

## **Gender diversity in the management field: Does it matter for research outcomes?**

**[PREPRINT]**

**Published version available here:**

Nielsen, M. W., & Börjeson, L. (2019). Gender diversity in the management field: Does it matter for research outcomes? *Research Policy*, 48(7), 1617-1632.

## Gender diversity in the management field: Does it matter for research outcomes?

**Mathias Wullum Nielsen<sup>a1</sup>, Love Börjeson<sup>b</sup>**

<sup>a</sup>The Danish Centre for Studies in Research and Research Policy, Department of Political Science, Aarhus University, Bartholins Alle 7, 8000, Aarhus, Denmark

<sup>b</sup>Graduate School of Education, Stanford University, Stanford, 94305, CA.

### Abstract

This study examines the relationship between gender diversity and research outcomes. Existing research on the topic primarily focuses on how team gender diversity influences scholarly productivity in terms of citations and publication rates. Far less attention has been devoted to the question of how the intellectual contents of research disciplines change as they become more gender diverse. Drawing on a global sample of more than 25,000 management papers, we use natural language processing techniques, correspondence analysis and regression models to illuminate impact-, content- and status-related dimensions of gender diversity in management research. In regression models adjusting for geographical setting, institutional prestige and collaboration patterns, we find no effects of team gender diversity on per-paper scientific impact. In contrast, our analyses converge to yield a broadly consistent pattern of gender-related variations in research focus: women are well-represented in social- and human-centered areas of management, while men comprise the vast majority in areas addressing more technical and operational aspects. Our findings corroborate recent sociological research suggesting that cultural norms and expectations are channeling women and men towards different areas of work and study. We argue that the broadened repertoire of perspectives, values and questions resulting from gender diversity may render management research more responsive to the full gamut of societal needs and expectations.

**Keywords:** Gender diversity; Research outcomes; Management research; Citations; Topic modeling; Bibliometrics; Research questions

**Classification codes:** Organizations; Research Institutions; University; Inequality

---

<sup>1</sup>Corresponding author at: Danish Centre for Studies in Research and Research Policy, Department of Political Science, Aarhus University, Bartholins Alle 7, 8000, Aarhus, Denmark. Email Address: mwn@ps.au.dk

## 1. Introduction

Universities and policymakers increasingly highlight a possible link between gender diversity (GD) and excellence as a key motive for promoting women's participation in research (European Commission, 2012; Global Research Council, 2016; Huyer, 2015; Maes et al., 2012; Royal Society, 2017; Valantine and Collins, 2015). Proponents argue that gender diversity is good for performance and creativity, but research in social psychology and management spurs inconsistent and conflicting findings on the matter (for a meta-analysis see Van Dijk et al., 2012). Little is also known about the potential links between GD and research outcomes.

Here, we develop a framework to examine the effects of GD in academic research. Most existing research on the topic focuses on how GD relates to team-performance in terms of publication rates and citation impact; but the particular nature of academic work makes it pertinent to combine such approaches with broader considerations about how research outcomes are reconfigured as traditionally underrepresented groups take their place in scientific disciplines (whether women entering traditionally male-dominated disciplines, or men entering traditionally female-dominated disciplines).

Research on sex segregation demonstrates how societal norms and expectations encourage gendered career choices in higher education and the labor market (Cech, 2013; Charles and Bradley, 2002, 2009). We argue that similar gendered processes may spur differences in researchers' choices of research topics within given disciplines. Illuminating such differences may add important new dimensions to the understanding of how GD influences the ideas, discoveries and innovations delivered by research organizations and vice-versa. Hence, we propose a dual-perspective on GD in research with a view to both the quantifiable performance outcomes at the team level (here measured by citation impact per paper) and the broader changes in knowledge outcomes at the systemic level (here measured by variations in research priorities and agendas).

We have three overarching objectives: (A) to examine the effect of team GD on the per-paper citation impact of author groups; (B) to explore potential gender-related variations in the scientific focus and interests of research teams; and (C) to illuminate potential differences in the status of research areas with varying male and female participation.

We limit our focus to the social-science literature on management (henceforth management research). While management research has played an important role in putting employee diversity

on the scholarly and public agenda, surprisingly little research examines how GD is connected to the work outcomes of management scholars themselves. Indeed, this article represents the first attempt to empirically explore the potential connections between GD and research outcomes (in terms of citation impact and research focus) in the management field.

Our analysis is based on a global sample of 27,676 management papers. We use topic modeling, a natural language-processing technique suitable for studying content in large corpora of academic texts (McFarland et al., 2013), to identify topics in the management literature and to illuminate variations in women's and men's participation across topics. Further, we use bibliometric data to explore differences in the scholarly status of female- and male-typical research topics, and to examine the effect of team GD on citation impact per paper. Here, we measure team GD based on the gender composition of the author group.

We find trivial effects of team GD on the citation impact of management researchers. Our analyses, however, converge to yield a broadly consistent pattern of gender-related content variations: Women authors and women-dominated author groups typically pose different questions and adopt different perspectives on management research than studies authored by men. Women are, for instance, more likely to engage in social and human-centered areas of management, while men (on average) gravitate towards the more technical and operational aspects. Finally, contrasting evidence reported for other social-science disciplines, we do not find that these differences can be fully explained by “ghettoization processes” that channel women into research areas of lower scholarly status (measured by journal prestige).

Compared to previous studies, our combination of bibliometric and content-focused approaches makes it possible to interpret associations between GD and research outcomes on a broader scale: GD may influence research in myriad ways, and not all of these are meaningfully captured using conventional performance metrics such as citations and publication rates. By exploring content-related variations in knowledge outcomes, we raise important questions about how to conceptualize and measure the impact of GD in research. Specifically, we argue that the broadened repertoire of perspectives, values and questions resulting from GD may render management research more responsive to the full gamut of societal needs and expectations.

## **2. Framework**



Inspired by the literature on cultural diversity in the workplace (Ely and Thomas, 2001; Merrill-Sands et al., 2000), we see gender as an identity category that shapes cognition, experience, world-view and perspective. Gendered aspects of identity, we posit, constitute embodied “ways of being” acquired through upbringing and invoked through culture and social interaction (McLeod, 2005); and researchers carry these ways of being into their day-to-day work activities. To borrow a concept from French sociologist Pierre Bourdieu (1990, pp. 54-64), the gendered aspects of identity are part of the researcher’s “habitus” – a set of cognitive and motivating dispositions, guiding future actions, thoughts and perceptions. Consequently, *gender diversity* refers to the combination of different beliefs, preferences, experiences and expressions structured along gendered lines. The potential benefits of GD are linked to cognitive diversity (Page, 2008). Cognitive diversity refers to the broader variety of values, beliefs, experiences and perspectives present in an entity, organization or system; and this type of diversity is generally considered desirable for the evolvment of knowledge production (Page, 2008; Stirling, 2009).

The outcomes of GD can be studied at different levels of analysis. While most of the existing literature is concerned with quantifiable team-level parameters such as publication productivity and scholarly impact, much can also be gained from adopting a content-focused perspective on the topic. Here, we briefly review both perspectives and specify the research questions guiding our analysis.

## ***2.1 Gender diversity and team performance***

The scholarship on GD and team performance in the academy is scarce and inconclusive. Based on a systematic review of the literature, we found five studies addressing this theme.

<sup>2</sup> As displayed in Table 1, two of these show benefits of GD at the team level (Campbell et al., 2013; Saá-Pérez et al., 2015), while the remaining report no discernable effects (Joshi, 2014; Lungeanu and Contractor, 2015; Stvilia et al., 2011). Even in studies that highlight positive effects, the potential benefits tend to vary depending on level of GD in the team and what type of performance is considered (Nielsen et al. 2018a).

The broader literature on gender and scientific performance is also characterized by inconsistencies. The majority of research on gender and productivity finds that women, on average,

---

<sup>2</sup>The literature survey was carried out based on systematic searches in SCOPUS and Web of Science. The search strategy is documented in Nielsen et al. (2018a).

publish fewer peer-reviewed papers than men (for an overview of the literature, see Mairesse and Pezzoni, 2015). However, studies from the US indicate that this average gender gap is diminishing over time (Sax et al., 2002; Xie and Shauman, 2003). Sociological research also suggest that structural variables such as employment rank, access to resources, university ranking, department prestige and type of appointment all partly explain the persistent gender gap in productivity (Allison and Long, 1990; Bland et al., 2006; Xie and Shauman, 1998). Others have related the so-called “productivity puzzle” to factors such as level of disciplinary specialization (Leahey, 2006), differences in co-authorship patterns (Abramo et al., 2013; Bozeman and Corley, 2004), a greater relative concentration of men among the most prolific authors (Abramo et al., 2009), and differences in teaching loads and relative time available for research (Leišytė, 2016; Taylor et al., 2006).

Extant research on gender and citation impact (the performance indicator employed in this study) also shows mixed results. Studies have reported slightly higher citation rates in favor of both women (e.g. Borrego et al., 2010; Long, 1992) and men (e.g. Aksnes et al., 2011; Caplar et al., 2017; Larivière et al., 2013); while other studies report no discernable gender difference (e.g. Lerchenmueller and Sorenson, 2018; Nielsen, 2016; Slyder et al., 2011; Symonds et al., 2006). Generally, the results of this literature tend to vary depending on study design, geographic location (Elsevier, 2017; Sugimoto et al., 2015), the discipline in focus (Duch et al., 2012; Gonzalez-Brambila and Veloso, 2007; Larivière et al., 2011; van Arensbergen et al., 2012) and the gender composition of the field (Ferber and Brün, 2011).

However, large-scale, multivariate analyses focusing on management science do not to report any discernable gender differences in citation impact per-paper (Judge et al., 2007; Nielsen 2017a; Podsakoff et al., 2008).

The vast literature assessing the link between GD and team performance outside the academy is also characterized by mixed findings. As a result of stereotype threat, negative status dynamics and in-group favoritism, GD is often argued to impede rather than strengthen team functioning and performance (Harrison et al., 1998; Horwitz, 2005; Hülshager et al., 2009; Mannix and Neale, 2005; Pelled et al., 1991; Williams and O'Reilly, 1998). However, a recent meta-analysis challenges the cumulative evidence backing this conclusion. Based on a comparison of 146 studies of team diversity, van Dijk and colleagues (2012) argue that the predominantly negative effects of demographic diversity (including GD) may result from a bias in how

performance is evaluated in the team-diversity literature. When performance assessments rely on subjective evaluations (e.g. assessments by a supervisor or leading group member), gender diverse teams are slightly outperformed by more homogeneous groups. However, when performance ratings are based on objective performance assessments (e.g. financial performance or number of correct answers in a problem-solving task) the difference between the groups become marginal statistically insignificant. These findings illustrate that social processes of negative stereotyping are more salient among evaluators than in the actual teams being evaluated, hence challenging the idea of diversity as process-loss. In summary, when juxtaposing the outcomes of the existing literature, GD appears to be a poor predictor of team performance in terms of per-paper citation impact. Therefore, we contend that:

- a. Differences in the per-paper citation impact of gender-homogeneous and gender-diverse author groups will be marginal to non-existing at the aggregate level.*

## **2.2 Broadening the perspective on GD**

As highlighted earlier, GD may influence research outcomes in myriad ways, and not all of these are meaningfully captured using quantifiable performance metrics at the team level. Here we propose a complementary approach for examining gender-related variations in research outcomes. This approach offers a bird's eye view that enables us to see how the changing gender demographics of scientific fields are connected to the viewpoints, questions and areas addressed by researchers.

A growing sociological literature demonstrates how cultural ideals and beliefs about gender are channeling women and men towards different occupations and fields of study. According to this literature, gender operates as a primary frame that shapes social relations and identities; and women and men draw on this frame in their expressions of personal values and beliefs (Cech, 2013; Charles and Bradley, 2009; Ridgeway, 2011). As Charles and Bradley (2009, p. 927) observe: “modern individuals are deeply invested in beliefs about gender difference, and these are embedded in virtually all organizational structures and interactional contexts, including families, labor markets and educational systems”. Cultural beliefs about gender hold women and men accountable to different stereotypical ideals about appropriate gendered behaviors, and this has implications for their curricular affinities and career aspirations (Cech, 2013; Charles and Bradley,

2009). For instance, people- and service oriented job-characteristics involving interpersonal interactions are typically seen as more suitable for women, while men are presumed to excel in tasks requiring strong analytical skills and interaction with things (Charles and Bradley, 2009; Eccles, 2007; Fernandez and Friedrich, 2011). Here it is important to note that such stereotypical presumptions often influence career choices on a subconscious level (Cech, 2013). As Ridgeway (2011, p. 156) notes, most of us “implicitly fall back on cultural beliefs about gender to frame what it means to make life choices that ‘express’ ourselves”.

The influence of the gender frame on curricular affinities is well-known. Charles and Bradley (2009) demonstrate that women’s participation in higher education is highest in the humanities, social sciences and health sciences, while men dominate science- and technology-related areas (see also Charles and Bradley, 2002, Charles and Grusky, 2004; England and Li, 2006). Existing research also documents gender differences in medical students’ specialty choices with women concentrated in areas such as pediatrics and gynecology and men in surgery and orthopedics (Alers et al., 2014).

Here we argue that the gender frame may also influence researchers’ selection of topics within given disciplines. Broader cultural influences about appropriate gender-typed work may draw a disproportionate number of women towards certain research topics and men towards others. Specifically, we posit that:

*b. Consistent gender variations can be detected in the research focus of management scholars at both individual and group-levels*

This proposition already finds some support in research pertaining to the neighboring social-science disciplines. Dolado and colleagues (2012) demonstrate clear variations in women’s and men’s primary areas of specialization in economic research: women mainly engage in areas such as health, education, welfare, labor and demographic economics, while men focus on agricultural economics, fluctuations/business cycles, general equilibrium and cooperative games, comparative systems, and corporate finance. Light (2013) distils ten specialization clusters among U.S. sociologists, of which women are overrepresented in gender-race-sexualities, family-demography-youth, and medical-mental-health-aging, and underrepresented in political-comparative-economic, and quantitative-demography-family. Both studies suggest that these differences are

linked to variations in the scholarly status of female- and male-typical subfields. Specifically, they show that female-typical subfields score lower on parameters of journal prestige than male-typical subfields. These findings point to the continuing relevance of Reskin's and Roos' (2009) pioneering work on the gendered ghettoization of occupational fields: when women make inroads into male-dominated occupations, subtle forms of gender segregation tend to persist. For instance, women may, self-select or be 'ghettoized' into less prestigious work areas. It also resonates with work in experimental economics, showing that women (on average) are less likely than men to select into competitive work environments (see e.g. Niederle and Vesterlund, 2007). These considerations lead us to the third and final proposition:

- c. Female authors and female-dominated author groups are more likely to be concentrated in lower-status areas of management research than male authors and male-dominated author groups.*

### **3. Data and Methods**

Data for this study were extracted from Web of Science's Social Science Citation Index (core collection). We collected all available information (including titles, abstracts, reference lists and author information) for peer-reviewed articles, written in English, indexed under the subject category *management*. Web of Science provides consistent full first-name author information for papers published since 2007, and for the purpose of gender analysis (see below), we limited our sample to studies published from 2007 through 2013, resulting in 46,549 papers. Figure S1.1. (Appendix S1) details the three data exclusion steps, including the assignment of gender to author first-names, leading to the final sample of 27,676 papers (59% of the full population) and 71,322 authorships. The somewhat fragmented coverage of anthology articles, conference proceedings and monographs in Web of Science represents a key limitation of our sample. However, since academic journals constitute the primary outlet for scholarly dissemination in the management field (Baruch, 2001), this bias may be less problematic than in other parts of the social sciences and humanities. On average, 55% of the references cited by papers included in this analysis go to

other papers covered by Web of Science (i.e. WoS coverage<sup>3</sup>). The average WoS coverage is slightly different for papers authored by women and men (male= 55.8%; female= 53.9%). Appendix S1 provides specifications on WoS coverage across gender groups and topics (Tables S1.1 and S1.2). The frequency distribution of papers per subject category and the average WoS coverage per subject category and research topic are also presented in this document (Tables S1.3 and S1.4).

### 3.1 Gender disambiguation

The name-to-gender assignment algorithm, Gender API (2016), was used to determine the gender of each author in the sample.<sup>4</sup> Gender API predicts the gender associated with a given first name, while accounting for variations in gender connotations across countries. In a recent bench-mark and comparison of five name-to-gender inference services, Gender API was evaluated as the best-performing service (Santamaría and Mihaljević, 2018). For each first name and country pair, Gender API provides a probability estimate of the accuracy of the gender prediction ranging from 50 to 100. We converted Gender API estimates into a single indicator specifying the probability of a name belonging to a female researcher, denoted  $f$ .  $f$  probability-scores range from 0 to 1 with values closer to 1 indicating a higher likelihood of the author being a woman.  $f$  probabilities were used to compute a weighted indicator,  $f\_weight$ , specifying the general gender composition for all author groups in the sample. This indicator also ranges from 0 to 1 with higher values indicating a higher representation of women.<sup>5</sup> A validation study based on a manual, web-based identification of author-gender for more than 500 authorships in our dataset, confirms the accuracy of Gender API. The false-positive rate for male authorships was 2.6% and the false-positive rate for female authorships was 5.8% (see Appendix S1).

### 3.2 Procedure

---

<sup>3</sup> The WoS coverage is calculated as the mean percentage of references cited by a WoS article to other articles covered by WoS. A WoS-coverage score of 0.5 specifies that 50% percent of the cited references in a paper go to other papers covered by WoS.

<sup>4</sup> Gender API supports gender assignment for 1,871,879 names from 178 countries.

<sup>5</sup> Due to the uncertainty of Gender API scores between .10 and .90, the use of this indicator is only applicable at the aggregate level. A paper with  $f_w = 0.8$  could, for instance, be authored by two women, while another all-male paper might have  $f_w = 0.2$ . However, these variations should level each other out at the aggregate level.

As highlighted by Rafols and colleagues (2012), no singular indicator is fully capable of capturing the multiple dimensions of research outcomes. Inspired by their approach, we rely on various partial indicators and complement these with computer-assisted methods for text analysis (topic modeling). This combination serves to broaden the empirical framework for interpreting the potential associations between GD and research outcomes. The analysis took place in three steps. First, we used logistic regressions and generalized linear models (henceforth GLMs) to examine the link between team GD and per-paper citation impact. Second, we employed topic modeling, correspondence analysis and logistic regressions to explore gender variations in the research focus of management scholars. Finally, we used GLMs and logistic regressions to illuminate gendered variations in the scholarly status of the research topics identified in our topic model. Here we provide specifications on data and measurement for each step in the analysis.

### **3.3 Citation impact**

Research collaborations in management research are typically not confined to stable, long-term team constellations. Management scholars are part of multiple constellations, of varying sizes, involving different colleagues, during the same time period. This makes it difficult to systematically estimate the impact of team GD on publication productivity. Thus, we limit our examination of the link between GD and research performance to citation rates per paper. Irrespective of methodological shortcomings (Gläser and Laudel, 2007), citation performance represents a widely used proxy for recognition and success in the academy (Cole and Cole, 1973; Cronin, 1981; Judge et al., 2007). Specifically, we rely on the following two proxies of citation impact per paper: PP top-10% and citation scores (CS). PP top-10% measures the proportion of papers in our sample which, “compared with other papers in the same field in the same year, belong to the top 10% most frequently cited” (CWTS, 2016). Both citation indicators are calculated with a four-year citation window.

To examine how team GD influences a paper’s probability of being among the top-10% most cited in its field, we use binary logistic regression analysis.

To investigate the relationship between the gender composition of author groups and CS, we use GLMs which are well suited for accommodating a variety of non-linear response distributions (McCullagh and Nelder, 1989). Due to overdispersion, the negative binomial distribution with log-link function is argued to be the best suitable model solution for citation data

(Bornmann et al., 2008; Mingers and Xu, 2010). Robustness checks have been carried out to examine the sensitivity of the results to alternative model specifications. Specifically, we ran linear regression models with a log-transformed CS variable, and negative binomial regressions with field-normalized citations as the outcome variable.<sup>6</sup>

The main predictors in our four models are *f\_weight* (model 1 and 3) and *Diversity index* (model 2 and 4). The *Diversity index* is based on transformed *f\_weight* probabilities. Values range from 0 to 1, with values closer to 1 indicating a more balanced share of women and men in the author group. Both models include covariates to adjust for geographical setting, institutional prestige, cross-institutional collaboration, international collaboration and self-citations. In addition, the GLM models with CS as outcome variable adjust for publication year.

