage of production—the review stage—may not be such a roadblock; we do not find IDR work more likely to be tossed in a file drawer or rejected. Rather, the first stage of production, during which authors plan, coordinate, and conduct their research, is the largest hurdle for interdisciplinary work and the main source of the production penalty. Although we acknowledge that these are not causal tests but associations, we subscribe to the view that documenting the role of a mechanism (or two) empirically strengthens claims of causal connections (Reskin, 2003; Gross, 2009).

Robustness Checks and Selection Issues

Sample selection concerns. Additional analyses alleviate concerns that the results are applicable only to center-affiliated scientists. It is true that our sample of academic scientists (described in table 1) includes only scientists affiliated with at least one research center, the IUCRCs we study. But center affiliation is relatively common: almost one-third (32 percent) of faculty at Research I universities nationwide are affiliated with a research center (Boardman and Corley, 2008). In addition, research centers are the main

mechanism by which universities and federal agencies seek to foster interdisciplinary research (Sá, 2008; Biancani, McFarland, and Dahlander, 2014), so they are ideal sites for investigating its impact.

Our center-based sample is distinctive in some ways but not in ways that alter the results we report. Compared with their counterparts unaffiliated with centers, university research center scientists tend to be more experienced (Bozeman, Dietz, and Gaughan, 2001) and productive (Biancani, McFarland, and Dahlander, 2014), and this may be particularly true when scholars are connected to industry (Carayol and Nguyen Thi, 2005) via IUCRCs. We find, however, that compared with the broader population of scientists in research universities represented in the NSF's surveys of Ph.D.-level scientists, our sample of scientists is not more productive—but is marginally younger and more male; see Online Appendix C.⁶ This suggests, with respect to productivity, that our results generalize to faculty in R1 institutions. The productivity penalty we document has also been documented among Stanford faculty: those who take a joint appointment (one measure of IDR) experience a decline in productivity (Biancani, McFarland, and Dahlander, 2014). Concerns about the younger age of our sample are alleviated by robustness checks with different age groups represented in our data. An interaction between IDR and age is negative and significant ($b = -.012$; $p < .01$), suggesting that if anything the penalty is stronger for older scientists—but the size of the coefficient is very small. Concerns about the preponderance of men in our sample are alleviated when interactions between female and IDR fail to reach statistical significance, suggesting that results hold for both men and women.

Other tests further suggest that our sample is representative of scientific papers more broadly. Although we cannot calculate our measure of IDR for Ph.D.-level scientists in the NSF's survey data, we can compare our sample of papers with the population of papers in Web of Science analyzed by Uzzi et al. (2013). Uzzi generously shared his measures for our sample of papers, so we can make direct comparisons in terms of both conventionality—a paper's tendency to reference journals that are commonly referenced together—and novelty—a paper's tendency to reference unusual journal pairings. The second panel of Online Appendix C shows that in the 1980s and 1990s, our papers are almost identical to the population of Web of Science papers in terms of conventionality. And, importantly, our papers exhibit slightly less novelty than the population, allaying concerns that center-based scholars are innovative superstars.⁷ This comparison demonstrates that our sample of papers is not exceptional, and our results are likely to apply to academic science more broadly. That said, and despite our additional efforts, we cannot completely rule out that there is something distinctive about center-based scientists.

To further address concerns about the selective nature of our center-based sample, we leveraged our longitudinal data, which include papers written by these scientists before they joined a center. Even after restricting the sample to papers published five years before and after a center's founding, and estimating fixed-effects models (Online Appendix D, models 1 and 2), the main

⁶ For the comparison reported here, we compared our sample of scientists with scientists represented in the NSF's SESTAT data in 2003, which collected self-reports of publications between 1998 and 2003.

⁷ Online Appendix D, model 3, also replicates the conventionality + novelty effect from Uzzi et al. (2013) for our sample and demonstrates that IDR has an independent and significant effect on visibility.

effects of IDR (positive on visibility; negative on productivity) remain significant. The interaction between IDR and center affiliation is not statistically significant, suggesting that center affiliation does not make IDR work more visible (model 1) or less difficult (model 2). This analysis suggests that center affiliation does not increase the benefits or reduce the cost of conducting IDR.