Additional regression models were carried out to examine possible curve-linear effects for *f\_weight* and *Diversity index*, and to test whether the main predictors interacted with the geographical groupings, self-citations and international collaboration (for variable specifications, see Table A1, Appendix). We also reran the generalized linear models and logistic regressions with the following quintile-based categories of gender compositions: *male dominated* (*f\_weight*: 0-0.2), *moderate male* (*f\_weight*: >0.2-0.4), *mixed* (*f\_weight*: >0.4-0.6), *moderate female* (*f\_weight*: >0.6-0.8) and *female dominated* (*f\_weight*: >0.8-1) (for an overview of the prevalence of each category in the data-set see Appendix S1, Figure S1.2.).<sup>7</sup>

### 3.4 Content variations

In our analysis of gender variations in the content of management research, we limit our focus to topics. Following Gläser and colleagues (2015, p. 1008), we conceptualize a topic as “a focus on theoretical, methodological or empirical knowledge that is shared by a number of researchers and thereby provides (...) a joint frame of reference for the formulation of a problem, the selection of methods or objects, the organizations of empirical data, or the interpretation of data”. The complex properties of topics limit the relevance of using bibliometric methods, which typically confine to a limited set of paper properties, such as authors, journals, references and key-terms (Gläser et al.,

---

<sup>6</sup>The field-normalized citation scores are calculated by dividing the annual number of citations assigned to a paper with the average number of citations given to all publications of the same type, in the same subject area, for the same year (excluding self-citations). This normalization makes it possible to compare the scholarly impact of papers across time-periods and sub-fields (Waltman et al., 2011).

<sup>7</sup> Single-authored papers have been assigned to categories based on their *f*-score.



2015). To overcome this limitation, we combine topic modeling, correspondence analysis and regression analysis. This combination allows us to delve deeper into the content of scholarly papers, while keeping in view the overarching associations between, and patterns across, documents.

Topic modeling is a natural language processing technique suitable for studying content in large samples of texts. The method has proven highly useful in extracting semantically meaningful topics from scholarly documents (see e.g. McFarland et al., 2013; Nichols, 2014; Talley et al., 2011). Put simply, this technique offers a probabilistic method for inferring the topic structure of a given sample of documents, through a “distant reading” (Moretti, 2005). Since the thematic structure of our text sample has not been studied previously, we use the unsupervised Latent Dirichlet Allocation algorithm (LDA) to provide a baseline statistical image of the topics in the corpus (Blei et al., 2003). Conceptually, each abstract in the corpus is regarded as a bag of words, where the distribution of the observable words in each bag is guided by a latent distribution of topics (Blei et al., 2003; Mohr and Bogdanov, 2013). In this sense, each text reflects a multinomial distribution over topics, and each topic a multinomial distribution over words (McFarland et al., 2013).

For this study, we use the LDA implementation available through the Stanford Topic Modeling Toolbox (Ramage and Rosen, 2011). To optimize model parameters, we have run numerous minor reconfigurations, testing various solutions with respect to stemming and stop-words. A crucial procedural step concerns determining the appropriate number of topics. In our case, perplexity parameters have been measured for topic solutions spanning from 25 to 100. Topics were trained in 1000 iterations (for a review of the perplexity procedure, see Asuncion et al., 2009). Parameters such as minimum document loadings per topic, minimum word lengths, level of topic smoothing and removal of non-meaningful words have been adjusted to minimize perplexity. Appendix S2 details the final parameters and outcomes of this procedure, resulting in an optimal model of 36 topics. Topics are given loadings that indicate their prevalence (given the size of the present corpus)<sup>8</sup>. Words are given loadings that indicate their probable frequency in the topic given the prevalence of the topic. We use the top-20 words per topic (in terms of loadings) to interpret and name the topics. Further, distinctive abstracts for each topic (i.e. abstracts with

---

<sup>8</sup> These loadings specify the per-document topic distributions.

particularly high loadings for a given topic) were used to qualify and corroborate our interpretations of the summary lists (summary lists and examples of distinctive abstracts are available in Appendix S3).

Making sense of topic models requires in-depth expertise in the area under study. One of the authors of this article is a management scholar by training, which proved highly beneficial in the naming and interpretation of topics. Further, to make robust and transparent conclusions on the strengths and weaknesses of our selected topic solution (and the plausibility of our interpretations), we juxtaposed the model with outcomes of complementary techniques for science mapping. Researchers have already documented how a combination of different techniques (e.g. co-word and co-citation analysis) can contribute to clearer view of the cognitive structures of a given scientific discipline (Braam et al., 1991a, 1991b). Specifically, we have analyzed differences and similarities between the outcomes of three techniques: topic modeling, co-citation analysis and co-word mapping (for specifications on the methods and outcomes of this analysis, see Appendix S5).

Our comparison shows clear overlaps between the outcomes of the three approaches, hence confirming the aptitude of our topic solution in capturing key research frontiers in the management literature. Further, a careful inspection of the thematic structure underpinning the co-citation network and co-word map serves as converging evidence for the plausibility of our interpretations of model results. However, a few topics in our model are too general (“time” and “literature reviews”) while others tend to be too specific or niche (“socio-economic policies” and “crisis management”) to be captured by the complementary science mapping techniques. The uncertainties associated with these topics should of course be taken into account in the analysis.

We used correspondence analysis (CA) to examine associations between the gender composition of author groups and topics. This technique is suitable for exploring the latent dimensionality and associations between data in the absence of a strong a priori theory (Borgatti et al., 2013). Results of the CA are displayed in a joint graphical space using symmetrical biplots. In this space, proximity between objects (i.e. gender categories and topics) indicates high degrees of correspondence, while distance represents low degrees of correspondence (Greenacre, 2007). Robustness and quality estimates for the CA (i.e. significance of the chi-squared distances, inertia-levels, and squared cosines for each object) are available in Appendix S4. Tables S4.7 and S4.8 (Appendix S4) specify the contribution biplots per topic (given as rows) and gender category

(given as columns) to the overall solution. In the CA, we rely on the quintile-based gender categories presented in Section 3.3.

Finally, we complement the CA with two logistic regression analyses predicting 1) the gender of each author contributing to a given paper, and 2) the gender of the first author contributing to a given paper<sup>9</sup>, based on topic-loadings.  $f$ -probabilities were used to compute the dichotomous outcome variable *Woman\_category*. We excluded all authors with  $f$ -probabilities higher than .10 and lower than .90, meaning that only male and female authors with high Gender API scores ( $>90$ ) were included in the analysis ( $N= 65,194$ ). Loadings for the 36 topics in our topic model were used as main predictors in these models. Additional covariates were included to prevent potential misspecifications of the model. 10 categorical covariates created based on geographical groupings, were used to account for area-specific differences in scholarly focus and variations in the participation of women authors. Further, categorical variables based on rankings of business schools and universities (*University Prestige* and *Business Prestige*) were used to capture covariation attributable to institutional prestige (for variable specifications, see Table A1 and A2, Appendix).

### **3.5 Status variations**

Journal prestige serves as an important status-based mechanism for allocating opportunities and rewards in the social sciences (Grant and Ward, 1991; Light, 2013; Nielsen, 2018b). To examine gender-related status variations across topics, we rely on journal scores per paper (henceforth JS). This measure specifies the average citation impact for the annual volume of research and review articles for a particular journal in the year it published a given study. JS scores correspond to Clarivate Analytics' Journal Impact Factor.

Despite fallacies, Journal impact factors are a commonly used proxy for status and visibility in the academy (Archambault and Larivière, 2009; Cole and Cole, 1973, pp. 46–60), and existing research suggests that prestigious journals in the social sciences tend to be biased towards certain topics, resulting in perceived within-field, topic-related status-variations (Light, 2013; Rafols et al., 2012). As a complementary approach for examining status-based variations across topics, we have considered fairly established journal lists and rankings in business and

---

<sup>9</sup> We ran a specific analysis for first authorships, since first authors are generally presumed to be leading authors in the social sciences (Lariviere et al. 2016).

management. Specifically, we use the lists of journals included in the UT Dallas (UTD) and Financial Times (FT) research rankings of business schools, and the designated top journals (level 4\*) in the Academic Journal Guide (AJG) developed by the Chartered Association of Business Schools in the UK.

To examine associations between topics and journal scores, we use linear regression. The outcome variable (JS) has been log-transformed to approximate a normal distribution (log-JS). To investigate variations in the presence of the 36 topics in papers published in top journals in UTD, FT and AJG, we use binary logistic regression analysis. Topic loadings for the 36 topics will serve as main predictors in all models. Additional covariates have been included to adjust for geographical area, institutional prestige, cross-institutional collaboration and international collaboration.

#### 4. Results

We begin our analysis by estimating the extent to which the scholarly impact of management papers is affected by the gender composition of the author groups. Table 2 displays odds ratios, confidence intervals and standard errors for the logistic regression analysis that predicts a paper's probability of being among the top-10% most cited in its field (PP top-10%). Table 3 displays the incidence rate ratios, confidence intervals and standard errors for the GLM predicting citations rates per paper (CS). Model 1 (Table 2) and 3 (Table 3) specify the effect attributable to women's participation in the author group (i.e. *f\_weight*), and Model 2 (Table 2) and 4 (Table 3) determine the effect of team-level GD (*Diversity index*). All models adjust for geographical setting, institutional prestige, self-citations, institutional collaboration, and international collaboration.<sup>10</sup> In addition, Models 3 and 4 (Table 3) adjust for publication year.

In accordance with our first conjecture, we find no indicative evidence of a bias in citation rates per paper attributable to the gender composition of author groups. The odds ratios (Table 2) and incidence rate ratios (Table 3) for the main predictors (*f\_weight* and *Diversity index*) are extremely close to 1.0 and the confidence intervals span the line of no difference.

As a robustness check, we ran linear regressions with a log-transformed CS as the outcome variable, and negative binomial regressions with field-normalized citation scores (NCS) as the

---

<sup>10</sup> All predictors and covariates in the regression models have Variance Inflation Factors < 5.

outcome variable. Both of these approaches exhibit similar results; the effects of *f\_weight* and *Diversity index* are trivial in all models and have confidence intervals that span the line of no difference (Table S1.5 and S1.6).

The negative binomial and logistic regressions with quintile-based categorical diversity variables as main predictors also exhibit similar results (Tables S1.7 and S1.8, Appendix S1). Additional regression models including quadratic terms show no curve-linear effects for the main predictors (*f\_weight* and *Diversity index*) (Tables S1.9 and S1.10, Appendix S1). And tests examining whether *f\_weight* and *Diversity index* interacted with the geographical groupings, self-citations and international collaboration do not reach statistical significance (Tables S1.11-S1.14, Appendix S1).

---

## INSERT Table 2 and 3 ABOUT HERE

---

As a second step in the analysis, we explore gender variations in the content of management papers. Table A3 in the Appendix reports the top-5 word lists for each of the 36 topics in the topic model. Significant abstracts and top-20 word lists (including loadings) are available in Appendix S3.<sup>11</sup> The dimensionality of the correspondences between topics and gender categories is shown in Figure 1. The proximity/distance between nodes specify the degree of correspondence between topics, and node-sizes indicate the relative prevalence of topics. Only one dimension in the CA is of statistical and pragmatic significance (metrics and a quality measure of the CA are reported in Appendix S4). This unidimensional distribution of topics can be interpreted as a gender scale spanning from male-dominated author groups and “male-typical” topics on the left side, to female-dominated author groups and “female-typical” topics on the right side. The correspondences between gender categories and topics are consistent along the scale: changing the gender composition will, on the aggregate level, imply a corresponding change in topic focus.<sup>12</sup>

The scale is more densely populated to the right than to the left, and topics gravitate towards the male-typical end of the scale, simply because there are more male-dominated author groups in

---

<sup>11</sup> An alternative topic solution could include more topics, but the distinctiveness of these topics would be lower, as indicated by the perplexity analysis (Appendix S2).

<sup>12</sup> This is true although the combination of authors only allows a certain number of discrete combinations of male-female authors limited by the total number of authors for a paper.

the sample (see Figure S1.2, Appendix S1). Topics located far to the left are thus closer to the discipline's intellectual center than topics located far to the right.

Topics on the left side are characterized by a quantitative methodological imprint (e.g. “consumer economics” and “corporate finance”) and typically address tangible and direct managerial ends (e.g. “operations algorithms” and “predictive modeling”). In general, these topics share an affinity with the Tayloristic origin of management studies, with a focus on production, efficiency and elimination of slack. The managerial activities captured on this side of the scale are technical and non-human: inventories, algorithms, statistical models, supply chains, etc.

The center of the scale is populated by topics gravitating towards the intellectual core of contemporary management research (e.g. “corporate governance”, “multinational business” and “strategic management”). Many of these topics focus on activities located at the upper echelons of the managerial “chain of commands”. They prescribe practices, processes and rules on how to balance the interests of stakeholders, implement new goals and initiatives, expand markets, coordinate collaborative ties, plan and monitor projects, develop new technologies and products, transfer knowledge between entities, etc. At the center of the scale, we also find second-order topics, such as “literature reviews”, and generic methodological topics such as “survey studies”.

On the right side of the scale, social aspects of management take precedence over the functional (e.g. “HRM”, “commitment”, “employee appraisals”). Human-centered perspectives (e.g. employee sentiments, organizational loyalty, interpersonal cooperation, learning, and leadership styles) are here acknowledged as crucial for the day-to-day functioning of organizations. The right side is also populated by more skeptical and sociologically informed perspectives. “Structural inequalities” adopts a critical gaze on the structural and cultural conditions perpetuating gender, race and age inequalities in organizations. “Constructionism” situates the management scholarship in a broader sociological framework seeing organizations as arenas that produce and constrain social identities, practices and change processes. Both of these topics are known to be inclined towards qualitative and interpretivist methodological approaches.

Finally, “healthcare management” is the outlier topic, situated far to the right on the margin of the scale. In addition to being human-centered, this topic's gravitation towards the margin, may be explained by a close affinity with another discipline heavily dominated by women, i.e. Nursing science.

---

## INSERT FIGURE 1. ABOUT HERE

---

Figures 2 and 3 display the outcomes of the complementary binary logistic regression models designed to examine associations between the 36 topics and first-author gender (Figure 2), and the gender of each author contributing to paper (Figure 3). As mentioned earlier, these models take individual authors as opposed to author groups as their basic unit of analysis and adjust for geographical setting and institutional prestige (for model specifications, see Tables S1.15 and S1.16, Appendix S1). The figures give odds ratios and confidence intervals for the loadings of each topic in the topic model, and topics have been ranked based on odds-ratio values. Model estimates should be interpreted as differentials from the reference variables (*Corporate finance* and *Oceania*) (i.e. Odds ratio=1). Hence, some predictors at the top of Figures 2 and 3 have confidence intervals spanning 1.0, due to odds-ratios that closely resemble the values of the excluded topic-variable *Corporate finance*.

Overall, the two regression models add converging evidence to the results of the CA. The distribution of topics along the gender scale in Figure 1 closely resembles the ranked odd ratios displayed in Figures 2 and 3. The ordering of topics is not completely identical across the three figures, but “male” topics (placed far to the left in Figure 1) are all located in the top of Figures 2 and 3, while “female” topics (placed to the right in Figure 1) are clustered together in the bottom of Figures 2 and 3. Healthcare management is the strongest predictor of the outcome variable in the logistic regression models, followed by Structural inequality, HRM and CSR. In contrast, Inventory management, Predictive modeling, Consumer Economics and Operations algorithms are the strongest predictors of the author being male.

To briefly summarize, the findings presented in Figures 1, 2 and 3 converge to yield a broadly consistent pattern. Conditioned on our model specifications, author gender, as posited in our second conjecture, appears to be a robust predictor of content variations in management research. This holds true when taking both author groups and individual authors as the unit of analysis, and when adjusting for geographical setting and institutional prestige.

---

## INSERT FIGURE 3. ABOUT HERE

---

As a third and final step in the analysis, we examine potential status differences between female-typical and male-typical topics. Figure 4 displays the unstandardized beta-coefficients and confidence intervals for the main predictors (i.e. the topic variables) in linear regression designed to illuminate status variations across topics. The model takes papers (not individual authors) as its unit of analysis. The outcome variable is the log-transformed journal score, and additional predictors have been included to adjust for geographical setting, institutional prestige, institutional and international collaboration, and the gender composition of author groups (for model specifications, see Table S1.17, Appendix S1). Once again, model estimates should be interpreted as differentials from the excluded reference variables (*Corporate finance* and *Oceania*). The 35 topics included in this figure are ranked based on the unstandardized beta coefficients. The topics most likely to accrue high journal scores are placed at the bottom of the figure. To visualize whether women are less likely to engage in high status areas of management research, topics located at the right side of Figure 1 and at the bottom of Figure 3 have been marked in bold.

---

### INSERT FIGURE 3. ABOUT HERE

---

Overall, the figure reveals high and inconsistent levels of status variance. Two “female-typical” topics are located at the upper part of the figure (“HRM” and “Leadership”), while the remaining “female-typical” topics are distributed from the middle to the bottom of the figure. Indeed, none of the three most female-dominated topics (“Healthcare management”, “Structural inequality” and “HRM”) are located at the bottom of Figure 3, but the same argument can be made for the “male-typical” topics (“Inventory management”, “Operations algorithms” and “Predictive modeling”) skewed most to the left in Figure 1 and located at the top of Figure 3. Further, the regression model (Table S1.17., Appendix S1) shows no discernable effect of the covariate *f\_weight* in predicting the log-transformed journal score ( $\beta = .004$ , CI:  $-.006$  to  $.013$ ). To examine whether the topic variables “absorb” the effect of the association between *f\_weight* and log-JS, we reran the same regression excluding all topics from the model (Table S1.18., Appendix S1). In this reduced model, the effect of *f\_weight* remains trivial with a 95% confidence intervals that span the line of no difference ( $\beta = .002$ , CI:  $-.008$  to  $.012$ ).



The logistic-regression models examining variations in the coverage of the 36 topics in “top journals” (UTD, FT and AJG) show a somewhat similar pattern (see Figures S1.3-1.5, and Tables S1.19-1.21, Appendix 1). Female-typical topics are scattered from the top to the lower-middle part of the three Figures (S1.3-1.5). However, they are consistently underrepresented among the topics with the highest likelihood of being covered by top journals. This suggests some level of status variance. Again, the regression models show no discernable effect of the covariate *f\_weight* in predicting coverage in top journals (i.e. UTD, FT and AJG). In reduced models without the topic variables, the effect of *f\_weight* remains inconsequential (Tables S1.22-1.24, Appendix S1).

In sum, we find insufficient evidence to confirm third conjecture – that female authors and female-dominated author groups are more likely to be concentrated in lower-status areas of management research than male authors and male-dominated author groups. However, female-typical research areas are consistently underrepresented among the topics with the highest likelihood of being covered by top journals according to journal lists developed by UT Dallas, Financial Times and the Chartered Association of Business Schools in the UK.

## **5. Concluding discussion**

This study was piqued by an assumption underpinning current policy efforts to promote women’s status and advancement in the academy – that gender diversity matters for research outcomes. Determining the links between gender diversity and research outcomes is a difficult task involving both conceptual and methodological challenges. The existing literature primarily focuses on how team GD influences citation and publication performance. This study has sought to broaden the notion of research outcomes to also encompass the intellectual contents of scholarly work. As we have attempted to illustrate, much can be gained from adopting a systemic perspective on GD. Such a perspective enables us to see how research outcomes are reconfigured as outsiders gain increasing representation in scientific disciplines.

Specifically, we have focused on the relationship between GD and research outcomes in the management field. Our investigations were guided by “converging partial indices” (Rafols et al., 2012) enabling a multi-faceted analytical focus on scholarly contributions. We have combined topic modeling, correspondence analysis and regression models to illuminate performance-, content- and status-related dimensions of GD in management research.

As a first step in the analysis, we investigated possible associations between the gender composition of author groups and the scholarly impact of their papers. Adjusting for geographical setting, institutional prestige, self-citation rates and collaboration patterns, we found trivial effects for both of our main predictors, *f\_weight* and the *Diversity index*.

This does, however, not imply that gender diversity is without possible benefits for scientific groups. Social psychological and management research demonstrates the important role played by contextual factors in shaping diversity effects - from the individual attitudes influencing collaborative problem-solving to the broader cultural norms shaping gender relations in the work place (see e.g. Homan et al., 2007; Joshi, 2014; Nishii, 2013; Schneid et al., 2015). Similar studies of the contextual determinants differentiating high-performing teams from medium- and low-performing teams in science, will be critical in moving the literature on team GD and research outcomes forward (Nielsen et al. 2018a).