Quality of IDR papers. We also find evidence suggesting that interdisciplinary papers are not distinctive in terms of their inherent quality. If IDR papers are lower in quality than other papers, this could explain the productivity penalty we document. If IDR papers are higher-quality papers, this could explain the reception benefit we document. Additional analyses allowed us to rule out these possibilities. We controlled for potential quality differences at the individual, journal, and field level. Because we specified models with person fixed effects that controlled for unmeasured attributes, we can be certain that the effects we document are not attributable to individual differences in aptitude or in propensity to engage in IDR. In other words, low-quality scholars are not attracted to IDR (which could explain the productivity penalty), nor are high-quality scholars (which could explain the reception benefit). With respect to the quality of the journal, we do see a positive correlation between IDR and journal impact ($r = .14$ at the paper-year level and $.09$ at the paper level). But the positive effect of IDR on citations holds even when we control for journal impact (table 3, model 1; table 4, model 2), when we include journal fixed effects (results not shown), and when we restrict our analysis to the subset of low-impact journals—those whose impact factor falls in the lowest quartile (results not shown). In other words, high-IDR papers in low-impact journals are more highly cited than other papers in those journals. Moreover, recall that we found no difference in turnaround times between papers published in a high-IDR journal and in a low-IDR journal. Thus we doubt that IDR papers are inherently higher quality, as then they would likely move more quickly through the review process; we also doubt that they are lower quality, as then they would likely spend more time in development in the review process or be more likely to be rejected. These robustness checks are consistent with our argument that IDR, and not some unobserved heterogeneity or selectivity, increases visibility and reduces productivity.

DISCUSSION

To empirically investigate the potential costs and the widely touted benefits of interdisciplinary scholarship, we collected and collated data from various sources for a sample of almost 900 center-based scientists and their 32,000 publications, providing the first systematic and mid-scale assessment of interdisciplinary research's impact. Our results demonstrate that IDR does benefit scientists: it improves their visibility in the scientific community, as indicated by cumulative citation counts. But we also document a productivity penalty associated with IDR: it depresses the number of articles that scientists publish. We see these effects in analyses of papers (i.e., high-IDR papers are more highly cited), in yearly analysis (e.g., scientists publishing high-IDR papers publish less in that year), and at the person level (e.g., scientists who publish IDR have more citations and higher variation in their citations). And when compared, the productivity penalty for IDR (standardized coefficient for direct effect = $-.09$)

outweighs the reception benefit (standardized coefficient for direct effect = $+ .03$) of engaging in this research. In other words, engaging in interdisciplinary research depresses productivity more than it increases citations. Compared with a scholar in the 20th percentile of IDR, a scholar in the 80th percentile of IDR produces seven fewer articles (20 rather than 27) and garners 40 more citations (230 rather than 190).⁸ The productivity penalty is strong enough to make the total effect of IDR on citations—i.e., the direct effects reported above, plus the indirect effect of IDR on citations via productivity—slightly negative (standardized coefficient for total effect = $-.04$). Apparently the learning curve is steep: it takes more time, effort, diligence, and perhaps coordination to master (at least aspects of) different fields and to work with scientists trained in disparate disciplines. That said, we cannot assess whether this trade-off is harmful or beneficial for a scientist—do 40 more citations offset seven fewer publications? We can say that the greater visibility this work receives is accompanied by fewer publications published.

We also make several theoretical contributions to the literature. First, we reorient away from reception penalties, which dominate the categories literature, toward production penalties, and we explore a new form of production penalty. The few recent papers that have examined production penalties focus on reduced quality (Kovács and Johnson, 2014), documenting through blind taste tests (i.e., controlling for audience perception), for example, that category-spanning wines are rated more negatively—presumably because of skill deficiencies on the part of the winemakers (Negro and Leung, 2013). In contrast, we focus on another form of production penalty that besets category-spanning work, at least in the realm of science: reduced productivity. Supplementary data on unpublished working papers, journal turnaround times, and a survey of authors' experiences, in addition to supplementary models incorporating individual fixed effects, suggest that productivity is hampered by the communication and coordination hurdles faced in the research process and not by the review process. This needs to be reconciled with the implicit production benefits found in the innovation and diversity literatures: better decisions (Beckman and Haunschild, 2002), heightened creativity (Pelled, Eisenhardt, and Xin, 1999), and organizational centrality (Powell, Koput, and Smith-Doerr, 1996). Our work suggests a potential short-run (fewer papers) and long-run (higher visibility) tradeoff.

Second, we demonstrate that production penalties and reception benefits depend on characteristics of the field. The production penalty (fewer papers) is less severe in a field that is increasingly interdisciplinary, like chemical engineering, than in a more stable field, whether that be a high-IDR field like life sciences or a low-IDR field like electrical engineering. The reception benefits, however, appear to accrue to high-IDR fields like life sciences. This suggests that the benefits and penalties of doing IDR work are not the same across fields. More broadly, the effect of category spanning depends on its typicality in the field (Lo and Kennedy, 2015). It may be that these high-IDR fields are shaped by high-status actors who reap the benefit of category spanning and normalize it for those who follow (Rao, Monin, and Durand, 2005; Sgourev and Althuizen, 2014).

⁸ These predicted values are obtained from table 3, model 1. We used the *predict* post-estimation command in Stata, then calculated the average prediction for each decile of IDR, which we then exponentiated because the outcomes in our analysis were logged.

Third, we move beyond mere category spanning to take into consideration the relationship between the spanned categories. Spanning two dissimilar entities is qualitatively different from spanning two similar entities, but extant research on both category spanning (Hsu, Hannan, and Kocak, 2009; Negro, Hannan, and Rao, 2010) and recombinant innovation (Fleming, Mingo, and Chen, 2007) largely ignores this (but see Leahey and Moody, 2014, and Kovács and Johnson, 2014). The IDR measure we use is sensitive to such differences, allowing connections between two unrelated disciplines to contribute more to the IDR score than connections between two related disciplines. The distinction is crucial and allows us to distinguish between the effects of variety—branching out into a number of other fields—and distance—branching out into cognitively distant and unrelated fields. The distance measure we use is consistent with the logic of diversity and brokerage—which presumes networks with structural holes contain more distant, non-redundant knowledge—but we actually measure distance in knowledge space. When we measure IDR as sheer variety, neglecting distance between fields, we find that the reception benefit holds—papers that span dissimilar and even similar subfields receive more citations—but the productivity penalty does not, suggesting that the cognitive challenges and coordination costs associated with producing IDR are largely a function of the cognitive distance among fields. This is also the case when we calculate a measure of multidisciplinaryity, which accounts for distance between fields represented in a scholar's oeuvre but not between fields represented in a given paper. Taken together, these results suggest that efforts to integrate, in a single paper, the cognitive distance across disparate fields contribute to the reception benefit (greater citations) and drive the productivity penalty (fewer publications).

Investigating academic science also allows us to extend Pontikes' (2012) insight about the nature of the audience to reconcile an even broader array of literatures than she recognized. As Pontikes suggested (2012: 111), scientists are likely "market-makers," eager to identify and develop innovative, game-changing ideas, who are thus drawn to rather than repelled by multi-category offerings like IDR. And this, of course, is what we find: interdisciplinary scientists do not experience the reception-side penalties (e.g., devalued, overlooked) that the category-spanning literature has widely documented among "market-takers." Rather, as the innovation literature suggests, without explicitly referring to it as such, scientists derive a reception-side benefit: greater visibility. But we also confirm that IDR is a high-risk, high-reward activity, as indicated by both more overall citations and higher variation of citations; we confirm that innovative ideas have longer tails (Singh and Fleming, 2010). Thus Pontikes' focus on the nature of the audience places an important scope condition on not just the category-spanning literature but also the innovation and brokerage literature, which has implicitly focused on market-makers and reception-side benefits. Given that the nature of the audience matters, it would behoove scholars of both category spanning and innovation to identify more explicitly whether the actors under study are market-takers or market-makers and to consider how this shapes both reception benefits, such as visibility, and risks, such as variance.

Because this is the first mid-scale and empirical assessment of IDR's effects on scientists' careers, extensions will be fruitful. Most illuminating might be a closer examination of the audience, following Lynn's (2014) lead. The audience does not just cite and thereby contribute to a cumulative citation count. Audience members themselves come from disciplinary or interdisciplinary homes. In this

paper we focused on the distribution of, and distance between, cited disciplines: those represented in scholars' bibliographies. But a comparison of referenced disciplines with the citing disciplines—that is, the disciplinary homes of scientists who reference a given paper—would allow us to assess whether certain fields or combinations of fields appeal to different audiences. In concurrent analyses, we find that papers with high IDR scores are likely to be cited by papers that themselves are interdisciplinary and hail from multiple disciplines. But partly because more interdisciplinary papers span more fields than less interdisciplinary papers, their audience is unlikely to represent new fields that were not included among referenced works. This preliminary analysis suggests that breadth breeds further breadth, but more exploration is warranted.

We also encourage extensions to broader samples and alternative research designs. Although we studied almost 900 scientists from a wide range of fields and universities, they are all affiliated with university research centers that foster connections with industry. Our analyses of potential selection biases and endogeneity concerns suggest that center affiliation does not drive the results we report here. But a prospective research design that follows a large sample or population of academic scientists through their careers as they move in and out of center affiliations and other interdisciplinary ventures would complement our search for mechanisms and evidence consistent with our theorizing. Efforts underway to link the NSF's Survey of Doctorate Recipients (SDR) to both Web of Science records and the National Bureau of Economic Research Patent database would be ideal in this regard. Another possibility is to supplement Zucker and Darby's NanoBank data by adding data on references (to calculate IDR).