Our results parallel previous research on the association between individual-author gender and citation impact in management research that also demonstrate trivial effects. Here, it is important to note that female-dominated author groups may still be at a disadvantage in performance assessments, due to gender-based double standards in evaluative judgements (see e.g. Botelho and Abraham, 2017; Moss-Racusin et al., 2012). For instance, recent research focusing on the life sciences suggest that women on average gain fewer returns to citations in terms of career advancement than their men colleagues (Lerchenmueller and Sorenson, 2018).

As a second step, we explored potential gender variations in the research content of management papers at both individual and group levels. Our models converged to yield a broadly consistent pattern: women authors and women-dominated author groups typically pose different questions and adopt different perspectives on management than men and male-dominated author groups. Women are, for instance, more likely to engage in social and human-centered areas of management, while men gravitate towards the more technical and operational aspects.

These findings corroborate recent research on sex-segregation demonstrating how cultural norms and expectations are channeling women and men towards different educational and occupational areas. Our results could indicate that similar gendered processes influence the selection of topics in management research.

Increasing GD (at the systemic level) may in this sense contribute to broaden the prevailing conceptions of what constitutes appropriate and desirable management solutions. Whether or not

this is beneficial for the evolvement of the field certainly depends on the eye of the beholder (Pfeffer, 1993; Van Maanen, 1995), but in the large picture, the broadened repertoire of perspectives, values and questions may render management research more responsive to the full gamut of societal needs and expectations (Medin et al., 2017; Nielsen 2018a).

Furthermore, our data allude to possible variations in the prevalence of different methodologies across the gender scale presented in Figure 1. Topics gravitating towards the “male” end of the scale carry a distinct quantitative methodological imprint, while topics on the “female” side encompass a more pluralistic arsenal of methodological approaches. In the future, scholars could consider exploring this methodological “divide” in more detail, with a particular view to its epistemological underpinnings and potential implications for research questions and knowledge claims.

Finally, we investigated whether female researchers and female dominated research groups were more likely to be ‘ghettoized’ into research areas of lower scholarly status. Our analysis linking topic loadings to journal scores did not provide sufficient grounds for making such conclusions. We found high and inconsistent levels of status variance across both typical “female” and “male” research topics, although female-typical research areas were consistently underrepresented among the topics with the highest likelihood of being covered by top journals in fairly established journal lists developed by UT Dallas, Financial Times and the Chartered Association of Business Schools in the UK.

Here, it is important to emphasize that journal prestige represents a somewhat narrow proxy of scholarly status and ghettoization. Other status differentials may play an important part in explaining the gender variations reported in this paper. A global survey of business and management schools accredited by the Association to Advance Collegiate Schools of Business, documents considerable differences in the average salaries of academic scholars across management sub-disciplines. Male dominated areas, such as finance, operations management, entrepreneurship and quantitative methods, top the list, while areas with higher female participation, such as HRM and business communications gravitate towards the bottom (AACSB, 2016). Further, some research topics may be more likely to accrue external funding than others, and this may produce gender-related status differences. These reservations refrain us from rejecting the “ghettoization argument”. In other words, we do not know to what extent women’s orientation towards certain research topics and men’s orientation towards others are moderated by

resource-related status differences. To shed light on these matters, future studies might consider how topic-related variations in researcher salaries and funding opportunities factor into the selection of topics and influence horizontal gender-segregation with respect to research preferences.

An important limitation of the study concerns the inability to account for the socialization aspect of topic selection in academic research. Research demonstrates that academics' socialization through their PhD training may influence research preferences and priorities (in terms of topics, methodologies and theoretical frameworks) (Golde, 2005; Hasrati and Street, 2009; Isaac et al., 1992; Neumann, 2007; Sugimoto, 2009). A closer examination of the potentially gendered aspect of such socialization processes thus represents a promising avenue for further research.

Another limitation concerns the varying coverage of management topics in WoS. As displayed in Table S1.3, several “female-typical” topics appear to have lower-than-average WoS-coverage. If “lower-status” publication outlets that fall outside WoS cover a larger proportion of “female-typical” topics (compared to “male-typical” topics), this could introduce a bias in the analysis of gender differences in scholarly impact and journal status. However, a notable part of the variation in WoS coverage may also be due to a greater emphasis on book-publications in the topics located at the upper half of Table S1.3.

In summary, our findings demonstrate the importance of studying the outcomes of GD according to a broad set of measures. Capturing the associations between GD and research outcomes requires a fine-grained and multifaceted methodology extending beyond traditional metrics of scholarly performance.

Our contribution is a small leap in this direction, but it illustrates that research priorities and outcomes in the management field, at least to some extent, will be determined by how academic organizations populate their faculty positions. Specifically, we have argued that GD, irrespective of whether this is a result of status related stratifications or not, begets cognitive diversity. In an ecological framework, this type of diversity has been argued to be desirable for the evolvment of knowledge production. It can prevent paradigmatic lock-in, accommodate cross-fertilization that leads to new types of knowledge, and make innovation systems more resilient to changing societal demands and expectations (Rafols et al., 2012; Rosenberg, 1994; Stirling, 1998, pp. 6-36; Stirling, 2007). However, a systematic focus on gender, merely captures one color in the gamut of diversity-related characteristics, which also span lines of ethnicity, religion, class, region

and nationality. In the future, important insights may be gained from studying the impact of diversity along these lines as well.

## **Acknowledgements**

Bibliometric indices (CS, NCS, PP top-10%, JS, self-citation rates, institutional collaboration, international collaboration and WoS coverage) were generously provided by the Centre for Science and Technology Studies (CWTS) at Leiden University. We thank Jesper Wiborg Schneider and Sergiy Proshiv for help with data acquisition and processing, and useful comments on the manuscript. We also received valuable feedback from audiences at SCANCOR, Stanford University.

**Competing interests:** We have no competing interests to declare.

## **References**

- Abramo, G., D'Angelo, C. A., Caprasecca, A., 2009. The contribution of star scientists to overall sex differences in research productivity. *Scientometrics* 81(1), 137-156. <https://doi.org/10.1007/s11192-008-2131-7>
- Abramo, G., D'Angelo, C. A., Murgia, G., 2013. Gender Differences in Research Collaboration. *Journal of Informetrics* 7, 811–822. <http://dx.doi.org/10.1016/j.joi.2013.07.002>
- Aksnes, D. W., Rorstad, K., Piro, F., Sivertsen, G., 2011. Are female researchers less cited? A large-scale study of Norwegian scientists. *Journal of the American Society for Information Science and Technology* 62(4), 628–636.
- Alers, M., Leerdam, L.V., Dielissen, P., Lagro-Janssen, A., 2014. Gendered specialties during medical education: a literature review. *Perspectives on Medical Education* 3(3), 163-178. doi: 10.1002/asi.21486
- Allison, P. D., Long, J. S., 1990. Departmental effects on scientific productivity. *American Sociological Review*, 55(4), 469–478. DOI: 10.2307/2095801
- Archambault, É., Larivière, V., 2009. History of the journal impact factor: Contingencies and consequences. *Scientometrics* 79(3), 635-649. DOI: 10.1007/s11192-007-2036-x
- Association to Advance Collegiate Schools of Business (AACSB), 2016. Business data guide 2016. (Downloaded on February 17 2017 from <http://www.aacsb.edu/-/media/aacsb/publications/data-trends-booklet/2016.ashx?la=en>)
- Asuncion, A., Welling, M., Smyth, P., Teh, Y. W., 2009. On smoothing and inference for topic models. *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press Arlington, Virginia, pp. 27-34.
- Baruch, Y., 2001. Global or North American? A geographical based comparative analysis of publications in top management journals. *International Journal of Cross Cultural Management* 1(1), 109-126. <https://doi.org/10.1177%2F1470595801111010>
- Bland, C. J., Center, B. A., Finstad, D. A., Risbey, K. R., Staples, J., 2006. The impact of appointment type on the productivity and commitment of full-time faculty in research and doctoral institutions. *The Journal of Higher Education* 77(1), 89–123.

- Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3, 993-1022.
- Borgatti, S. P., Everett, M. G., Johnson, J. C., 2013. *Analyzing social networks*. SAGE Publications Limited, New York.
- Bornmann, L., Mutz, R., Neuhaus, C., Daniel, H. D., 2008. Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in science and environmental politics* 8(1), 93-102. doi: 10.3354/esep00084
- Borrego, A., Barrios, M., Villarroya, A., Ollé, C., 2010. Scientific output and impact of postdoctoral scientists: a gender perspective. *Scientometrics* 83(1), 93–101. DOI 10.1007/s11192-009-0025-y
- Botelho, T. L., Abraham, M., 2017. Pursuing Quality: How Search Costs and Uncertainty Magnify Gender-based Double Standards in a Multistage Evaluation Process. *Administrative Science Quarterly* 62(4), 698-730. DOI: 10.1177/0001839217694358
- Bourdieu, P., 1990. *The Logic of Practice*, transl. by R. Nice. Polity Press, Cambridge.
- Bozeman, B., Corley, E., 2004. “Scientists’ Collaboration Strategies: Implications for Scientific and Technical Human Capital.” *Research Policy* 33, 599–616.
- Braam, R. R., Moed, H. F., van Raan, A. F. J., 1991a. Mapping of science by combined co-citation and word analysis. I. Structural aspects. *Journal of the American Society for Information Science* 42(4), 233-251.
- Braam, R. R., Moed, H. F., van Raan, A. F. J., 1991b. Mapping of science by combined co-citation and word analysis. II: Dynamical aspects. *Journal of the American Society for Information Science* 42(4), 252-266.
- Brooks, C., Fenton, E. M., & Walker, J. T. (2014). Gender and the evaluation of research. *Research Policy*, 43(6), 990-1001.
- Campbell, L. G., Mehtani, S., Dozier, M. E., Rinehart, J., 2013. Gender-heterogeneous working groups produce higher quality science. *PLOS ONE* 8(10). DOI: 10.1371/journal.pone.0079147
- Caplar, N., Tacchella, S., Birrer, S., 2017. Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy* 1(6), 0141. <https://doi.org/10.1038/s41550-017-0141>
- Cech, E. A., 2013. The Self-Expressive Edge of Occupational Sex Segregation 1. *American Journal of Sociology* 119(3), 747-789. DOI: 10.1086/673969
- Charles, M., Bradley, K., 2002. Equal but separate? A cross-national study of sex segregation in higher education. *American Sociological Review* 67(4), 573-599. DOI: 10.2307/3088946
- Charles, M., Bradley, K., 2009. Indulging our gendered selves? Sex segregation by field of study in 44 countries 1. *American journal of sociology* 114(4), 924-976. DOI: 10.1086/595942
- Charles, M., Grusky, D. B., 2004. *Occupational ghettos: The worldwide segregation of women and men*. Stanford University Press, Stanford.
- Cole J. R., Cole S., 1973. *Social Stratification in Science*. University of Chicago Press, Chicago.
- Cronin, B., 1981. The need for a theory of citing. *Journal of Documentation* 37, 16-24. <https://doi.org/10.1108/eb026703>

CWTS, 2016. CWTS Leiden Ranking: Indicators. (Downloaded on February 27 2016 from <http://www.leidenranking.com/information/indicators>).

Dolado, J., J., Felgueroso, F., Almunia, M., 2012. Are men and women-economists evenly distributed across research fields? Some new empirical evidence. *SERIEs* 3(3), 367-393. DOI 10.1007/s13209-011-0065-4

Duch, J., Zeng, X. H. T., Sales-Pardo, M., Radicchi, F., Otis, S., Woodruff, T. K., et al., 2012. The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PLoS One*, 7(12), e51332. <https://doi.org/10.1371/annotation/7f54a3e6-6dcf-4825-9eb9-201253c1e25>

Eccles, J. S., 2007. Where Are All the Women? Gender Differences in Participation in Physical Science and Engineering. In: Ceci, S. J., Williams, W. M. (Eds.), *Why Aren't More Women in Science? Top Researchers Debate the Evidence*. American Psychological Association, Washington, pp. 199–210.

Elsevier, 2017. Gender in the global research landscape. (Downloaded on April 14 2017 from [https://www.elsevier.com/\\_\\_data/assets/pdf\\_file/0008/265661/ElsevierGenderReport\\_final\\_for-web.pdf](https://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf)).

Ely, R. J., Thomas, D. A., 2001. Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly* 46(2), 229–273.

England, P., Li, S., 2006. Desegregation stalled: The changing gender composition of college majors, 1971–2002. *Gender & Society* 20(5), 657–677.

European Commission, 2012. Communication from the commission to the European Parliament, the Council and The European Economic and Social Committee and the Committee of the Regions: A reinforced European Research Area Partnership for Excellence and Growth. (Downloaded on May 2 2016 from [https://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/era-communication-partnership-excellence-growth\\_en.pdf](https://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-partnership-excellence-growth_en.pdf))

Ferber, M. A., Brün, M., 2011. The gender gap in citations: does it persist? *Feminist Economics*, 17(1), 151–158.

Fernandez, R. M., Friedrich, C. 2011. Gender sorting at the application interface. *Industrial Relations: A Journal of Economy and Society* 50(4), 591-609.

Gender API, 2016. Gender API - Determines the gender of a first name. (Downloaded on January 21 2016 from <https://gender-api.com/>).

Gläser, J., Heinz, M., Havemann, F., 2015. Epistemic Diversity as Distribution of Paper Dissimilarities. (Downloaded on January 21 2016 from <https://pdfs.semanticscholar.org/ec5f/7f25603a4300308c79c5c826245198c760fc.pdf>).

Gläser, J., Laudel, G., 2007. The social construction of bibliometric evaluations. In: Whitley, R., Gläser, J. (Eds.), *The Changing Governance of The Sciences: The Advent of Research Evaluation Systems*. Springer, Doodrecht, pp. 101-123.

Global Research Council, 2016. Statement of Principles and Actions Promoting the Equality and Status of Women in Research. (Downloaded on February 23 2017 from [https://www.globalresearchcouncil.org/fileadmin/documents/GRC\\_Publications/Statement\\_of\\_Principles\\_and\\_Actions\\_Promoting\\_the\\_Equality\\_and\\_Status\\_of\\_Women\\_in\\_Research.pdf](https://www.globalresearchcouncil.org/fileadmin/documents/GRC_Publications/Statement_of_Principles_and_Actions_Promoting_the_Equality_and_Status_of_Women_in_Research.pdf)).

Golde, C. M., 2005. The role of the department and discipline in doctoral student attrition: Lessons from four departments. *The Journal of Higher Education* 76(6), 669-700.

Gonzalez-Brambila, C., Veloso, F. M., 2007. The determinants of research output and impact: A study of Mexican researchers. *Research Policy* 36(7), 1035–1051. <https://doi.org/10.1016/j.respol.2007.03.005>

Grant, L., Ward, K. B., 1991. Gender and publishing in sociology. *Gender & Society* 5(2), 207-223.

Greenacre, M., 2007. Correspondence Analysis in Practice. Chapman & Hall, Florida.

Harrison, D. A., Price, K. H., Bell, M. P., 1998. Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion. *Academy of management journal* 41(1), 96-107. DOI: 10.2307/256901

Hasrati, M., Street, B., 2009. PhD topic arrangement in 'D'iscourse communities of engineers and social sciences/humanities. *Journal of English for Academic Purposes* 8(1), 14-25. <https://doi.org/10.1016/j.jeap.2009.01.002>

Homan, A. C., Van Knippenberg, D., Van Kleef, G. A., De Dreu, C. K., 2007. Bridging faultlines by valuing diversity: diversity beliefs, information elaboration, and performance in diverse work groups. *Journal of Applied Psychology* 92(5), pp. 1189–1199. DOI:10.1037/0021-9010.92.5.1189

Horwitz S. K., 2005. The compositional impact of team diversity on performance: Theoretical consideration. *Human Resource Development Review* 4(2), 219-245. DOI: 10.1177/1534484305275847

Hülsheger, U. R., Anderson, N., Salgado, J. F., 2009. Team-level predictors of innovation at work: A comprehensive meta-analysis spanning three decades of research. *Journal of Applied Psychology* 94(5), 1128–1145. doi: 10.1037/a0015978

Huyer, S., 2015. Is the Gender Gap Narrowing in Science and Engineering? In: UNESCO Science Report: Towards 2030. UNESCO Publishing, Paris. (Downloaded on May 21 2017 from <http://unesdoc.unesco.org/images/0023/002354/235406e.pdf>).

Isaac, P. D., Quinlan, S. V., Walker, M. M., 1992. Faculty perceptions of the doctoral dissertation. *The Journal of Higher Education* 63(3), 241-268. DOI: 10.2307/1982014

Joshi, A., 2014. By whom and when is women's expertise recognized? The interactive effects of gender and education in science and engineering teams. *Administrative Science Quarterly* 59(2), 202-239. DOI: 10.1177/0001839214528331

Judge, T. A., Cable, D. M., Colbert, A. E., Rynes, S. L., 2007. What causes a management article to be cited—article, author, or journal? *Academy of Management Journal* 50(3), 491-506. DOI: 10.5465/AMJ.2007.25525577

Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., Sugimoto, C. R., 2016. Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3), 417-435. DOI: 10.1177/0306312716650046

Larivière, V., Ni, C., Gingras, Y., Cronin, B., Sugimoto, C. R., 2013. Bibliometrics: global gender disparities in science. *Nature* 504(7479), 211–213. doi:10.1038/504211a

Larivière, V., Vignola-Gagné, É., Villeneuve, C., Gelinas, P., Gingras, Y., 2011. Sex differences in research funding, productivity and impact: An analysis of Quebec university professors. *Scientometrics*, 87(3), 483–498. DOI 10.1007/s11192-011-0369-y

Leahey, E., 2006. Gender differences in productivity research specialization as a missing link. *Gender & Society* 20(6), 754-780.