Practically, should university research administrators and federal agencies like the National Academies of Science and the National Science Foundation continue to invest in IDR? Given that IDR scholars produce useful and noteworthy research that has more impact on the scientific community, the enthusiasm for IDR is not premature. Our analysis suggests, however, that it is not attuned to the implications—especially the negative implications—that IDR has for individuals. Scholars who produce interdisciplinary work may be more likely to publish in top-tier journals—the person-level correlation between IDR and journal impact factor is .25—but their overall productivity is hampered. There is a clear production penalty associated with interdisciplinary work. Even though our analysis of working papers suggests that IDR does not hamper subsequent publication, and if anything seems to help, additional analysis on a larger sample is warranted. The penalties of IDR may be more far-reaching than we document here.

To this point, we encourage more direct examination of IDR's impact on scientific careers. We do not examine important career outcomes, like receiving tenure. Including more and younger scholars would be an important and useful extension, especially as universities reorganize to train scientists to be interdisciplinary. It is important to understand whether IDR enhances or damages career prospects for those starting out their careers; recent research does suggest younger scholars engage in IDR less (Sobey et al., 2013). Certainly it appears that interdisciplinary scholars bear additional burdens: they struggle to master multiple domains of knowledge, to integrate them in a single work, and to coordinate with coauthors from different backgrounds. But these projects also appear to be received well by fellow scientists. More-focused examinations of the career outcomes of individual scientists are still needed. Future research should examine forms of productivity other than publishing, such as

patenting, consulting, and advising, to examine whether scientists are compensating for their fewer publications with other activities.

Given the continued interest and enthusiasm for interdisciplinarity (National Research Council, 2014), it is important that the costs and benefits of this type of research be evaluated empirically. We have taken a step in that direction. Even if the individual-level costs of such work are substantial, the societal-level benefits—in the form of more useful and valuable science—seem clear. This suggests that scientists and the scientific community need to reassess how to evaluate scholarship if scientists are to be encouraged to continue to engage in interdisciplinary research.

Acknowledgments

This research was supported by NSF award #0332051 in 2003. We thank the many research assistants who assisted in data collection. We thank Brian Uzzi for generously sharing his data. We also thank seminar participants at University of Chicago, Penn State, McGill, Carnegie Mellon, University of Maryland, University of Iowa, UC Irvine, UC Riverside, INSEAD, IESE, University of Colorado at Boulder, NYU Stern, and University of Arizona for their helpful comments. Earlier versions of this paper were presented at the American Sociological Association in 2012 and the Academy of Management meetings in 2014.

REFERENCES

Abbott, A.

2001 *Chaos of Disciplines*. Chicago: University of Chicago Press.

Aharonson, B. S., and M. A. Schilling

2016 "Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution." *Research Policy*, 45: 81–96.

Allison, P. D., and J. S. Long

1987 "Interuniversity mobility of academic scientists." *American Sociological Review*, 52: 643–652.

Allison, P. D., and R. Waterman

2002 "Fixed effects negative binomial regression models." In R. M. Stolzenberg (ed.), *Sociology Methodology*: 247–265. Oxford: Basil Blackwell.

Anderson, T. W.

1957 "Maximum likelihood estimates for multivariate normal distribution when some observations are missing." *Journal of the American Statistical Association*, 52: 200–203.

Azoulay, P., T. Stuart, and Y. Wang

2014 "Matthew: Effect or fable?" *Management Science*, 60: 92–109.

Beckman, C. M.

2006 "The influence of founding team company affiliations on firm behavior." *Academy of Management Journal*, 49: 741–758.

Beckman, C. M., and P. R. Haunschild

2002 "Network learning: The effects of partners' heterogeneity of experience on corporate acquisitions." *Administrative Science Quarterly*, 47: 92–124.

Biancani, S., D. A. McFarland, and L. Dahlander

2014 "The semiformal organization." *Organization Science*, 25: 1306–1324.

Birnbaum, P. H.

1981 "Integration and specialization in academic research." *Academy of Management Journal*, 24: 487–503.

Bjork, B., and D. Solomon

2013 "The publishing delay in scholarly peer-reviewed journals." *Journal of Informetrics*, 7: 914–923.

Boardman, C., and B. Bozeman

2007 "Role strain in university research centers." *Journal of Higher Education*, 78: 430–463.

Boardman, C. P., and E. A. Corley

2008 "University research centers and the composition of research collaborations." *Research Policy*, 37: 900–913.

Boix Mansilla, V.

2006 "Assessing expert interdisciplinary work at the frontier: An empirical exploration." *Research Evaluation*, 15: 17–29.

Boix Mansilla, V., M. Lamont, and K. Sato

2016 "Shared cognitive–emotional–interactional platforms: Markers and conditions for successful interdisciplinary collaborations." *Science, Technology, and Human Values*, 41: 571–612.

Bollen, K. A.

1989 *Structural Equations with Latent Variables*. New York: Wiley.