Leišytė, L., 2016. New public management and research productivity—a precarious state of affairs of academic work in the Netherlands. *Studies in Higher Education* 41(5), 828-846. <https://doi.org/10.1080/03075079.2016.1147721>

Lerchenmueller, M. J., Sorenson, O., 2018. The gender gap in early career transitions in the life sciences. *Research Policy* 47(6), 1007-1017. <https://doi.org/10.1016/j.respol.2018.02.009>



- Light, R., 2013. Gender inequality and the structure of occupational Identity: The case of elite sociological publication, in: McDonald, S. (Eds.), *Networks, Work, and Inequality* (Research in the Sociology of Work, volume 24), Emerald Group Publishing Limited, 239-268.
- Long, J. S., 1992. Measures of sex-differences in scientific productivity. *Social Forces*, 71(1), 159–178. DOI: 10.2307/2579971
- Lungeanu, A., Contractor, N. S., 2015. The effects of Diversity and Network Ties on Innovations: The Emergence of a New Scientific Field. *American Behavioral Scientist* 59(5), 548-564. DOI: 10.1177/0002764214556804
- Maes, K., Gvozdanovic, J., Buitendijk, S., Hallberg, I. R., Mantilleri, B., 2012. Women, research and universities: excellence without gender bias. League of European Universities, Leuven. (Downloaded on November 21 2018 from <https://www.leru.org/files/Women-Research-and-Universities-Excellence-without-Gender-Bias-Full-paper.pdf>)
- Mairesse, J., Pezzoni, M., 2015. Does gender affect scientific productivity? *Revue économique*, 66(1), 65-113.
- Mannix, E., Neale, M. A., 2005. What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological science in the public interest* 6(2), 31-55. <https://doi.org/10.1111/j.1529-1006.2005.00022.x>
- McCullagh, P., J. Nelder., 1989. *Generalized Linear Models*. Chapman & Hall, Boca Raton.
- McFarland, D., Ramage, D., Chuang, J., Heer, J., Manning, C., Jurafsky, D., 2013. Differentiating language usage through topic models. *Poetics* 41(6), 607–625. <https://doi.org/10.1016/j.poetic.2013.06.004>
- McLeod, J., 2005. Feminists re-reading Bourdieu: Old debates and new questions about gender habitus and gender change. *Theory and Research in Education* 3(1), 11-30. DOI: 10.1177/1477878505049832
- Medin, D., Ojalehto, B., Marin, A., Bang, M., 2017. Systems of (non-)diversity. *Nature Human Behaviour*, 1, 0088. <https://doi.org/10.1038/s41562-017-0088>
- Merrill-Sands, D., Holvino, E., Cumming, J., 2000. Working with diversity: A framework for action. Gender and Diversity Program of the Consultative Group on International Agricultural Research (CGIAR). (Downloaded on February 14 2016 from [https://library.cgiar.org/bitstream/handle/10947/2726/24\\_Working%20with%20Diversity.%20A%20Framework%20for%20Action\\_genderdiversity\\_WP.pdf?sequence=1](https://library.cgiar.org/bitstream/handle/10947/2726/24_Working%20with%20Diversity.%20A%20Framework%20for%20Action_genderdiversity_WP.pdf?sequence=1))
- Mingers, J., Xu, F., 2010. The drivers of citations in management science journals. *European Journal of Operational Research* 205(2), 422-430. <https://doi.org/10.1016/j.ejor.2009.12.008>
- Mohr, J. W., Bogdanov, P., 2013. Introduction-Topic models: What they are and why they matter. *Poetics* 41(6), 545–569. <https://doi.org/10.1016/j.poetic.2013.10.001>
- Moretti, F., 2005. *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso, London.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., Handelsman, J., 2012. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences* 109(41), 16474-16479. doi: 10.1073/pnas.1211286109
- Neumann, R., 2007. Policy and practice in doctoral education. *Studies in Higher Education* 32(4), 459-473. DOI: 10.1080/03075070701476134
- Nichols, L. G., 2014. A topic model approach to measuring interdisciplinarity at the National Science Foundation. *Scientometrics* 100(3), 741-754. DOI 10.1007/s11192-014-1319-2

- Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics* 122(3), 1067-1101.
- Nielsen, M. W., 2016. Gender inequality and research performance: moving beyond individual-meritocratic explanations of academic advancement. *Studies in Higher Education* 41(11), 2044-2060. DOI: 10.1080/03075079.2015.1007945
- Nielsen, M. W. (2017a). Gender and citation impact in management research. *Journal of Informetrics*, 11(4), 1213-1228.
- Nielsen, M. W. (2017b). Gender consequences of a national performance-based funding model: new pieces in an old puzzle. *Studies in Higher Education*, 42(6), 1033-1055.
- Nielsen, M. W., Bloch, C. W., Schiebinger, L., 2018a. Making gender diversity work for scientific discovery and innovation. *Nature Human Behaviour*, 2, 726-734. <https://doi.org/10.1038/s41562-018-0433-1>
- Nielsen, M. W., 2018b. Scientific performance assessments through a gender lens: A case study on evaluation and selection practices in academia. *Science & Technology Studies*, 31(1), 2-30. DOI: 10.23987/sts.60610
- Nishii, L. H., 2013. The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, 56(6), 1754-1774. DOI: 10.5465/amj.2009.0823
- Page, S. E., 2008. *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton University Press, Princeton.
- Pelled, L. H., Eisenhardt, K. M., Xin, K. R., 1999. Exploring the black box: An analysis of work group diversity, conflict and performance. *Administrative science quarterly* 44(1), 1-28.
- Pfeffer, J., 1993. Barriers to the advance of organizational science: Paradigm development as a dependent variable. *Academy of Management Review* 18 (4), 599-620. DOI: 10.5465/AMR.1993.9402210152
- Podsakoff, P. M., MacKenzie, S. B., Podsakoff, N. P., Bachrach, D. G., 2008. Scholarly influence in the field of management: A bibliometric analysis of the determinants of university and author impact in the management literature in the past quarter century. *Journal of Management*, 34(4), 641-720. <https://doi.org/10.1177/0149206308319533>
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A., 2012. How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy* 41(7), 1262-1282. DOI: 10.1016/j.wasman.2012.01.017
- Ramage, D., Rosen, E., 2011. Stanford topic modeling toolbox. (Downloaded on January 21 2016 from <https://nlp.stanford.edu/software/tmt/tmt-0.4/>).
- Reskin, B. F., Roos, P. A., 2009. *Job queues, gender queues: Explaining women's inroads into male Occupations*. Temple University Press, Philadelphia.
- Ridgeway, C. L., 2011. *Framed by gender: How gender inequality persists in the modern world*. Oxford University Press, New York.
- Rosenberg, N., 1994. Uncertainty and technological change. Conference on growth and development: The economics of the 21st century. Stanford, CA: Stanford University, Center for Economic Policy Research.
- Royal Society, 2017. Diversity in Science. (Downloaded on April 10 2017 from <https://royalsociety.org/topics-policy/diversity-in-science/topic/>).

- Saá-Pérez, D., Díaz-Díaz, N. L., Aguiar-Díaz, I., Ballesteros-Rodríguez, J. L., 2015. How diversity contributes to academic research teams performance. *R&D Management* 47(2), 165-179. <https://doi-org.ez.statsbiblioteket.dk:12048/10.1111/radm.12139>
- Santamaría, L., Mihaljević, H., 2018. Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science*, 4, e156. <https://doi.org/10.7717/peerj-cs.156>
- Sax, L. J., Hagedorn, L. S., Arredondo, M., Dicrisi III, F. A., 2002. Faculty research productivity: Exploring the role of gender and family-related factors. *Research in higher education*, 43(4), 423-446.
- Schneid, M., Isidor, R., Li, C., Kabst, R., 2015. The influence of cultural context on the relationship between gender diversity and team performance: a meta-analysis. *The International Journal of Human Resource Management* 26(6), 733-756. DOI: 10.1080/09585192.2014.957712
- Slyder, J. B., Stein, B. R., Sams, B. S., Walker, D. M., Beale, B. J., Feldhaus, J. J., et al. (2011). Citation pattern and lifespan: a comparison of discipline, institution, and individual. *Scientometrics*, 89(3), 955–966. DOI 10.1007/s11192-011-0467-x
- Stirling, A., 1998. On the economics and analysis of diversity, SPRU Electronic Working Papers, 28. (Downloaded on June 27 2017 from <https://pdfs.semanticscholar.org/08e8/fb40fdbfd91fec63fc632d91b7dbc5779ea1.pdf>).
- Stirling, A., 2007. A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface* 4 (15), 707–719. doi: 10.1098/rsif.2007.0213
- Stirling, A., 2009. Direction, distribution and diversity! pluralising progress in innovation, sustainability and development. Working Paper. STEPS Centre. (Downloaded on June 27 2017 from <http://anewmanifesto.org/wp-content/uploads/stirling-paper-32.pdf>).
- Stvilia, B., Hinnant, C. C., Schindler, K., Worrall, A., Burnett, G., Burnett, K., Kazmer, M. M., Marty, P. F., 2011. Composition of scientific teams and publication productivity at a national science lab. *Journal of the American Society for Information Science and Technology* 62(2), 270-283. <https://doi.org/10.1002/asi.21464>
- Sugimoto, C. R., 2009. Mentoring, collaboration, and interdisciplinarity: An evaluation of the scholarly development of Information and Library Science doctoral students (Doctoral dissertation, University of North Carolina at Chapel Hill).
- Sugimoto, C. R., Ni, C., Larivière, V., 2015. On the relationship between gender disparities in scholarly communication and country-level development indicators. *Science and Public Policy*, 42(6), 789–810. <https://doi.org/10.1093/scipol/scv007>
- Symonds, M. R., Gemmell, N. J., Braisher, T. L., Gorringer, K. L., Elgar, M. A., 2006. Gender differences in publication output: towards an unbiased metric of research performance. *PLoS One*, 1(1), e127. doi: 10.1371/journal.pone.0000127
- Talley, E. M., Newman, D., Mimno, D., Herr II, B. W., Wallach, H. M., Burns, G. A., Leenders, A. G., McCallum, A., 2011. Database of NIH grants using machine-learned categories and graphical clustering. *Nature Methods* 8(6), 443-444. doi: 10.1038/nmeth.1619
- Taylor, S. W., Fender, B. F., Burke, K. G., 2006. Unraveling the academic productivity of economists: The opportunity costs of teaching and service. *Southern Economic Journal*, 72(4), 846–859. DOI: 10.2307/20111856
- Valantine, H. A., Collins, F. S., 2015. National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences* 112(40), 12240-12242. doi: 10.1073/pnas.1515612112

- van Arensbergen, P., van der Weijden, I., Van den Besselaar, P., 2012. Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*, 93(3), 857–868. DOI 10.1007/s11192-012-0712-y
- van Dijk, H., van Engen, M. L., van Knippenberg, D., 2012. Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes* 119(1), 38-53. <https://doi.org/10.1016/j.obhdp.2012.06.003>
- Van Maanen, J., 1995. Style as theory. *Organization science* 6(1), 133-143.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F., 2011. Towards a new crown indicator: Some theoretical considerations. *Journal of informetrics*, 5(1), 37-47.
- Williams, K. Y., O'Reilly, C. A., 1998. Demography and diversity in organizations. In: Staw, B. M., Sutton, R. M., (Eds.), *Research in organizational behavior*. JAI, Stamford, CT, pp. 77-140.
- Xie, Y., Shauman, K. A., 1998. Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review*, 63(6) 847-870. DOI: 10.2307/2657505
- Xie, Y., Shauman, K. A., 2003. *Women in science: Career processes and outcomes*. Harvard University Press, Cambridge (MA).

## TABLES AND FIGURES

**Table 1.** Summary of gender diversity research focusing on academic work-settings, 2006-2015

Author(s)	Field	Nation	N	Focus	Period	Data	Result
Campbell et al. (2013)	Environmental sciences	United States	157 research teams	Team GD and citation-performance	1996-1998, 2005-2007	Citations accrued by 157 gender-diverse and same-gender research groups	Papers authored by both men and women receive more citations than those written by same-gender author groups. The positive effect of GD decreases with the share of women in the author-group.
Joshi (2014)	Multi-disciplinary	United States	60 teams, 550 team members	Team GD and publication/citation performance	Two years	Multiple data sources including survey and performance data (publications, conference proceedings, presentation and citations)	Finds trivial, statistically insignificant effects of women's participation on the performance of research groups.
Lungeanu and Contractor (2015)	Medicine	Global	469 papers, 1,354 scholars	Team GD and the emergence of a new biomedical subfield	2007-2010	Bibliometric data on early medical research papers addressing <i>oncofertility</i> – a disease-term introduced in 2007	Neither GD nor gender similarity are significant factors influencing the emergence of oncofertility as a biomedical subfield.
Saá-Pérez et al. (2015)	Multiple fields	Spain	Approx. 155 research teams	Team GD and publication performance	2006-2009	Bibliometric data on the publication rates of 155 teams with varying levels of GD at a Spanish university	Finds a moderate, positive link between GD and scholarly publication rates in national scientific journals, whereas no discernable diversity effect is found for international journals.
Stvilia et al. (2011)	Physics	United States	1,415 experiments	Team GD and publication performance	2005-2009	Publication data from 1,415 experiments conducted at the National High Magnetic Field Laboratory	Finds a slightly negative, statistically insignificant relationship between increased GD in research teams and scientific publication rates.

**Table 2.** Logistic regression models with PP top-10% as the outcome variable

		MODEL 1			MODEL 2				
		Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
	f_weight	0.949	0.823	1.094	0.072	---			
CONTEXT	Diversity Index	---				1.040	0.829	1.304	0.115
	Univ. Prestige	1.148	1.018	1.294	0.061	1.148	1.018	1.294	0.061
	Business prestige	1.404	1.261	1.564	0.055	1.404	1.261	1.564	0.055
	Arab States	0.556	0.195	1.583	0.534	0.557	0.196	1.585	0.534
	East Asia	2.766	0.976	7.838	0.531	2.762	0.975	7.824	0.531
	Common Wealth St.	0.736	0.101	5.359	1.013	0.738	0.102	5.354	1.011
	Latin America	0.769	0.439	1.349	0.286	0.769	0.439	1.348	0.286
	North America	2.705	2.185	3.350	0.109	2.704	2.184	3.349	0.109
	SW Asia	0.613	0.336	1.119	0.307	0.612	0.336	1.116	0.307
	SCE Europe	0.712	0.442	1.149	0.244	0.713	0.442	1.149	0.244
	Sub-Saharan Africa	0.244	0.071	0.834	0.627	0.244	0.072	0.835	0.627
	Western Europe	1.516	1.226	1.875	0.108	1.515	1.225	1.874	0.108
PUBLICATION BEHAVIOR									
	Self-citations	1.475	1.444	1.507	0.011	1.475	1.444	1.507	0.011
	Collaboration	1.140	1.015	1.279	0.059	1.140	1.015	1.279	0.059
	International Collab.	1.307	1.184	1.442	0.050	1.307	1.184	1.442	0.050
MODEL SUMMARY									
	-2 Log Likelihood	15386.817				15387.221			
	Nagelkerke R <sup>2</sup>	0.155				0.155			
	N	25,297				25,297			

Notes: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables.

**Table 3.** Generalized linear models with CS as the outcome variable

MODEL 3					MODEL 4				
		Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
CONTEXT	Intercept	2.779	2.570	3.005	0.0400	2.763	2.557	2.987	0.0397
	f_weight	0.978	0.926	1.034	0.0283	---			
	Diversity Index	---				0.993	0.912	1.083	0.0439
	Univ. Prestige	1.105	1.055	1.158	0.0239	1.105	1.055	1.159	0.0239
	Business prestige	1.185	1.139	1.234	0.0204	1.185	1.139	1.234	0.0204
	Arab States	0.944	0.798	1.116	0.0854	0.944	0.798	1.116	0.0855
	East Asia	1.259	1.062	1.492	0.0866	1.259	1.063	1.493	0.0867
	Common Wealth St.	0.572	0.344	0.950	0.2592	0.572	0.344	0.952	0.2597
	Latin America	0.793	0.646	0.974	0.1049	0.793	0.645	0.974	0.1050
	North America	1.630	1.523	1.746	0.0349	1.630	1.522	1.745	0.0349
PUBLICATION CHARACTERISTICS	SW Asia	0.652	0.567	0.751	0.0718	0.652	0.567	0.751	0.0718
	SCE Europe	0.870	0.756	1.001	0.0717	0.870	0.756	1.001	0.0718
	Sub-Saharan Africa	0.447	0.374	0.533	0.0899	0.447	0.375	0.533	0.0899
	Western Europe	1.238	1.160	1.322	0.0335	1.239	1.160	1.323	0.0335
	Self-citations	1.256	1.241	1.272	0.0062	1.256	1.241	1.272	0.0061
	Institutional collab.	1.070	1.029	1.112	0.0200	1.070	1.029	1.112	0.0200
	International collab.	1.149	1.104	1.196	0.0205	1.149	1.104	1.196	0.0205
	Pub_year (2007)	0.614	0.542	0.695	0.0636	0.614	0.542	0.696	0.0637
	Pub_year (2008)	0.931	0.877	0.988	0.0304	0.931	0.877	0.988	0.0304
	Pub_year (2009)	0.909	0.860	0.961	0.0285	0.910	0.860	0.962	0.0285
GOODNESS OF FIT	Pub_year (2010)	0.990	0.937	1.047	0.0282	0.990	0.937	1.047	0.0282
	Pub_year (2011)	0.915	0.865	0.968	0.0289	0.915	0.865	0.969	0.0289
	Pub_year (2012)	0.956	0.904	1.011	0.0285	0.956	0.904	1.011	0.0285
	Deviance Value/DF	1.116					1.116		
	χ² Value /DF	1.359					1.359		
OMNIBUS TEST	Log Likelihood	-65210.976					-65211.374		
	Log-likelihood χ²	4062.221					4061.424		
		(P<0.0001)					(P<.0001)		
	DF	21					21		
	N	25,297					25,297		

Notes: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables. 2013 is the reference group for the Pub\_year variables.

**Table A1.** Variable specifications

	Explanation	Measurement type
<b>Outcome variables</b>		
Woman_category	Binary gender variable for first authors (first authors with $f$ -score<0.10=0, first authors with $f$ -score $\geq$ .90= 1).	Binary (0, 1)
JS	Higher values indicate higher average levels of citations accrued by the journal in which a given article has been published. This variable is calculated with a four-year citation window.	Continuous (0.04-31.2)
Log-JS	Log-transformed values for JS LN(JS+1).	Continuous (0.02-1.51)
CS	Raw per-paper citation scores. This variable is calculated with a four-year citation window.	Count (0-198)
PP top-10%	Measures the proportion of papers among the top-10% most frequently cited compared with other papers published in the same field the same year (papers that do not belong to top-10%=0, papers that belong to top-10%=1). This variable is calculated with a four-year citation window.	Binary (0, 1)
Financial times	Binary variable (papers not published in journals used in the Financial Times Research rank= 0, papers published in journals used in the Financial Times Research rank=1). Based on the journal list published in 2016.	Binary (0, 1)
UT Dallas	Binary variable (papers not published in journals used in the UT Dallas Business-school rankings= 0, papers published in journals used in the UT Dallas Business-school rankings=1). Based on the most recent UT Dallas journal list.	Binary (0, 1)
AJG	Binary variable (papers not published in journals of distinction (4*) in the Academic Journal Guide= 0, papers published in journals of distinction (4*) in the Academic Journal Guide). Based on the 2015 ranking.	Binary (0, 1)
<b>Main predictors</b>		
$f$ weight	Values closer to 1 indicate a higher share of women in the author group.	Continuous (0-1)
Diversity Index	Values closer to 1 indicate a more equal share of women and men in the author-group (0= 100% women or 100% men, 1=50% women and men)	Continuous (0-1)
Topic variables	36 topic variables based on loadings from our topic model. Higher values for each topic indicate a higher loading of that topic in a given paper	Continuous (0-1)
<b>Covariates</b>		
University prestige	Measures whether a given author is affiliated with one or more of the top-100 universities in the world according to the 2008 Shanghai ranking (Shanghai Ranking 2008).	Binary (0, 1)
Business prestige	Measures whether a given author is affiliated with one or more of the top-50 universities in the QS World University's subject based ranking of Business and Management Studies for the year 2016 (QS 2016).	Binary (0, 1)
Arab States	Geographical location of the author's institutional affiliation (1=Arab states)	Binary (0, 1)
East Asia	Geographical location of the author's institutional affiliation (1=East Asia)	Binary (0, 1)
Common Wealth St.	Geographical location of the author's institutional affiliation (1=Commonwealth independent states)	Binary (0, 1)
Latin America	Geographical location of the author's institutional affiliation (1=Latin America)	Binary (0, 1)
North America	Geographical location of the author's institutional affiliation (1=North America)	Binary (0, 1)
Oceania	Geographical location of the author's institutional affiliation (1=Oceania)	
SW Asia	Geographical location of the author's institutional affiliation (1=South & West Asia)	Binary (0, 1)
SCE Europe	Geographical location of the author's institutional affiliation (1=Eastern Europe)	Binary (0, 1)
Sub-Saharan Africa	Geographical location of the author's institutional affiliation (1=Sub-Saharan Africa)	Binary (0, 1)
Western Europe	Geographical location of the author's institutional affiliation (1=Western Europe)	Binary (0, 1)
Self-citations	Measures the self-citation frequency of a given paper	Count (0-29)
Institutional Collaboration	Measures whether a given paper involves collaboration between authors from two or more institutions (0=no collaboration, 1=collaboration).	Binary (0, 1)
International Collab.	Measures whether a given paper involves collaboration between authors from two or more institutions in different countries (0=no collaboration, 1=collaboration).	Binary (0, 1)



**Table A2.** Geographical groupings

**Arab States**

Algeria  
Egypt  
Jordan  
Kuwait  
Lebanon  
Oman  
Qatar  
Saudi Arabia  
Syria  
Tunisia  
United Arab Emirates  
Morocco

**East Asia**

Brunei  
Cambodia  
China  
Indonesia  
Japan  
Malaysia  
Myanmar  
Philippines  
South Korea  
Singapore  
Mongol Peoples Republic  
Taiwan  
Thailand  
Vietnam

**Commonwealth Independent States**

Kazakhstan  
Kyrgyzstan  
Republic of Georgia  
Russia  
Tajikistan  
Uzbekistan

**Latin America**

Argentina  
Bahamas  
Belize  
Bolivia  
Brazil  
Chile  
Columbia  
Costa Rica  
Cuba  
Dominican Republic  
Ecuador  
El Salvador  
French Guyana  
Guyana

Haiti  
Honduras  
Jamaica  
Mexico  
Nicaragua  
Panama  
Paraguay  
Peru  
St. Lucia  
Trinidad & Tobago  
Uruguay  
Venezuela