Bozeman, B., J. Dietz, and M. Gaughan

2001 "Scientific and technical human capital: An alternative model for research evaluation." *International Journal of Technology Management*, 22: 636–655.

Brint, S.

2005 "Creating the future: 'New directions' in American research universities." *Minerva*, 43: 23–50.

Carayol, N., and T. U. Nguyen Thi

2005 "Why do academic scientists engage in interdisciplinary research?" *Research Evaluation*, 14: 70–79.

Carnabuci, G., and J. Bruggeman

2009 "Knowledge specialization, knowledge brokerage, and the uneven growth of knowledge domains." *Social Forces*, 88: 607–642.

Carroll, G. R.

1985 "Concentration and specialization: Dynamics of niche width in populations of organizations." *American Journal of Sociology*, 90: 1262–1283.

Clemens, E. S., W. W. Powell, K. McIlwaine, and D. Okamoto

1995 "Careers in print: Books, journals, and scholarly reputation." *American Journal of Sociology*, 101: 433–494.

Cole, J. R., and H. Zuckerman

1984 "The productivity puzzle: Persistence and change in patterns of publication of men and women scientists." *Advances in Motivation and Achievement*, 2: 217–258.

Corley, E., and M. Gaughan

2005 "Scientists' participation in university research centers: What are the gender differences?" *Journal of Technology Transfer*, 30: 371–381.

Cummings, J. N., and S. Kiesler

2005 "Collaborative research across disciplinary and organizational boundaries." *Social Studies of Science*, 35: 703–722.

Cummings, J. N., and S. Kiesler

2007 "Coordination costs and project outcomes in multi-university collaborations." *Research Policy*, 36: 1620–1634.

Dahlander, L., and L. Frederiksen

2012 "The core and cosmopolitans: A relational view of innovation in user communities." *Organization Science*, 23: 988–1007.

Diamond, A. M.

1986 "What is a citation worth?" *Journal of Human Resources*, 21: 200–215.

Evans, J. A.

2008 "Electronic publication and the narrowing of science and scholarship." *Science*, 321: 395–399.

Ferber, M.

1986 "Citations: Are they an objective measure of scholarly merit?" *Signs*, 11: 381–389.

Fleming, L.

2001 "Recombinant uncertainty in technological search." *Management Science*, 47: 117–132.

Fleming, L., S. Mingo, and D. Chen

2007 "Collaborative brokerage, generative creativity, and creative success." *Administrative Science Quarterly*, 52: 443–475.

Fleming, L., and O. Sorenson

2004 "Science as a map in technological search." *Strategic Management Journal*, 25: 909–928.

Fleming, L., and D. M. Waguespack

2007 "Brokerage, boundary spanning, and leadership in open innovation communities." *Organization Science*, 18: 165–180.

Foster, J. G., A. Rzhetsky, and J. A. Evans

2015 "Tradition and innovation in scientists' research strategies." *American Sociological Review*, 80: 875–908.

Freeman, J., and M. T. Hannan

1983 "Niche width and the dynamics of organizational populations." *American Journal of Sociology*, 88: 1116–1145.

Gans, J. S., and F. Murray

2015 "Credit history: The changing nature of scientific credit." In A. Jaffe and B. Jones (eds.), *The Changing Frontier: Rethinking Science and Innovation Policy*: 107–134. Chicago: University of Chicago Press.

Gaughan, M., and B. Ponomariov

2008 "Faculty publication productivity, collaboration, and grants velocity: Using curricula vitae to compare center-affiliated and unaffiliated scientists." *Research Evaluation*, 17: 103–110.

Gross, N.

2009 "A pragmatist theory of social mechanisms." *American Sociological Review*, 74: 358–379.

Hamilton, K. S., F. Narin, and D. Olivastro

2005 *Using Bibliometrics to Measure Multidisciplinarity*. Westmont, NJ: iplQ, Inc.

Hannan, M. T., and J. H. Freeman

1989 *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Hannan, M. T., L. Pólos, and G. R. Carroll

2007 *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton, NJ: Princeton University Press.

Hargadon, A. B.

2002 "Brokering knowledge: Linking learning and innovation." In R. M. Kramer and B. M. Staw (eds.), *Research in Organizational Behavior*, 24: 41–85. New York: Elsevier/JAI.

Hsu, G.

2006a "Evaluative schemas and the attention of critics in the US film industry." *Industrial and Corporate Change*, 15: 467–496.

Hsu, G.

2006b "Jack of all trades and master of none: Audiences' reactions to spanning genres in feature film production." *Administrative Science Quarterly*, 51: 420–450.

Hsu, G., M. T. Hannan, and O. Kocak

2009 "Multiple category memberships in markets: An integrative theory and two empirical tests." *American Sociological Review*, 74: 150–169.