**North America**

United States of America  
Canada

**Oceania**

Australia  
Micronesia  
Fiji Islands  
New Caledonia  
New Zealand  
Solomon Islands

**South and West Asia**

Afghanistan  
Bangladesh  
Bhutan  
India  
Iran  
Nepal  
Pakistan  
Sri Lanka

**South-Central and Eastern Europe**

Albania  
Bosnia Hercegovina  
Bulgaria  
Croatia  
Czech Republic  
Estonia  
Hungary  
Latvia  
Lithuania  
Macedonia  
Poland  
Romania  
Serbia  
Slovakia  
Slovenia  
Turkey  
Ukraine

**Sub-Saharan Africa**

Benin  
Botswana  
Burkina Faso  
Burundi  
Cameroon  
Congo  
Eritrea  
Ethiopia  
Ghana  
Ivory Coast  
Kenya  
Lesotho  
Malawi  
Mali  
Mauritius  
Mozambique  
Namibia  
Niger  
Reunion  
Rwanda  
Senegal  
Seychelles  
Sierra Leone  
South Africa  
Swaziland  
Togo  
Tanzania  
Uganda  
Zambia  
Zimbabwe

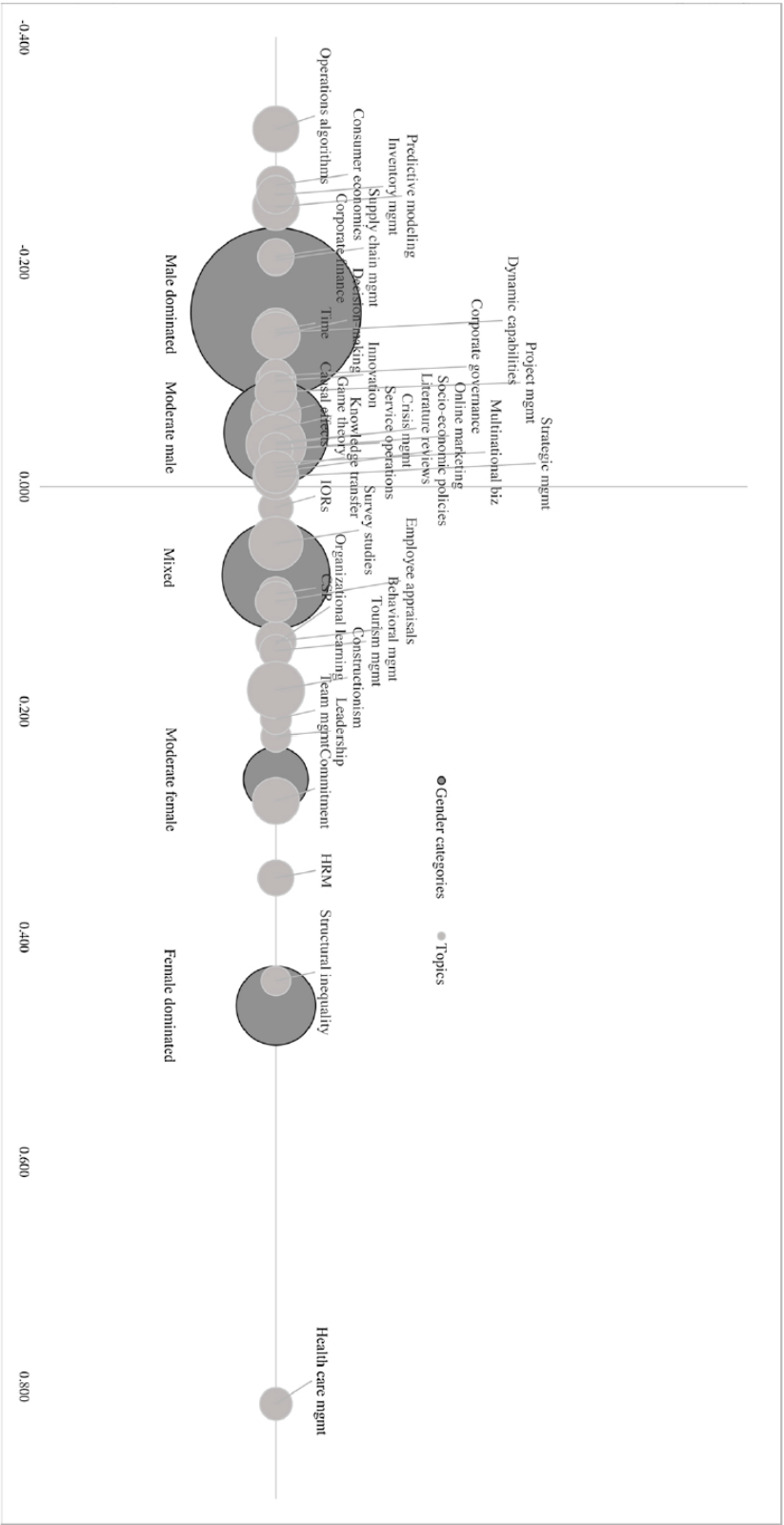
**Western Europe**

Austria  
Belgium  
Cyprus  
Denmark  
Finland  
France  
Germany  
Great Britain  
Greece  
Iceland  
Ireland  
Israel  
Italy  
Luxembourg  
Malta  
Monaco  
Netherlands  
Norway  
Portugal  
Spain  
Sweden  
Switzerland

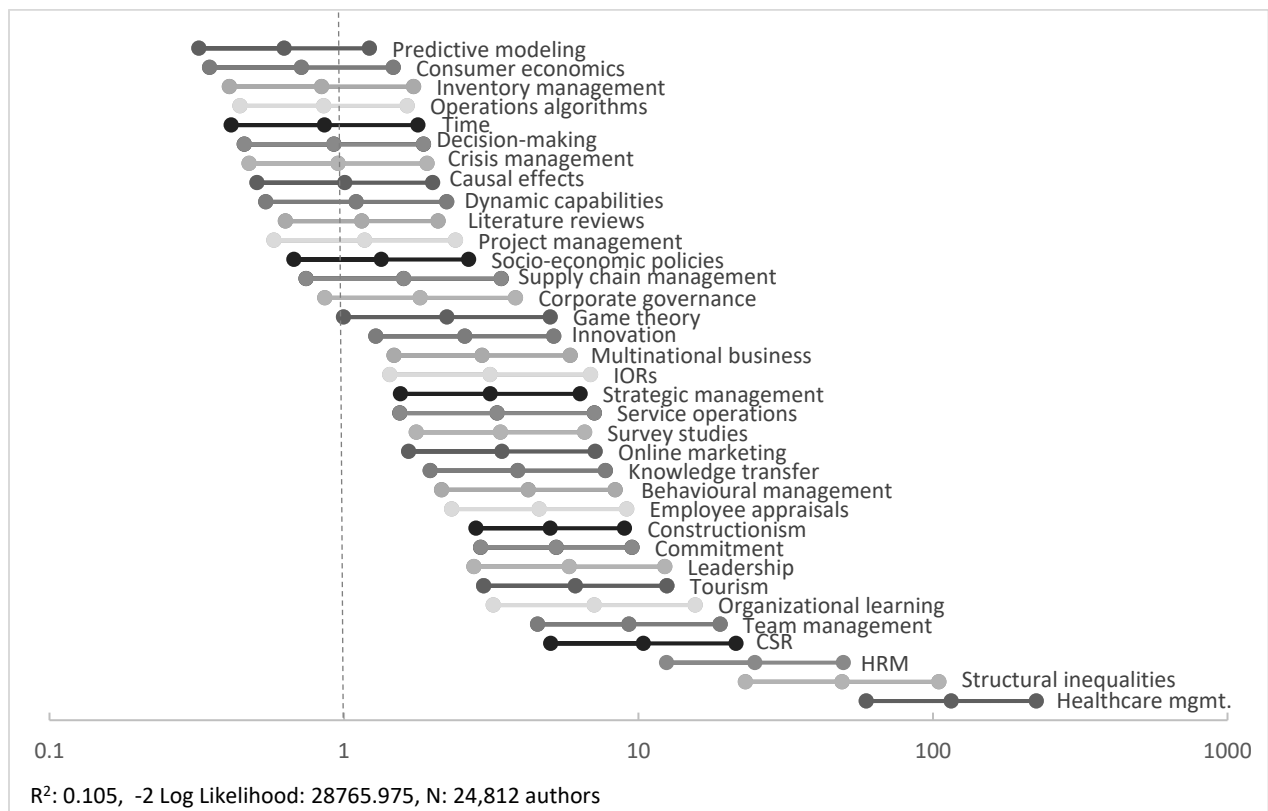
**Table A3.** Top-5 words for each topic in the topic model

Innovation	Dynamic capabilities	Organizational learning	Time	IORs	Strategic management
Innov	firm	organiz	growth	inform	manag
Product	capabl	organ	time	collabor	busi
technolog	competit	learn	period	trust	strategi
develop	industri	cultur	chang	share	compani
market	resourc	creativ	effect	relationship	strateg
Leadership	Commitment	Decision-making	HRM	Survey studies	Multinational business
leadership	employe	model	work	factor	intern
leader	job	decis	human	differ	global
student	work	process	resourc	data	chines
educ	relationship	approach	career	purpos	china
ethic	commit	propos	employ	analysi	entrepreneuri
Causal effects	Game theory	Structural inequality	Behavioral management	Knowledge transfer	Corporate governance
perform	contract	group	behavior	knowledg	govern
relationship	motiv	women	person	network	control
effect	power	gender	individu	transfer	corpor
posit	negoti	differ	emot	patent	board
impact	cooper	status	influenc	social	firm
Project management	Consumer economics	Operations algorithms	Crisis management	Service operations	Predictive modeling
project	price	problem	challeng	servic	measur
system	market	optim	failur	custom	model
design	product	solut	peopl	qualiti	forecast
develop	retail	effici	crisi	satisfact	estim
implement	profit	function	need	provid	test
Literature reviews	Inventory management	CSR	Tourism management	Online marketing	Supply chain management
literatur	cost	environment	organis	user	suppli
theori	time	sustain	communiti	onlin	chain
review	demand	social	behaviour	consum	supplier
develop	capac	respons	tourism	inform	manufactur
discuss	polic	stakehold	event	brand	logist
Socio-economic policies	Team management	Employee appraisals	Constructionism	Healthcare management	Corporate finance
polic	team	effect	chang	Nurs	risk
public	member	particip	process	care	valu
countri	conflict	percept	social	health	invest
sector	task	experi	practic	manag	financi
region	divers	perceiv	organ	hospit	market

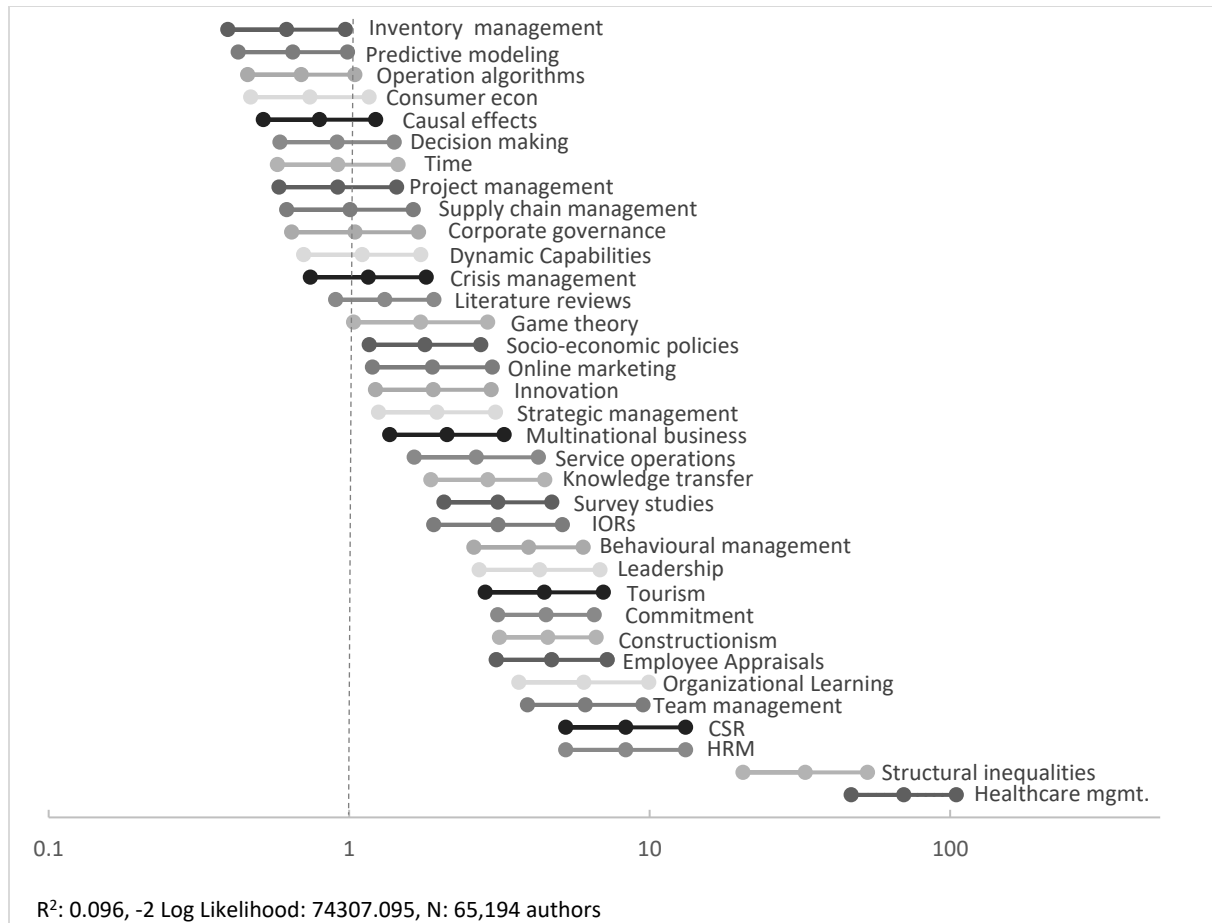
Figure 1. Topic dimension of management studies (Correspondence analysis).



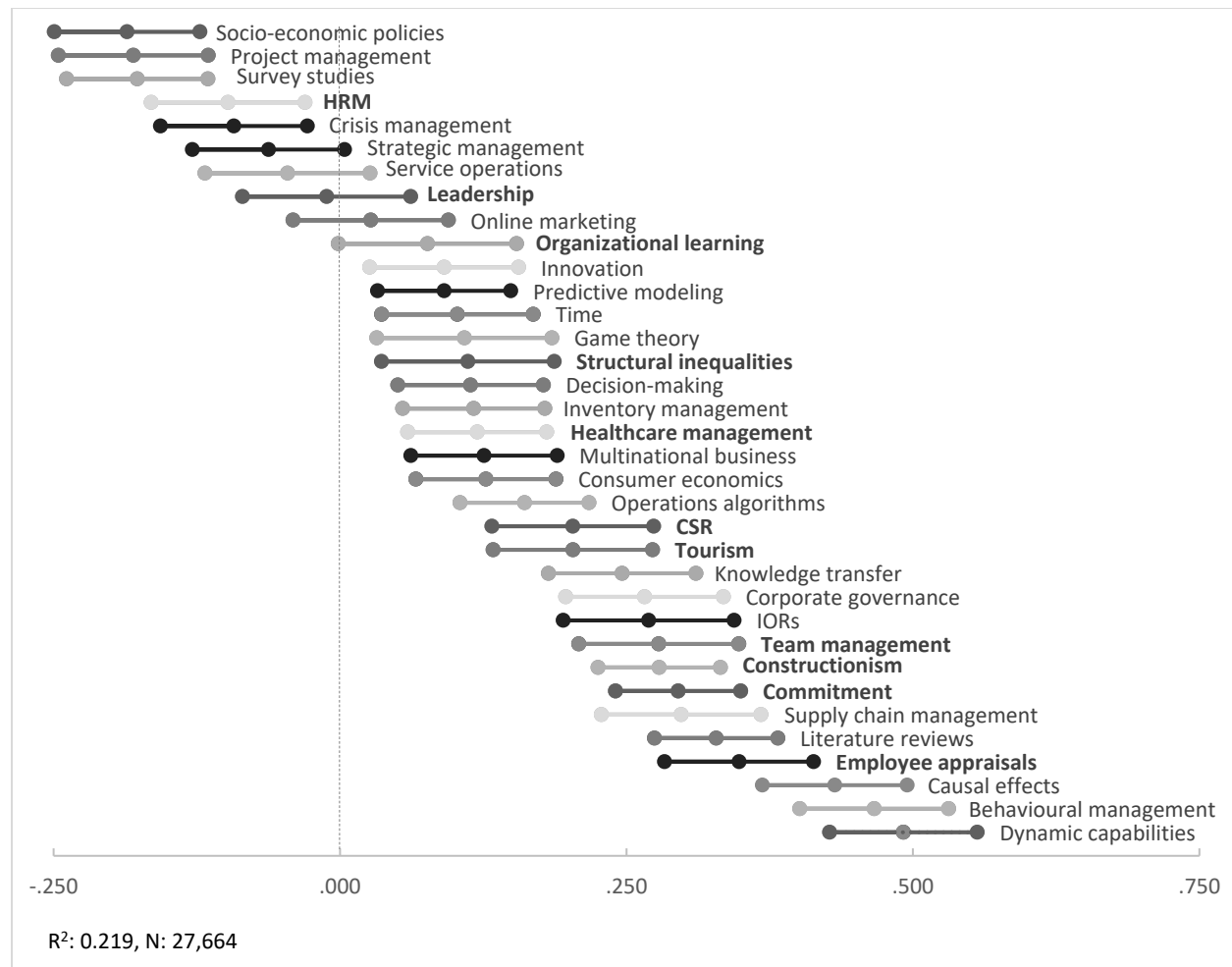
**Figure 2.** Odds ratios and 95% confidence intervals for binary logistic regression with first-author gender as the outcome variable



**Figure 3.** Odds ratios and 95% confidence intervals for the binary logistic regression with author gender as the outcome variable



**Figure 4.** Unstandardized beta-coefficients and 95% confidence intervals for OLS with Log-Journal Scores as outcome variable (log-JS)



## SUPPLEMENTARY MATERIAL: APPENDICES S1-S5

Nielsen, M. W., Börjeson, L. "Gender diversity in the management field: Does it matter for research outcomes?"

### Contents

APPENDIX S1: RESULTS.....	2
APPENDIX S2: TOPIC MODELING PARAMETERS.....	29
APPENDIX S3: TOPIC SUMMARY.....	31
APPENDIX S4: ROBUSTNESS AND QUALITY ESTIMATES FOR THE CA.....	40
APPENDIX S5: INSPECTING THE VALIDITY OF THE TOPIC MODEL.....	43

## APPENDIX S1: RESULTS

### Gender API Validation study

Our validation of the gender assignment procedure was carried out in two steps. First, we tested for potential sampling bias with respect to gender composition in our final sample. Second, we assessed the accuracy of the Gender API in inferring author gender.