Hsu, G., G. Negro, and F. Perretti

2012 "Hybrids in Hollywood: A study of genre spanning in the U.S. film industry." *Industrial and Corporate Change*, 21: 1427–1450.

Jacobs, J. A., and S. Frickel

2009 "Interdisciplinarity: A critical assessment." *Annual Review of Sociology*, 35: 43–65.

Kaplan, S., J. Milde, and R. S. Cowan

2016 "Symbiotic practices in boundary spanning: Bridging the cognitive and political divides in interdisciplinary research." *Academy of Management Journal*, published online ahead of print. DOI: 10.5465/amj.2015.0809.

Knorr Cetina, K.

1980 "The scientist as an analogical reasoner." In K. Knorr-Cetina, R. G. Krohn, and R. Whitley (eds.), *The Social Process of Scientific Investigation*: 25–52. Boston: Kluwer.

Kovács, B., and M. T. Hannan

2015 "Conceptual spaces and the consequences of category spanning." *Sociological Science*, 2: 252–286.

Kovács, B., and R. Johnson

2014 "Contrasting alternative explanations for the negative effects of category spanning: A study of restaurant reviews and menus in San Francisco." *Strategic Organization*, 12: 7–37.

Lamont, M.

2009 *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press.

Lamont, M., G. Mallard, and J. Guetzkow

2006 "Beyond blind faith: Overcoming obstacles to interdisciplinary evaluation." *Research Evaluation*, 15: 43–55.

Larivière, V., and Y. Gingras

2010 "On the relationship between interdisciplinarity and scientific impact." *Journal of the American Society for Information Science and Technology*, 61: 126–131.

Larivière, V., S. Haustein, and K. Börner

2015 "Long-distance interdisciplinarity leads to higher scientific impact." *PLoS ONE* 10:e0122565.

Latour, B.

1987 *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

Leahey, E.

2006 "Gender differences in productivity: Research specialization as a missing link." *Gender and Society*, 20: 754–780.

Leahey, E.

2007 "Not by productivity alone: How visibility and specialization contribute to academic earnings." *American Sociological Review*, 72: 533–561.

Leahey, E.

2016 "From solo investigator to team scientist: Trends in the practice and study of research collaboration." *Annual Review of Sociology*, 42: 81–100.

Leahey, E., and J. Moody

2014 "Sociological innovation through subfield integration." *Social Currents*, 1: 228–256.

Lingo, E. L., and S. O'Mahony

2010 "Nexus work: Brokerage on creative projects." *Administrative Science Quarterly*, 55: 47–81.

Lo, J. Y.-C., and M. T. Kennedy

2015 "Approval in nanotechnology patents: Micro and macro factors that affect reactions to category blending." *Organization Science*, 26: 119–139.

Lynn, F. B.

2014 "Diffusing through disciplines: Insiders, outsiders, and socially influenced citation behavior." *Social Forces*, 93: 355–382.

Maliniaka, D., R. Powers, and B. F. Walter

2013 "The gender citation gap in international relations." *International Organization*, 67: 889–922.

McBee, D. J., and E. Leahey

2016 "New directions in interdisciplinary training: Trials and tribulations." In S. Frickel, M. Albert, and B. Prainsack (eds.), *Investigating Interdisciplinary Collaboration: Theory and Practice across Disciplines*: ch. 1. Rutgers, NJ: Rutgers University Press (in press).

McBrier, D. B.

2003 "Gender and career dynamics within a segmented professional labor market: The case of law academia." *Social Forces*, 81: 1201–1266.

McPherson, M., L. Smith-Lovin, and J. M. Cook

2001 "Birds of a feather: Homophily in social networks." *Annual Review of Sociology*, 27: 415–444.

Merton, R. K.

1977 "The sociology of science: An episodic memoir." In R. K. Merton and J. Gaston (eds.), *The Sociology of Science in Europe*: 3–144. Carbondale, IL: Southern Illinois University Press.

Millar, M. M., and D. A. Dillman

2012 "Trends in interdisciplinary dissertation research: An analysis of the survey of earned doctorates." NCSSES Working paper #12-200.

Montpetit, É., A. Blais, and M. Foucault

2008 "What does it take for a Canadian political scientist to be cited?" *Social Science Quarterly*, 89: 802–816.

Murray, F.

2010 "The Oncomouse that roared: Hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions." *American Journal of Sociology*, 116: 341–388.

Najman, J. M., and B. Hewitt

2003 "The validity of publication and citation counts for sociology and other selected disciplines." *Journal of Sociology*, 39: 62–80.

National Academies of Science, National Academy of Engineering, and Institute of Medicine

2005 *Facilitating Interdisciplinary Research*. Washington, DC: National Academies Press.

National Research Council

2014 *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*. Washington, DC: National Academies Press.

Negro, G., M. T. Hannan, and H. Rao

2010 "Categorical contrast and audience appeal: Niche width and critical success in winemaking." *Industrial and Corporate Change*, 19: 1397–1425.

Negro, G., and M. D. Leung

2013 "Actual and perceptual effects of category spanning." *Organization Science*, 24: 684–696.

Neter, J., W. Wasserman, and M. H. Kutner

1985 *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*. Homewood, IL: Irwin.

Owen-Smith, J., and W. W. Powell

2003 "The expanding role of university patenting in the life sciences: Assessing the importance of experience and connectivity." *Research Policy*, 32: 1695–1711.

Pelled, L. H., K. M. Eisenhardt, and K. R. Xin

1999 "Exploring the black box: An analysis of work group diversity, conflict and performance." *Administrative Science Quarterly*, 44: 1–28.

Phillips, D. J.

2011 "Jazz and the disconnected: City structural disconnectedness and the emergence of the jazz canon (1897 to 1933)." *American Journal of Sociology*, 117: 420–483.

Phillips, D. J., and E. Zuckerman

2001 "Middle status conformity: Theoretical restatement and evidence from two markets." *American Journal of Sociology*, 107: 379–429.

Poincare, H.

1952 "Mathematical creation." In B. Ghiselin (ed.), *The Creative Process*. Berkeley, CA: University of California Press.

Ponomarev, B. L., and P. C. Boardman

2010 "Influencing scientists' collaboration and productivity patterns through new institutions: University research centers and scientific and technical human capital." *Research Policy*, 39: 613–624.

Pontikes, E. G.

2012 "Two sides of the same coin: How ambiguous classification affects multiple audiences' evaluations." *Administrative Science Quarterly*, 57: 81–118.

Porter, A. L., A. S. Cohen, J. D. Roessner, and M. Perreault

2007 "Measuring researcher interdisciplinarity." *Scientometrics*, 72: 117–147.

Porter, A. L., and I. Rafols

2009 "Is science becoming more interdisciplinary? Measuring and mapping six research fields over time." *Scientometrics*, 81: 719–745.

Powell, W. W., K. W. Koput, and L. Smith-Doerr

1996 "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology." *Administrative Science Quarterly*, 41: 116–145.

Pray, L.

2002 "Interdisciplinarity in science and engineering: Academia in transition." *Science Career Magazine*. Accessed at <http://www.sciencemag.org/careers/2002/01/interdisciplinarity-science-and-engineering-academia-transition>.

Prpic, K.

2002 "Gender and productivity differentials in science." *Scientometrics*, 55: 27–58.

Rafols, I., and M. Meyer

2010 "Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience." *Scientometrics*, 82: 263–287.

Rao, H., P. Monin, and R. Durand

2005 "Border crossing: Bricolage and the erosion of categorical boundaries in French gastronomy." *American Sociological Review*, 70: 968–991.

Reskin, B.

2003 "Including mechanisms in our models of ascriptive inequality." *American Sociological Review*, 68: 1–21.

Rhoten, D., and A. Parker

2006 "Risks and rewards of an interdisciplinary research path." *Science*, 306: 2046.

Rhoten, D., and S. Pfirman

2007 "Women in interdisciplinary science: Exploring preferences and consequences." *Research Policy*, 36: 56–75.

Rinia, E., T. van Leeuwen, and A. F. van Raan

2002 "Impact measures of interdisciplinary research in physics." *Scientometrics*, 53: 241–248.

Rosenkopf, L., and P. Almeida

2003 "Overcoming local search through alliances and mobility." *Management Science*, 49: 751–766.

Ruef, M., and K. Patterson

2009 "Credit and classification: Defining industry boundaries in nineteenth-century America." *Administrative Science Quarterly*, 54: 486–520.

Sá, C. M.

2008 "'Interdisciplinary strategies' in US research universities." *Higher Education*, 55: 537–552.

Sanz, L. M., M. Bordons, and M. A. Zulueta

2001 "Interdisciplinarity as a multidimensional concept: Its measure in three different research areas." *Research Evaluation*, 10: 47–58.

Sauer, R. D.

1988 "Estimates of the returns to quality and coauthorship in economic academia." *Journal of Political Economy*, 96: 855–866.

Schilling, M. A.

2005 "A 'small-world' network model of cognitive insight." *Creativity Research Journal*, 17: 131–154.

Schilling, M. A., and E. Green

2011 "Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences." *Research Policy*, 40: 1321–1331.

Schumpeter, J. A.