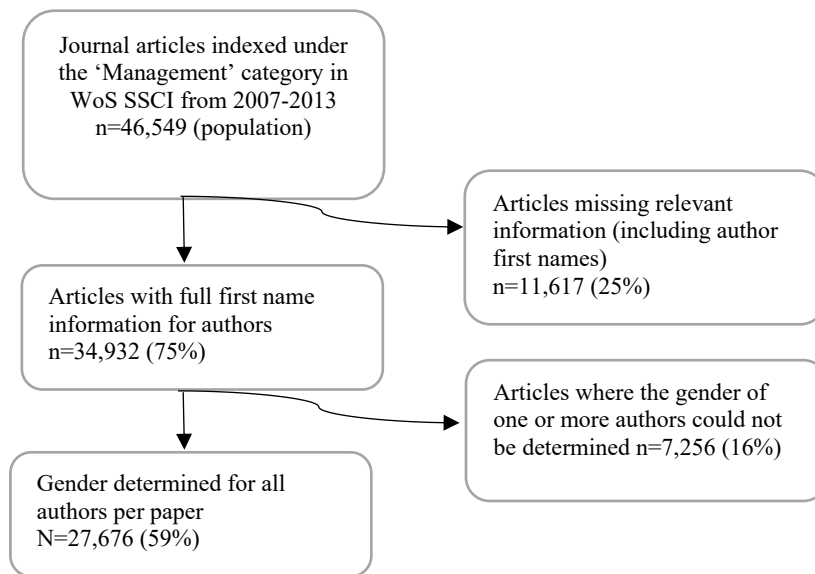
1. We used a method similar to Gonzalez-Alvarez and Cervera-Crespo (2017a; 2017b) (see also Nielsen 2017 that builds on the same validation sample) to test for potential sampling bias (i.e. whether the gender composition of our final sample accurately reflects the gender composition of the original dataset). We selected a random validation sample of 382 papers (including 939 authorships) from the original dataset of 46,549 papers (including 114,096 authorships). We used a sample-size calculator (95% confidence interval, 5% margin of error) to compute the appropriate size of the validation sample. The validation sample was subjected to manual examination based on photos, resumes, biographical descriptions and other relevant information. Based on a comprehensive web-search, we were able to infer the gender of 887 (95%) of the 939 authorships in the validation sample. Of these, 242 (27.3%) were women, and 645 (72.7%) were men. We used a 2X2 contingency table (calculated with Chi-square with Yates correction) to compare the gender composition of the validation sample with the final sample of authorships included in the analysis (the final sample consists of 21,354 (29.9%) women and 49,968 (70.1%) men). The contingency table indicated no statistically significant difference between the gender composition of the validation sample and the final data set ( $\chi^2(1) = 2.826$ ,  $P = .0928$ ).
2. The accuracy of Gender API was examined by comparing its predictions with the outcomes of the manual examination. We were able to match 545 (58%) authorships from the validation sample to data in our final sample of 71,322 authorships. Of these, 391 authorships were classified as male by Gender API and 381 authorships were found to be male in the manual web-search (false positive rate = 2.6%). Gender API classified 154 authorships as female of which 135 were found to be female in the manual search (false positive rate = 5.8%).

In Figures 2 and 3 (in the manuscript), and Tables S1.14 and S1.15, we used the gender information on each authorship in the dataset. In this part of the analysis, we have excluded all authorships with Gender API scores  $\geq 90$  from the analysis. 496 authorships from the validation sample were matched to data in the dataset restricted to authorships with Gender API scores  $\geq 90$ . Of these, Gender API classified 364 as male and 132 as female. Of the 364 authorships classified as male, 363 were found to be male and 1 was found to be female in the manual web-search (false positive rate = 0.27%). Of the 132 authorships classified as female by Gender API, 131 were found to be female and 1 was found to be male in the manual search (false positive rate = 0.76%).

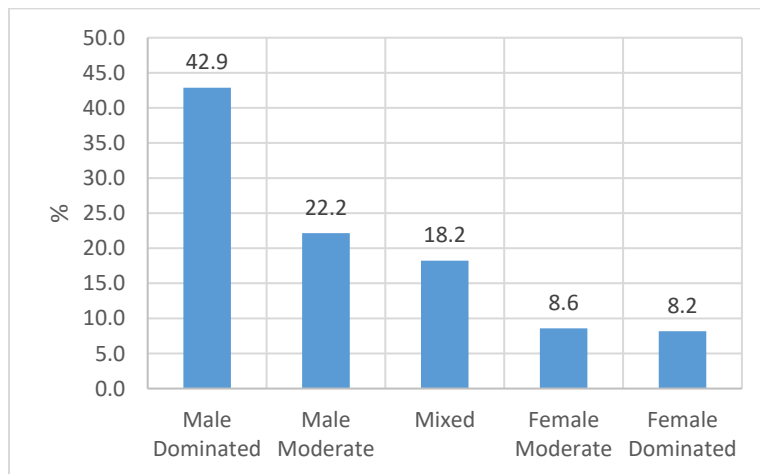


## FIGURES

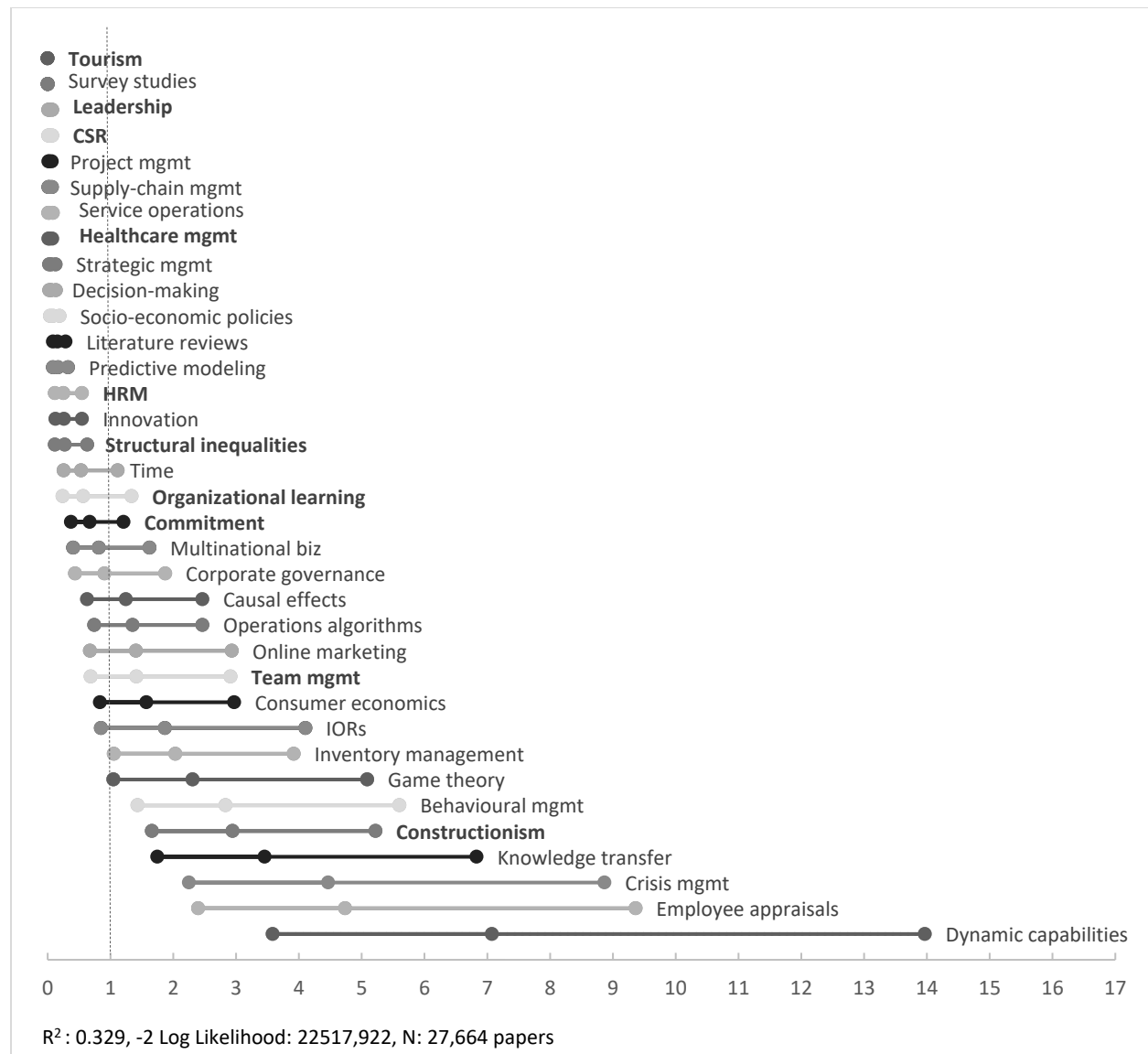
**Figure S1.1.** Flowchart of data inclusion and exclusion



**Figure S1.2.** Distribution across gender categories

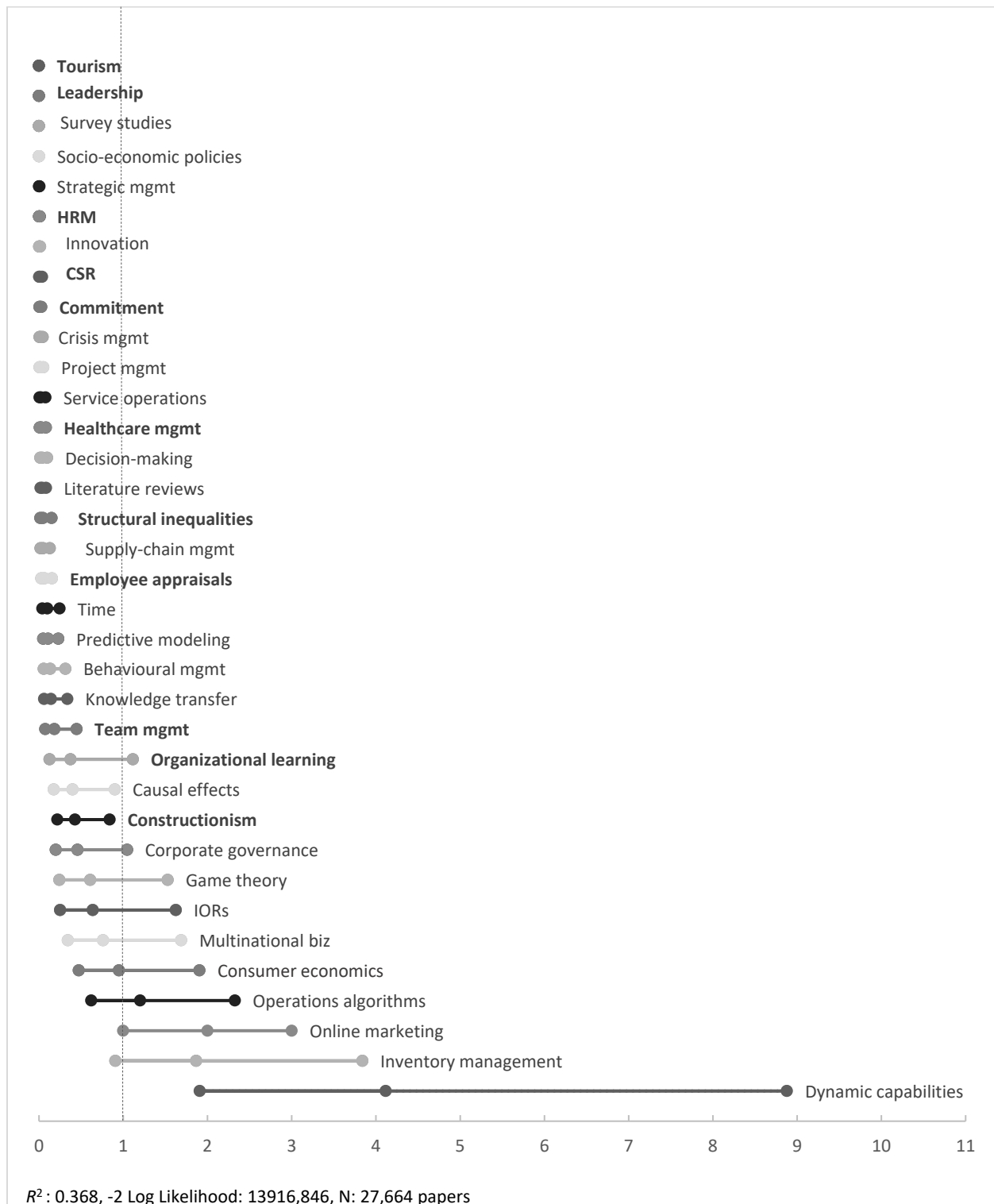


**Figure S1.3.** Odds ratios and 95% confidence intervals for the logistic regression with ‘Financial times’ as the outcome variable (1=journal included in the FT journal list)



Notes: *Oceania* is the reference group for the geographical variables. *Corporate finance* is the reference group for the topic variables. Covariates in the model: geographical groupings, university prestige, business prestige, f\_weight, cross-institutional collaboration, international collaboration.

**Figure S1.4.** Odds ratios and 95% confidence intervals for the logistic regression with ‘UT Dallas’ as the outcome variable (1=journal included in the UT Dallas journal list)



Notes: *Oceania* is the reference group for the geographical variables. *Corporate finance* is the reference group for the topic variables. Covariates in the model: geographical groupings, university prestige, business prestige, f\_weight, cross-institutional collaboration, international collaboration.

**Figure S1.5.** Odds ratios and 95% confidence intervals for the logistic regression with AJG top journals (4\*) as the outcome variable (1=journal included in the AJG top journal list)



Notes: *Oceania* is the reference group for the geographical variables. *Corporate finance* is the reference group for the topic variables. Covariates in the model: geographical groupings, university prestige, business prestige, f\_weight, cross-institutional collaboration, international collaboration.

**Table S.1.1.** Average WoS coverage per diversity group

Diversity group	WoS Coverage
Male dominated	0.549
Moderate male	0.550
Mixed	0.549
Moderate female	0.556
Female dominated	0.548

Note: WoS coverage is calculated as the mean percentage of references cited by WoS articles to other articles covered by WoS, per diversity group. A WoS-coverage score of 0.5 specifies that 50% percent of the cited references in papers pertaining to a given diversity group go to other papers covered by WoS.

**Table S.1.2.** Average WoS coverage per gender group

Gender	WoS Coverage
Male	0.558
Female	0.539

Note: WoS coverage is calculated as the mean percentage of references cited by WoS articles to other articles covered by WoS, per gender group. A WoS-coverage score of 0.5 specifies that 50% percent of the cited references in papers pertaining to a given gender group go to other papers covered by WoS.

**Table S.1.3.** Bivariate correlation coefficients for topics and WoS coverage

Topic	Pearson's r	Kendall's t
Socio-economic policies	-0.269	-0.295
<b>Tourism</b>	<b>-0.193</b>	<b>-0.253</b>
Crisis mgmt	-0.191	-0.203
Project mgmt	-0.181	-0.207
<b>Constructionism</b>	<b>-0.156</b>	<b>-0.176</b>
Strategic mgmt	-0.149	-0.204
<b>CSR</b>	<b>-0.136</b>	<b>-0.191</b>
Survey studies	-0.127	-0.104
<b>Healthcare mgmt</b>	<b>-0.087</b>	<b>-0.217</b>
Literature reviews	-0.073	-0.147
<b>Leadership</b>	-0.070	-0.147
<b>HRM</b>	<b>-0.061</b>	<b>-0.120</b>
Decision-making	-0.049	-0.066
Online marketing	-0.033	-0.089
Service operations	-0.026	-0.071
<b>Structural inequalities</b>	<b>-0.014</b>	<b>-0.059</b>
<b>Organizational learning</b>	<b>-0.011</b>	<b>-0.054</b>
Supply-chain mgmt	-0.009	-0.060
Time	0.009	0.007
Inventory mgmt	0.027	-0.029
Corporate finance		
Multinational biz	0.029	-0.052
Corporate governance	0.040	-0.020
Innovation	0.044	0.006
Knowledge transfer	0.058	-0.022
Game theory	0.071	0.053
Predictive modeling	0.073	0.069
IORs	0.079	0.053
Consumer economics	0.091	0.075
Operations algorithms	0.098	0.010
<b>Team mgmt</b>	<b>0.121</b>	<b>0.086</b>

<b>Employee appraisals</b>	<b>0.151</b>	<b>0.097</b>
<b>Commitment</b>	<b>0.171</b>	<b>0.109</b>
Behavioural mgmt	0.178	0.151
Dynamic capabilities	0.196	0.151
Causal effects	0.242	0.258

Note: The topics ordered from highest to lowest Pearson's r correlation. To visualize whether women are more likely to engage in areas of management research with low WoS coverage, topics located at the right side of Figure 1 and at the bottom of Figure 2 have been marked in bold. The correlations are calculated based on the full sample. WoS coverage is calculated as the mean percentage of references cited by a WoS article to other articles covered by WoS. A WoS-coverage score of 0.5 specifies that 50% percent of the cited references in a paper go to other papers covered by WoS.

**Table S.1.4.** Paper frequency and average reference coverage per Web of Science Subject Category

<b>Web of Science - Subject Category</b>	<b>Frequency</b>	<b>Percent</b>	<b>Average WoS Coverage</b>
Arts & Humanities - Other Topics	43	0.1	0.3144
Behavioral Sciences	72	0.2	0.4889
Business & Economics	27676	63.0	0.5499
Communication	140	0.3	0.5373
Computer Science	691	1.6	0.6246
Education & Educational Research	80	0.2	0.5200
Engineering	1604	3.7	0.5572
Environmental Sciences & Ecology	824	1.9	0.3911
Government & Law	76	0.2	0.5953
Information Science & Library Science	1443	3.3	0.5621
Mathematical Methods In Social Sciences	131	0.3	0.5402
Mathematics	219	0.5	0.5653
Nursing	475	1.1	0.4463
Operations Research & Management Science	3488	7.9	0.6299
Psychology	3426	7.8	0.6638
Public Administration	1127	2.6	0.4907
Public, Environmental & Occupational Health	112	0.3	0.2795
Science & Technology - Other Topics	214	0.5	0.5032
Social Sciences - Other Topics	1651	3.8	0.4287
Sociology	113	0.3	0.4031
Sport Sciences	150	0.3	0.4257
Transportation	61	0.1	0.3510
Women's Studies	93	0.2	0.3526
Total	43909	100.0	

Note: Most papers have several subject categories assigned to them. The numerical values are based on a full, not fractionalized, counting of WoS subject categories assigned to papers. The results are calculated based on the full sample. WoS coverage is calculated as the mean percentage of references cited by WoS articles to other articles covered by WoS. For instance, a per-paper WoS-coverage score of .7 specifies that 70% percent of the cited references in a paper go to other papers covered by WoS.

**Table S1.5.** Linear regression model with log-transformed citation scores (log-CS) as the outcome variable

		<b>B</b>	<b>CI (95%)</b>		<b>S.E.</b>	<b>B</b>	<b>CI (95%)</b>		<b>S.E.</b>
	Constant	0.967	0.916	1.017	0.026	0.961	0.911	1.011	0.025
	f_weight	-0.021	-0.056	0.014	0.018	---			
	Diversity Index	---				-0.005	-0.061	0.052	0.029
<b>CONTEXT</b>									
	Univ. Prestige	0.089	0.056	0.122	0.017	0.089	0.056	0.122	0.017
	Business prestige	0.168	0.140	0.197	0.015	0.168	0.140	0.197	0.015
	Arab States	0.027	-0.124	0.178	0.077	0.027	-0.124	0.178	0.077
	East Asia	0.082	-0.069	0.234	0.077	0.083	-0.069	0.234	0.077
	Common Wealth St.	-0.240	-0.597	0.117	0.182	-0.240	-0.598	0.117	0.182
	Latin America	-0.210	-0.303	-0.117	0.047	-0.210	-0.303	-0.118	0.047
	North America	0.341	0.297	0.385	0.023	0.341	0.297	0.385	0.023
	SW Asia	-0.260	-0.346	-0.173	0.044	-0.260	-0.347	-0.173	0.044
	SCE Europe	-0.122	-0.199	-0.045	0.039	-0.122	-0.199	-0.045	0.039
	Sub-Saharan Africa	-0.412	-0.522	-0.303	0.056	-0.412	-0.522	-0.302	0.056
	Western Europe	0.166	0.123	0.209	0.022	0.166	0.123	0.209	0.022
<b>PUBLICATION CHARACTERISTICS</b>									
	Self-citations	0.179	0.172	0.185	0.003	0.179	0.172	0.185	0.003
	Institutional collab.	0.084	0.058	0.111	0.014	0.084	0.058	0.111	0.014
	International collab.	0.090	0.064	0.115	0.013	0.090	0.064	0.115	0.013
	Pub_year (2007)	-0.337	-0.424	-0.249	0.045	-0.336	-0.423	-0.248	0.045
	Pub_year (2008)	-0.062	-0.101	-0.022	0.020	-0.061	-0.101	-0.022	0.020
	Pub_year (2009)	-0.047	-0.084	-0.010	0.019	-0.046	-0.083	-0.009	0.019
	Pub_year (2010)	-0.007	-0.043	0.028	0.018	-0.007	-0.042	0.028	0.018
	Pub_year (2011)	-0.113	-0.147	-0.079	0.017	-0.113	-0.147	-0.079	0.017
	Pub_year (2012)	-0.029	-0.064	0.005	0.018	-0.029	-0.064	0.006	0.018
<b>MODEL SUMMARY</b>									
	R <sup>2</sup>	0.169					1.169		
	Adjusted R <sup>2</sup>	0.168					1.168		
	N	25,297					25,297		

Note: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables. B refers to the unstandardized beta-coefficients. 2013 is the reference group for the Pub\_year variables. CS has been log transformed by calculating the natural logarithm of CS+1.