1912 *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Sgourev, S. V., and N. Althuisen

2014 "'Notable' or 'not able': When are acts of inconsistency rewarded?" *American Sociological Review*, 79: 282–302.

Shi, X., L. A. Adamic, B. L. Tseng, and G. S. Clarkson

2009 "The impact of boundary spanning scholarly publications and patents." *PloS ONE* 4:e6547.

Shrum, W., J. Genuth, and I. Chompalov

2007 *Structures of Scientific Collaboration*. Cambridge, MA: MIT Press.

Singh, J., and L. Fleming

2010 "Lone inventors as sources of breakthroughs: Myth or reality?" *Management Science*, 56: 41–56.

Sobey, A. J., N. C. Townsend, C. D. Metcalf, K. D. Bruce, and F. M. Fazi

2013 "Incorporation of early career researchers within multidisciplinary research at academic institutions." *Research Evaluation*, 22: 169–178.

Trapido, D.

2015 "How novelty in knowledge earns recognition: The role of consistent identities." *Research Policy*, 44: 1488–1500.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones

2013 "Atypical combinations and scientific impact." *Science*, 342: 468–472.

van Dalen, H., and K. Henkens

2005 "Signals in science—On the importance of signaling in gaining attention in science." *Scientometrics*, 64: 209–233.

van Rijnsoever, F. J., and L. K. Hessels

2011 "Factors associated with disciplinary and interdisciplinary research collaboration." *Research Policy*, 40: 463–472.

Wagner, C. S., J. D. Roessner, K. Bobb, J. Thompson Klein, K. W. Boyack, J. Keyton, I. Rafols, and K. Börner

2011 "Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature." *Journal of Informetrics*, 5: 14–26.

Wagner-Dobler, R.

1997 "Self-organization of scientific specialization and diversification: A quantitative case study." *Social Studies of Science*, 27: 147–170.

Wang, J.

2013 "Citation time window choice for research impact evaluation." *Scientometrics*, 94: 851–872.

Weick, K. E.

1979 *The Social Psychology of Organizing*. Reading, MA: Addison-Wesley.

Weitzman, M. L.

1998 "Recombinant growth." *Quarterly Journal of Economics*, 113: 331–360.

Whitley, R.

2000 *The Intellectual and Social Organization of the Sciences*. Oxford: Oxford University Press.

Wuchty, S., B. F. Jones, and B. Uzzi

2007 "The increasing dominance of teams in production of knowledge." *Science*, 316: 1036–1039.

Yegros-Yegros, A., I. Rafols, and P. D'Este

2015 "Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity." *PLoS ONE*, 10: e0135095.

Zitt, M., S. Ramanana-Rahary, and E. Bassecoulard

2005 "Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation." *Scientometrics*, 63: 373–401.

Zuckerman, E. W.

1999 "The categorical imperative: Securities analysts and the illegitimacy discount." *American Journal of Sociology*, 104: 1398–1438.

Zuckerman, E. W.

2005 "Typecasting and generalism in firm and market: Genre-based career concentration in the feature film industry, 1933–1995." In C. Jones and P. H. Thornton (eds.), *Research in the Sociology of Organizations*, 23: 173–216. Bingley, UK: Emerald Insight.

Zuckerman, E. W., T.-Y. Kim, K. Ukanwa, and J. von Rittmann

2003 "Robust identities or nonentities? Typecasting in the feature-film labor market." *American Journal of Sociology*, 108: 1018–1074.

Authors' Biographies

Erin Leahey is a professor in the School of Sociology at the University of Arizona, Tucson, AZ 85721 (e-mail: leahey@arizona.edu). Her research focuses on scientific practice and scientific careers, with a special emphasis on specialization and interdisciplinarity. In a current NSF-funded project (award #1461989), she is moving up to an organizational level of analysis to examine universities' commitment to interdisciplinarity and how it affects valued social and economic outcomes. She received her Ph.D. in sociology from UNC Chapel Hill.

Christine M. Beckman is a professor of management and organization at The Robert H. Smith School of Business, University of Maryland, 7621 Mowatt Lane, College Park, MD 20740 (e-mail: cbeckman@umd.edu). Her research focuses on organizational learning, interorganizational networks, and entrepreneurship, particularly on how collaborative relationships and diverse experiences facilitate organizational change. She received her Ph.D. in organizational behavior from the Graduate School of Business at Stanford University.

Taryn L. Stanko is an assistant professor of management at The Orfalea College of Business, California Polytechnic State University, 1 Grand Avenue, San Luis Obispo, CA 93407 (e-mail: tstanko@calpoly.edu). Her research focuses on virtual work collaboration, organizational control, and the management of work–nonwork identities, as well as the role that the use of communication technology plays in each of these areas. She received her Ph.D. in management from the Paul Merage School of Business at the University of California, Irvine.