**Table S1.6.** Generalized linear models with field-normalized citation scores as the outcome variable (NCS integer)

		Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
	Intercept	606.901	566.178	650.554	0.0354	604.546	563.898	648.124	0.0355
	f_weight	0.971	0.919	1.027	0.0284	---			
	Diversity Index	---				0.975	0.893	1.063	0.0444
<b>CONTEXT</b>	Univ. Prestige	1.093	1.043	1.145	0.0238	1.093	1.043	1.146	0.0238
	Business prestige	1.177	1.131	1.225	0.0203	1.177	1.131	1.225	0.0203
	Arab States	0.978	0.820	1.168	0.0902	0.978	0.819	1.167	0.0903
	East Asia	1.232	1.030	1.473	0.0912	1.234	1.032	1.475	0.0912
	Common Wealth St.	0.637	0.373	1.089	0.2736	0.638	0.373	1.092	0.2741
	Latin America	0.810	0.662	0.991	0.1028	0.809	0.661	0.990	0.1029
	North America	1.623	1.513	1.741	0.0358	1.622	1.512	1.741	0.0359
	SW Asia	0.661	0.572	0.763	0.0735	0.661	0.572	0.763	0.0735
	SCE Europe	0.871	0.755	1.004	0.0726	0.871	0.755	1.004	0.0727
	Sub-Saharan Africa	0.431	0.361	0.514	0.0904	0.431	0.361	0.515	0.0903
	Western Europe	1.239	1.159	1.325	0.0343	1.240	1.159	1.326	0.0343
<b>PUBLICATION CHARACTERISTICS</b>									
	Self-citations	1.257	1.242	1.272	0.0062	1.257	1.242	1.272	0.0062
	Institutional collab.	1.067	1.026	1.110	0.0200	1.067	1.026	1.110	0.0200
	International collab.	1.149	1.103	1.197	0.0209	1.148	1.102	1.196	0.0209
<b>GOODNESS OF FIT</b>									
	Deviance Value/DF	1.250					1.250		
	$\chi^2$ Value /DF	0.803					0.803		
	Log Likelihood	-191163.164					-191163.462		
<b>OMNIBUS TEST</b>									
	Log-likelihood $\chi^2$	2308.225 (P<0.0001)					2307.630 (P<0.0001)		
	DF	15					15		
	N	25,297					25,297		

Notes: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables. The field-normalized citation scores have been transformed to integer values (NCS integer). Specifically, we have multiplied each NCS score with 1000 and rounded to nearest integer.



**Table S1.7.** Logistic regression model with PP top-10% as the outcome variable (quintile-based categorical diversity variables)

		<b>Exp(B)</b>	<b>CI (95%)</b>		<b>S.E.</b>
	Moderate male	0.974	0.873	1.087	0.056
	Mixed	1.010	0.899	1.135	0.059
	Moderate female	0.989	0.846	1.156	0.080
	Female dominated	0.920	0.780	1.086	0.084
<b>CONTEXT</b>					
	Univ. Prestige	1.147	1.018	1.294	0.061
	Business prestige	1.404	1.261	1.563	0.055
	Arab States	0.556	0.195	1.584	0.534
	East Asia	2.765	0.976	7.835	0.531
	Common Wealth St.	0.735	0.101	5.348	1.013
	Latin America	0.769	0.439	1.349	0.286
	North America	2.705	2.184	3.349	0.109
	SW Asia	0.613	0.336	1.118	0.306
	SCE Europe	0.713	0.442	1.150	0.244
	Sub-Saharan Africa	0.244	0.071	0.833	0.627
	Western Europe	1.515	1.225	1.874	0.108
<b>PUBLICATION CHARACTERISTICS</b>					
	Self-citations	1.475	1.444	1.507	0.011
	Collaboration	1.140	1.015	1.279	0.059
	International Collab.	1.306	1.184	1.442	0.050
<b>MODEL SUMMARY</b>					
	-2 Log Likelihood	15386.054			
	Nagelkerke R <sup>2</sup>	0.155			
	N	25,297			

Note: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables. *Male dominated* is the reference group for the categorical diversity variables.

**Table S1.8.** Generalized linear model with citation scores as the outcome variable (quintile-based categorical diversity variables)

		Exp(B)	CI (95%)		S.E.
	Intercept	2.772	2.563	2.997	0.0400
	Moderate male	1.008	0.966	1.051	0.0215
	Mixed	0.983	0.941	1.027	0.0224
	Moderate female	0.988	0.931	1.049	0.0306
	Female dominated	0.982	0.919	1.050	0.0339
<b>CONTEXT</b>					
	Univ. Prestige	1.106	1.055	1.159	0.0239
	Business prestige	1.185	1.139	1.234	0.0204
	Arab States	0.944	0.798	1.116	0.0855
	East Asia	1.260	1.063	1.493	0.0867
	Common Wealth St.	0.572	0.344	0.950	0.2590
	Latin America	0.793	0.646	0.974	0.1049
	North America	1.630	1.522	1.745	0.0349
	SW Asia	0.652	0.567	0.751	0.0718
	SCE Europe	0.870	0.755	1.001	0.0719
	Sub-Saharan Africa	0.447	0.375	0.533	0.0899
	Western Europe	1.238	1.160	1.322	0.0335
<b>PUB CHARACTERISTICS</b>					
	Self-citations	1.256	1.241	1.271	0.0062
	Collaboration	1.070	1.029	1.112	0.0199
	International Collab.	1.150	1.104	1.197	0.0205
	Pub_year (2007)	0.614	0.542	0.695	0.0636
	Pub_year (2008)	0.931	0.877	0.988	0.0304
	Pub_year (2009)	0.909	0.859	0.961	0.0285
	Pub_year (2010)	0.990	0.937	1.046	0.0282
	Pub_year (2011)	0.915	0.864	0.968	0.0290
	Pub_year (2012)	0.956	0.904	1.011	0.0285
<b>GOODNESS OF FIT</b>					
	Deviance Value/DF	1.116			
	$\chi^2$ Value /DF	1.359			
	Log Likelihood	-65210.480			
<b>OMNIBUS TEST</b>					
	Log-likelihood $\chi^2$	4063.211			
	DF	24			
	N	25,297			

Note: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables. *Male dominated* is the reference group for the categorical diversity variables.

**Table S1.9.** Logistic regression model with PP top-10% as the outcome variable (quadratic terms)

		Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
	f_weight	1.103	0.734	1.657	0.208				
	Diversity Index	---				1.042	0.437	2.484	0.443
	f_weight (quadratic)	0.834	0.526	1.323	0.235	---			
	Diversity (quadratic)	---				0.996	0.155	6.382	0.948
<b>CONTEXT</b>	Univ. Prestige	1.148	1.018	1.294	0.061	1.148	1.018	1.294	0.061
	Business prestige	1.405	1.261	1.564	0.055	1.404	1.261	1.564	0.055
	Arab States	0.557	0.196	1.587	0.534	0.557	0.196	1.585	0.534
	East Asia	2.758	0.973	7.813	0.531	2.762	0.975	7.824	0.531
	Common Wealth St.	0.739	0.102	5.367	1.012	0.738	0.102	5.354	1.011
	Latin America	0.770	0.439	1.349	0.286	0.769	0.439	1.348	0.286
	North America	2.704	2.184	3.349	0.109	2.704	2.184	3.348	0.109
	SW Asia	0.612	0.336	1.117	0.307	0.612	0.336	1.116	0.307
	SCE Europe	0.712	0.442	1.149	0.244	0.713	0.442	1.149	0.244
	Sub-Saharan Africa	0.244	0.071	0.833	0.627	0.244	0.072	0.835	0.627
	Western Europe	1.515	1.225	1.874	0.108	1.515	1.225	1.874	0.108
<b>PUBLICATION CHARACTERISTICS</b>									
	Self-citations	1.475	1.443	1.507	0.011	1.475	1.444	1.507	0.011
	Collaboration	1.140	1.016	1.279	0.059	1.140	1.015	1.279	0.059
	International Collab.	1.307	1.184	1.442	0.050	1.307	1.184	1.442	0.050
<b>MODEL SUMMARY</b>									
	-2 Log Likelihood	15386.223				15387.221			
	Nagelkerke R <sup>2</sup>	0.155				0.155			
	N	25,297				25,297			

Note: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables.

**Table S1.10.** Generalized linear model with citation scores (CS) as the outcome variable (quadratic terms)

		Exp(B)	CI (95%)	S.E.	Exp(B)	CI (95%)	S.E.
	Intercept	2.773	2.563 3.001	0.0403	2.755	2.548 2.979	0.0399
	f_weight	1.007	0.861 1.177	0.0797	---		
	Diversity Index	---			1.081	0.775 1.508	0.1699
	F weight (quadratic)	0.967	0.807 1.158	0.0921	---		
	Diversity (quadratic)	---			0.830	0.408 1.688	0.3625
<b>CONTEXT</b>							
	Univ. Prestige	1.105	1.055 1.158	0.0239	1.106	1.055 1.159	0.0239
	Business prestige	1.185	1.139 1.234	0.0204	1.185	1.139 1.234	0.0204
	Arab States	0.944	0.799 1.117	0.0855	0.943	0.797 1.115	0.0855
	East Asia	1.258	1.061 1.491	0.0866	1.261	1.064 1.494	0.0867
	Common Wealth St.	0.572	0.344 0.950	0.2590	0.572	0.344 0.951	0.2591
	Latin America	0.793	0.646 0.974	0.1049	0.793	0.645 0.974	0.1049
	North America	1.630	1.523 1.745	0.0348	1.630	1.522 1.745	0.0349
	SW Asia	0.652	0.567 0.751	0.0718	0.652	0.566 0.751	0.0718
	SCE Europe	0.870	0.756 1.001	0.0717	0.870	0.756 1.001	0.0719
	Sub-Saharan Africa	0.447	0.375 0.533	0.0899	0.447	0.375 0.533	0.0899
	Western Europe	1.238	1.160 1.322	0.0335	1.239	1.160 1.322	0.0335
<b>PUB CHARACTERISTICS</b>							
	Self-citations	1.256	1.241 1.271	0.0062	1.256	1.241 1.271	0.0061
	Institutional collab.	1.070	1.029 1.112	0.0200	1.070	1.029 1.112	0.0199
	International collab.	1.150	1.104 1.197	0.0205	1.149	1.104 1.196	0.0205
	Pub_year (2007)	0.613	0.542 0.695	0.0637	0.614	0.542 0.696	0.0637
	Pub_year (2008)	0.931	0.877 0.988	0.0304	0.931	0.877 0.988	0.0304
	Pub_year (2009)	0.909	0.859 0.961	0.0285	0.910	0.860 0.962	0.0284
	Pub_year (2010)	0.990	0.937 1.046	0.0282	0.991	0.937 1.047	0.0282
	Pub_year (2011)	0.915	0.864 0.968	0.0290	0.915	0.865 0.969	0.0289
	Pub_year (2012)	0.955	0.903 1.010	0.0285	0.956	0.904 1.011	0.0285
<b>GOODNESS OF FIT</b>							
	Deviance Value/DF	1.116			1.116		
	$\chi^2$ Value /DF	1.359			1.359		
	Log Likelihood	-65210.879			-65211.198		
<b>OMNIBUS TEST</b>							
	Log-likelihood $\chi^2$	4062.414 (P<0.0001)			4061.777 (P<0.0001)		
	DF	22			22		
	N	25,297			25,297		

Note: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables. 2013 is the reference group for the Pub\_year variables.

**Table S1.11.** Logistic regression model with PP top-10% as the outcome variable (interaction terms for self-citations and international collaboration)

		Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
	f_weight	0.942	0.765	1.161	0.107	---			
	Diversity Index	---				1.220	0.874	1.701	0.170
<b>CONTEXT</b>	Univ. Prestige	1.148	1.018	1.294	0.061	1.148	1.018	1.295	0.061
	Business prestige	1.405	1.262	1.564	0.055	1.404	1.261	1.564	0.055
	Arab States	0.555	0.195	1.581	0.534	0.556	0.195	1.583	0.534
	East Asia	2.771	0.978	7.852	0.531	2.769	0.977	7.850	0.532
	Common Wealth St.	0.734	0.100	5.364	1.015	0.726	0.099	5.335	1.018
	Latin America	0.769	0.439	1.349	0.286	0.770	0.439	1.349	0.286
	North America	2.706	2.185	3.351	0.109	2.706	2.185	3.350	0.109
	SW Asia	0.613	0.336	1.118	0.307	0.612	0.336	1.116	0.306
	SCE Europe	0.711	0.441	1.147	0.244	0.713	0.442	1.150	0.244
	Sub-Saharan Africa	0.244	0.071	0.834	0.627	0.244	0.071	0.835	0.627
	Western Europe	1.516	1.226	1.875	0.108	1.516	1.226	1.875	0.108
<b>PUBLICATION CHARACTERISTICS</b>									
	Self-citations	1.482	1.436	1.529	0.016	1.487	1.439	1.535	0.016
	Collaboration	1.140	1.015	1.279	0.059	1.139	1.015	1.279	0.059
	International Collab.	1.278	1.119	1.458	0.068	1.375	1.201	1.573	0.069
	f_weight*self-citations	0.985	0.915	1.060	0.038				0.059
	f_weight*Int. collab.	1.077	0.805	1.440	0.148				0.236
	Diversity Index*self-citations					0.963	0.857	1.082	0.016
	Diversity Index*Int. collab.					0.775	0.488	1.232	0.059
<b>MODEL SUMMARY</b>									
	-2 Log Likelihood	15386.440				15385.524			
	Nagelkerke R <sup>2</sup>	0.155				0.155			
	N	25,297				25,297			

Note: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables.

**Table S1.12.** Generalized linear model with citation scores (CS) as the outcome variable (interaction terms for self-citations and international collaboration)

		Exp(B)	CI		S.E.	Exp(B)	CI (95%)		S.E.
			(95%)						
	Intercept	2.819	2.603	3.053	0.0407	2.769	2.559	2.996	0.0402
	f_weight	0.932	0.864	1.005	0.0388	---			
<b>CONTEXT</b>	Diversity Index	---				0.984	0.876	1.104	0.0590
	Univ. Prestige	1.106	1.055	1.159	0.0239	1.105	1.055	1.158	0.0239
	Business prestige	1.185	1.138	1.233	0.0204	1.185	1.139	1.234	0.0204
	Arab States	0.943	0.798	1.116	0.0855	0.944	0.798	1.116	0.0856
	East Asia	1.261	1.064	1.494	0.0866	1.259	1.062	1.492	0.0867
	Common Wealth St.	0.578	0.348	0.959	0.2588	0.573	0.344	0.953	0.2598
	Latin America	0.794	0.647	0.975	0.1047	0.793	0.645	0.974	0.1050
	North America	1.630	1.523	1.746	0.0349	1.630	1.522	1.746	0.0349
	SW Asia	0.653	0.567	0.751	0.0717	0.652	0.567	0.751	0.0718
	SCE Europe	0.869	0.755	1.000	0.0715	0.870	0.756	1.001	0.0717
<b>PUB CHARACTERISTICS</b>	Sub-Saharan Africa	0.445	0.374	0.531	0.0896	0.447	0.375	0.533	0.0898
	Western Europe	1.239	1.160	1.323	0.0335	1.239	1.160	1.323	0.0335
	Self-citations	1.246	1.225	1.268	0.0087	1.257	1.235	1.280	0.0090
	Institutional collab.	1.069	1.028	1.112	0.0199	1.070	1.029	1.112	0.0199
	International collab.	1.113	1.055	1.175	0.0275	1.138	1.077	1.202	0.0278
	Pub_year (2007)	0.613	0.542	0.695	0.0636	0.614	0.542	0.696	0.0637
	Pub_year (2008)	0.931	0.877	0.988	0.0304	0.931	0.877	0.988	0.0304
	Pub_year (2009)	0.910	0.860	0.962	0.0284	0.909	0.860	0.962	0.0285
	Pub_year (2010)	0.991	0.938	1.047	0.0282	0.990	0.937	1.047	0.0282
	Pub_year (2011)	0.916	0.866	0.969	0.0289	0.915	0.865	0.969	0.0289
<b>GOODNESS OF FIT</b>	Pub_year (2012)	0.956	0.904	1.011	0.0285	0.956	0.904	1.011	0.0285
	f_weight*self-citations	1.027	0.985	1.070	0.0212				
	f_weight*Int. collab.	1.110	0.984	1.251	0.0611				
	Diversity Index*self-citations					0.996	0.932	1.064	0.0337
	Diversity Index*Int. collab.					1.051	0.867	1.273	0.0979
	Deviance Value/DF	1.116					1.116		
	$\chi^2$ Value /DF	1.359					1.359		
	Log Likelihood	-65207.920					-65211.207		
	Log-likelihood $\chi^2$	4068.332					4061.758		
	DF	23					23		
<b>OMNIBUS TEST</b>	N	25,297					25,297		

Note: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables.

**Table S1.13.** Logistic regression model with PP top-10% as the outcome variable (interaction terms for the geographical groupings)

	Exp(B)	CI (95%)		S.E.	Exp(B)	CI (95%)		S.E.
f_weight	0.873	0.443	1.720	0.346	---			
Diversity Index	---				1.318	0.450	3.862	0.548
Context								
Univ. Prestige	1.148	1.018	1.295	0.061	1.149	1.019	1.295	0.061
Business prestige	1.405	1.262	1.564	0.055	1.404	1.261	1.564	0.055
Arab States	0.585	0.145	2.364	0.712	0.409	0.081	2.064	0.826
East Asia	2.489	0.620	9.984	0.709	4.044	0.807	20.279	0.823
Common Wealth St.	0.833	0.065	10.733	1.304	0.885	0.072	10.851	1.279
Latin America	0.975	0.459	2.070	0.384	0.936	0.430	2.037	0.397
North America	2.596	1.927	3.499	0.152	2.719	1.993	3.709	0.158
SW Asia	0.510	0.202	1.286	0.472	0.524	0.195	1.405	0.503
SCE Europe	0.616	0.309	1.228	0.352	0.558	0.263	1.185	0.384
Sub-Saharan Africa	0.160	0.024	1.068	0.968	0.153	0.018	1.294	1.091
Western Europe	1.515	1.123	2.044	0.153	1.695	1.242	2.314	0.159
Publication Behavior								
Self-citations	1.475	1.444	1.508	0.011	1.476	1.444	1.508	0.011
Collaboration	1.139	1.015	1.279	0.059	1.141	1.016	1.280	0.059
International Collab.	1.307	1.184	1.442	0.050	1.305	1.183	1.440	0.050
f_weight*Arab States	0.829	0.027	25.028	1.739				
f_weight*East Asia	1.444	0.049	42.823	1.729				
f_weight*Common Wealth St.	0.586	0.000	703.608	3.617				
f_weight*Latin America	0.412	0.052	3.253	1.054				
f_weight*North America	1.147	0.566	2.327	0.361				
f_weight*SW Asia	1.737	0.238	12.701	1.015				
f_weight*SCE Europe	1.607	0.337	7.670	0.798				
f_weight*Sub-Saharan Africa	3.676	0.081	167.252	1.948				
f_weight*Western Europe	1.002	0.489	2.053	0.366				
Diversity I.*Arab States					4.986	0.018	1377.234	2.868
Diversity I.*East Asia					0.139	0.001	37.485	2.855
Diversity I.*Common Wealth St.					0.323	0.000	15471.206	5.498
Diversity I.*Latin America					0.347	0.016	7.629	1.576
Diversity I.*North America					1.914	0.071	51.754	1.682
Diversity I.*SW Asia					0.967	0.315	2.968	0.572
Diversity I.*SCE Europe					3.111	0.240	40.325	1.307
Diversity I.*Sub-Saharan Africa					7.826	0.010	6163.560	3.403
Diversity I.*Western Europe					0.564	0.181	1.760	0.581
Model Summary								
-2 Log Likelihood	15383.915				15379.402			
Nagelkerke R <sup>2</sup>	0.155				0.156			
N	25,297				25,297			

Note: Single-authored papers are excluded. *Oceania* is the reference group for the geographical variables.

**Table S1.14.** Generalized linear model with citation scores (CS) as the outcome variable (interaction terms for the geographical groupings)

		Exp(B)	CI (95%)	S.E.	Exp(B)	CI (95%)	S.E.
	Intercept	2.756	2.489 3.052	0.0520	2.798	2.527 3.097	0.0519
	f_weight	1.006	0.829 1.219	0.0984	---		
	Diversity Index	---			0.926	0.666 1.287	0.1682
<b>CONTEXT</b>	Univ. Prestige	1.105	1.055 1.158	0.0239	1.107	1.056 1.160	0.0239
	Business prestige	1.185	1.138 1.233	0.0204	1.185	1.138 1.233	0.0204
	Arab States	1.027	0.809 1.303	0.1218	0.974	0.763 1.242	0.1242
	East Asia	1.137	0.894 1.447	0.1230	1.123	0.878 1.438	0.1259
	Common Wealth St.	0.724	0.363 1.445	0.3522	0.483	0.248 0.938	0.3391
	Latin America	0.744	0.556 0.995	0.1487	0.704	0.531 0.934	0.1440
	North America	1.643	1.487 1.816	0.0511	1.572	1.422 1.738	0.0512
	SW Asia	0.652	0.532 0.800	0.1039	0.603	0.489 0.744	0.1070
	SCE Europe	0.910	0.750 1.104	0.0985	0.873	0.718 1.062	0.0999
	Sub-Saharan Africa	0.499	0.400 0.622	0.1125	0.486	0.387 0.609	0.1155
	Western Europe	1.250	1.134 1.378	0.0498	1.260	1.141 1.390	0.0502
<b>PUBLICATION CHARACTERISTICS</b>							
	Self-citations	1.256	1.241 1.272	0.0062	1.256	1.241 1.271	0.0062
	Institutional collab.	1.069	1.028 1.112	0.0199	1.071	1.030 1.114	0.0199
	International collab.	1.150	1.105 1.197	0.0205	1.149	1.104 1.196	0.0204
	Pub_year (2007)	0.614	0.542 0.696	0.0638	0.614	0.542 0.696	0.0637
	Pub_year (2008)	0.930	0.877 0.988	0.0304	0.933	0.879 0.991	0.0304
	Pub_year (2009)	0.909	0.859 0.961	0.0285	0.910	0.861 0.963	0.0284
	Pub_year (2010)	0.990	0.937 1.046	0.0282	0.991	0.938 1.047	0.0282
	Pub_year (2011)	0.915	0.865 0.968	0.0289	0.916	0.866 0.969	0.0289
	Pub_year (2012)	0.956	0.904 1.011	0.0285	0.957	0.906 1.012	0.0284
	f_weight*Arab States	0.763	0.448 1.300	0.2716			
	f_weight*East Asia	1.390	0.807 2.393	0.2771			
	f_weight*Common Wealth St.	0.434	0.139 1.358	0.5821			
	f_weight*Latin America	1.219	0.676 2.200	0.3011			
	f_weight*North America	0.975	0.791 1.201	0.1068			
	f_weight*SW Asia	0.998	0.627 1.588	0.2371			
	f_weight*SCE Europe	0.860	0.570 1.297	0.2096			
	f_weight*Sub-Saharan Africa	0.660	0.344 1.267	0.3326			
	f*weight*Western Europe	0.969	0.782 1.201	0.1094			
	Diversity I.*Arab States				0.799	0.309 2.064	0.4841
	Diversity I.*East Asia				1.892	0.724 4.944	0.4900
	Diversity I.*Common Wealth St.				2.575	0.348 19.073	1.0217
	Diversity I.*Latin America				1.875	0.658 5.346	0.5346
	Diversity I.*North America				1.199	0.840 1.713	0.1819
	Diversity I.*SW Asia				1.438	0.676 3.059	0.3852
	Diversity I.*SCE Europe				0.978	0.412 2.318	0.4403
	Diversity I.*Sub- Saharan Africa				0.623	0.225 1.727	0.5204
	Diversity I.*Western Europe				0.919	0.645 1.310	0.1807
<b>GOODNESS OF FIT</b>							
	Deviance Value/DF	1.116			1.116		
	$\chi^2$ Value /DF	1.361			1.358		



	Log Likelihood	-65207.493	-65200.558
<b>OMNIBUS TEST</b>			
	Log-likelihood $\chi^2$	4069.186 (P<0.0001)	4083.057 (P<0.0001)
	DF	30	30
	N	25,297	25,297

Note: Single-authored papers are excluded. The models are computed with robust covariance matrix estimation. *Oceania* is the reference group for the geographical variables.

**Table S.1.15.** Binary logistic regression with first-author gender as the outcome variable (1=female)

	Exp(B)	CI (95%)	S.E.
University Prestige	0.950	0.869 1.039	0.046
Business Prestige	0.889	0.822 0.961	0.040
Arab States	0.321	0.200 0.516	0.242
East Asia	1.604	0.985 2.612	0.249
Common Wealth Independent States	1.374	0.525 3.597	0.491
Latin America	0.680	0.524 0.883	0.133
North America	0.721	0.642 0.811	0.060
South and West Asia	0.404	0.310 0.528	0.136
South-Central and Eastern Europe	1.238	1.016 1.509	0.101
Sub-Saharan Africa	0.634	0.471 0.855	0.152
Western Europe	0.821	0.735 0.918	0.057
Innovation	2.574	1.284 5.160	0.355
Dynamic Capabilities	1.101	0.543 2.234	0.361
Organizational Learning	7.085	3.217 15.601	0.403
Time	0.860	0.414 1.785	0.373
IOR	3.138	1.432 6.879	0.400
Strategic Mgmt	3.142	1.557 6.342	0.358
Leadership	5.829	2.764 12.293	0.381
Commitment	5.265	2.912 9.518	0.302
Decision-making	0.925	0.459 1.864	0.357
HRM	24.876	12.466 49.641	0.353
Survey Studies	3.402	1.761 6.571	0.336
Multinational biz	2.948	1.480 5.871	0.352
Causal effects	1.007	0.507 2.002	0.351
Game theory	2.238	0.997 5.020	0.412
Structural Inequality	49.192	23.102 104.745	0.386
Behav. Mgmt	4.232	2.147 8.340	0.346
Knowledge transfer	3.896	1.964 7.725	0.349
Corporate Governance	1.816	0.862 3.826	0.380
Project Mgmt	1.177	0.579 2.393	0.362
Consumer econ.	0.717	0.350 1.471	0.366
Operations algorithms	0.853	0.444 1.641	0.334
Crisis Mgmt	0.956	0.477 1.916	0.355
Service oper.	3.314	1.549 7.090	0.388
Predictive Modeling	0.627	0.322 1.221	0.340
Lit reviews	1.150	0.633 2.087	0.304
Inventory mgmt.	0.840	0.409 1.726	0.367
CSR	10.402	5.042 21.461	0.370
Tourism	6.106	2.983 12.501	0.366
Online Marketing	3.441	1.660 7.134	0.372
Supply chain mgmt.	1.597	0.744 3.425	0.389
Socio econ pol	1.340	0.677 2.652	0.348
Team mgmt.	9.289	4.550 18.964	0.364
Employee appraisals	4.604	2.323 9.124	0.349

Constructionism	5.022	2.810	8.975	0.296
Healthcare mgmt.	115.399	59.352	224.375	0.339
<b>Model Summary</b>		-2 Log Likelihood		28765.975
		Nagelkerke		.105
		R <sup>2</sup>		
		N		24,812

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables.

**Table S1.16.** Binary logistic regression with author gender as the outcome variable (1=female)

	Exp(B)	CI (95%)		S.E.
University Prestige	.910	.861	.962	.028
Business Prestige	.893	.850	.938	.025
Arab States	.402	.299	.539	.150
East Asia	.533	.474	.598	.059
Common Wealth Independent States	1.732	1.017	2.951	.272
Latin America	.774	.665	.901	.078
North America	.775	.721	.833	.037
South and West Asia	.443	.378	.520	.081
South-Central and Eastern Europe	1.329	1.177	1.500	.062
Sub-Saharan Africa	.636	.522	.774	.100
Western Europe	.802	.748	.859	.035
Innovation	1.907	1.224	2.971	.226
Dynamic Capabilities	1.106	.706	1.733	.229
Organizational Learning	6.041	3.672	9.940	.254
Time	.917	.577	1.456	.236
IOR	3.133	1.914	5.127	.251
Strategic Mgmt	1.963	1.253	3.073	.229
Leadership	4.308	2.712	6.842	.236
Commitment	4.525	3.126	6.548	.189
Decision-making	.912	.589	1.413	.223
HRM	18.363	11.929	28.267	.220
Survey Studies	3.13	2.071	4.733	.211
Multinational biz	2.118	1.367	3.284	.224
Causal effects	.797	.519	1.266	.219
Game theory	1.73	1.035	2.892	.262
Structural Inequality	32.989	20.469	53.166	.244
Behav. Mgmt	3.957	2.604	6.013	.213
Knowledge transfer	2.895	1.871	4.480	.223
Corporate Governance	1.047	.644	1.702	.248
Project Mgmt	.917	.585	1.439	.230
Consumer econ.	.741	.470	1.166	.232
Operations algorithms	.693	.460	1.045	0.21
Crisis Mgmt	1.159	.744	1.807	.226
Service oper.	2.653	1.648	4.269	.243
Predictive Modeling	.65	.428	.988	.214
Lit reviews	1.315	.902	1.915	.192
Inventory mgmt.	.62	.395	.973	.230
CSR	8.33	5.261	13.190	.234
Tourism	4.467	2.842	7.021	.231
Online Marketing	1.894	1.196	2.999	.234
Supply chain mgmt.	1.007	.620	1.636	.247
Socio econ pol	1.790	1.168	2.744	.218
Team mgmt.	6.113	3.926	9.519	.226
Employee appraisals	4.725	3.090	7.225	.217
Constructionism	4.587	3.169	6.640	.189
Healthcare mgmt.	70.149	46.894	104.935	.205

# Model Summary

-2 Log Likelihood	74307.095
Nagelkerke R <sup>2</sup>	.096
N	65,194

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables.

**Table S1.17.** Linear regression model with log-journal score as the outcome variable

	B	CI (95%)	S.E.
Constant	0.442	0.394 0.491	0.025
<b>CONTEXT</b>			
Univ. Prestige	0.039	0.030 0.048	0.005
Business prestige	0.054	0.047 0.062	0.004
Arab States	0.044	0.004 0.085	0.021
East Asia	0.020	-0.021 0.060	0.021
Common Wealth St.	0.012	-0.084 0.107	0.049
Latin America	-0.053	-0.078 -0.028	0.013
North America	0.141	0.129 0.153	0.006
SW Asia	-0.131	-0.155 -0.107	0.012
SCE Europe	-0.049	-0.071 -0.028	0.011
Sub-Saharan Africa	-0.154	-0.183 -0.124	0.015
Western Europe	0.083	0.072 0.095	0.006
Institutional collab.	0.029	0.021 0.036	0.004
International collab.	0.036	0.029 0.043	0.004
f_weight	0.004	-0.006 0.013	0.005
Innovation	0.091	0.026 0.156	0.033
Dynamic Capabilities	0.492	0.427 0.556	0.033
Organizational Learning	0.076	-0.001 0.154	0.040
Time	0.103	0.036 0.169	0.034
IOR	0.269	0.195 0.344	0.038
Strategic Mgmt	-0.062	-0.129 0.004	0.034
Leadership	-0.012	-0.085 0.062	0.037
Commitment	0.295	0.241 0.350	0.028
Decision making	0.114	0.051 0.178	0.032
HRM	-0.098	-0.165 -0.031	0.034
Survey Studies	-0.177	-0.239 -0.115	0.031
Multinational biz	0.126	0.062 0.190	0.033
Causal effects	0.432	0.369 0.495	0.032
Game theory	0.109	0.032 0.185	0.039
Structural Inequality	0.112	0.036 0.187	0.038
Behav. Mgmt	0.466	0.401 0.531	0.033
Knowledge transfer	0.246	0.182 0.311	0.033
Corporate Governance	0.266	0.197 0.335	0.035
Project Mgmt	-0.180	-0.246 -0.115	0.033
Consumer econ.	0.127	0.066 0.188	0.031

	Operations algorithms	0.161	0.105	0.217	0.029
	Crisis Mgmt	-0.093	-0.157	-0.028	0.033
	Service oper.	-0.046	-0.118	0.026	0.037
	Predictive Modeling	0.091	0.033	0.149	0.030
	Lit reviews	0.328	0.274	0.382	0.027
	Inventory mgmt	0.117	0.055	0.179	0.032
	CSR	0.203	0.133	0.274	0.036
	Tourism	0.203	0.134	0.273	0.035
	Online Market	0.027	-0.041	0.095	0.035
	Supply chain mgmt.	0.298	0.228	0.368	0.036
	Socio econ pol	-0.186	-0.249	-0.122	0.032
	Team mgmt	0.278	0.208	0.348	0.036
	Employee appraisals	0.348	0.283	0.413	0.033
	Constructionism	0.279	0.225	0.332	0.027
	Healthcare mgmt	0.120	0.059	0.181	0.031
<b>MODEL SUMMARY</b>					
	R <sup>2</sup>	0.219			
	Adjusted R <sup>2</sup>	0.217			
	N	27,664			

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables. B refers to unstandardized beta-coefficients.

**Table S1.18.** Linear regression model with log-journal score as the outcome variable (without topics).

		<b>B</b>	<b>CI (95%)</b>		<b>S.E.</b>
<b>CONTEXT</b>	Constant	0.561	0.549	0.574	0.006
	Univ. Prestige	0.040	0.030	0.049	0.005
	Business prestige	0.063	0.055	0.071	0.004
	Arab States	0.035	-0.007	0.077	0.022
	East Asia	0.048	0.005	0.090	0.022
	Common Wealth St.	-0.001	-0.101	0.099	0.051
	Latin America	-0.081	-0.108	-0.055	0.013
	North America	0.170	0.157	0.182	0.006
	SW Asia	-0.164	-0.189	-0.139	0.013
	SCE Europe	-0.081	-0.103	-0.059	0.011
	Sub-Saharan Africa	-0.205	-0.236	-0.174	0.016
	Western Europe	0.091	0.079	0.103	0.006
	Institutional collab.	0.043	0.035	0.050	0.004
	International collab.	0.040	0.033	0.048	0.004
	<i>f_weight</i>	0.002	-0.008	0.012	0.005
<b>MODEL SUMMARY</b>					
	R <sup>2</sup>	0.137			
	Adjusted R <sup>2</sup>	0.137			
	N	27,664			

Note: *Oceania* is the reference group for the geographical variables. B refers to unstandardized beta-coefficients.

**Table S1.19.** Logistic regression with ‘Financial times’ as the outcome variable (1=journal included in the FT journal list)

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.031	1.870	2.206 0.042
University Prestige	1.722	1.574	1.884 0.046
Arab States	1.190	0.622	2.279 0.331
East Asia	1.477	0.779	2.802 0.326
Common Wealth Independent St.	0.829	0.184	3.744 0.769
Latin America	0.821	0.524	1.288 0.230
North America	4.071	3.405	4.867 0.091
South & West Asia	0.869	0.549	1.376 0.235
South-Central & Eastern Europe	0.894	0.592	1.349 0.210
Sub-Saharan Africa	0.176	0.048	0.636 0.657
Western Europe	1.753	1.467	2.095 0.091
Institutional Collaboration	1.479	1.357	1.614 0.044
International Collaboration	1.354	1.252	1.465 0.040
F_weight	1.025	0.918	1.144 0.056
Innovation	0.261	0.125	0.548 0.378
Dynamic Capabilities	7.075	3.584	13.968 0.347
Organizational Learning	0.565	0.239	1.338 0.440
Time	0.533	0.255	1.115 0.377
IOR	1.868	0.851	4.105 0.402
Strategic Mgmt	0.057	0.026	0.128 0.410
Leadership	0.006	0.002	0.015 0.497
Commitment	0.671	0.372	1.210 0.301
Decision making	0.061	0.028	0.131 0.392
HRM	0.251	0.115	0.548 0.398
Survey Studies	0.001	0.001	0.003 0.431
Multinational biz	0.814	0.408	1.623 0.352
Causal effects	1.246	0.629	2.465 0.348
Game theory	2.310	1.048	5.089 0.403
Structural Inequality	0.272	0.117	0.632 0.431
Behav. Mgmt	2.832	1.432	5.603 0.348
Knowledge transfer	3.455	1.747	6.830 0.348
Corporate Governance	0.906	0.438	1.872 0.370
Project Mgmt	0.029	0.013	0.066 0.422
Consumer econ.	1.573	0.833	2.971 0.324
Operations algorithms	1.353	0.742	2.465 0.306
Crisis Mgmt	4.467	2.250	8.868 0.350
Service oper.	0.031	0.011	0.083 0.504
Predictive Modeling	0.167	0.086	0.325 0.339
Lit reviews	0.158	0.087	0.287 0.306
Inventory mgmt	2.033	1.055	3.919 0.335
CSR	0.028	0.011	0.073 0.485
Tourism	0.000	0.000	0.000 0.792
Online Market	1.406	0.674	2.933 0.375
Supply chain mgmt.	0.029	0.012	0.070 0.441
Socio econ pol	0.087	0.040	0.191 0.402
Team mgmt	1.414	0.686	2.913 0.369
Employee appraisals	4.738	2.398	9.362 0.347
Constructionism	2.946	1.661	5.223 0.292
Healthcare mgmt	0.031	0.013	0.074 0.440
<b>Model summary</b>	-2 Log likelihood	22517.922	
	Cox & Snell R <sup>2</sup>	0.214	
	Nagelkerke R	0.329	
	N	27,664	

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables.

**Table S1.20.** Logistic regression with ‘UT Dallas’ as outcome the variable (1=journal included in the UTD journal list)

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.301	2.064	2.565 0.055
University Prestige	1.684	1.499	1.893 0.060
Arab States	1.519	0.643	3.588 0.439
East Asia	1.326	0.574	3.063 0.427
Common Wealth Independent St.	0.964	0.136	6.816 0.998
Latin America	0.748	0.401	1.393 0.317
North America	4.838	3.720	6.293 0.134
South & West Asia	1.134	0.621	2.072 0.308
South-Central & Eastern Europe	1.247	0.718	2.165 0.281
Sub-Saharan Africa	0.375	0.073	1.927 0.835
Western Europe	1.399	1.071	1.827 0.136
Institutional Collaboration	1.887	1.671	2.132 0.062
International Collaboration	1.336	1.208	1.478 0.051
F_weight	0.890	0.769	1.030 0.075
Innovation	0.006	0.002	0.015 0.519
Dynamic Capabilities	4.116	1.908	8.878 0.392
Organizational Learning	0.375	0.126	1.115 0.556
Time	0.099	0.040	0.245 0.463
IOR	0.639	0.251	1.625 0.476
Strategic Mgmt	0.002	0.001	0.007 0.574
Leadership	0.000	0.000	0.001 0.866
Commitment	0.012	0.005	0.028 0.420
Decision making	0.039	0.016	0.097 0.466
HRM	0.003	0.001	0.012 0.639
Survey Studies	0.000	0.000	0.001 0.601
Multinational biz	0.761	0.343	1.689 0.407
Causal effects	0.399	0.177	0.901 0.416
Game theory	0.609	0.243	1.528 0.469
Structural Inequality	0.048	0.015	0.150 0.579
Behav. Mgmt	0.134	0.057	0.315 0.437
Knowledge transfer	0.143	0.061	0.338 0.438
Corporate Governance	0.458	0.200	1.047 0.422
Project Mgmt	0.018	0.007	0.052 0.526
Consumer econ.	0.949	0.473	1.906 0.356
Operations algorithms	1.202	0.621	2.326 0.337
Crisis Mgmt	0.018	0.007	0.045 0.468
Service oper.	0.023	0.007	0.079 0.620
Predictive Modeling	0.108	0.051	0.230 0.385
Lit reviews	0.041	0.020	0.083 0.368
Inventory mgmt	1.867	0.908	3.840 0.368
CSR	0.011	0.003	0.037 0.638
Tourism	0.000	0.000	0.000 1.130
Online Market	1.519	0.658	3.504 0.427
Supply chain mgmt.	0.049	0.019	0.129 0.490
Socio econ pol	0.001	0.000	0.003 0.634
Team mgmt	0.184	0.075	0.447 0.454
Employee appraisals	0.066	0.028	0.154 0.437
Constructionism	0.428	0.218	0.839 0.344
Healthcare mgmt	0.028	0.010	0.082 0.544
<b>Model summary</b>	-2 Log likelihood 13916.846		
	Cox & Snell R <sup>2</sup> 0.187		
	Nagelkerke R 0.368		
	N 27,664		

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables.

**Table S1.21.** Logistic regression with AJG top journals (4\*) as the outcome variable (1=AJG 4\*)

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.147	1.925	2.395 0.056
University Prestige	1.622	1.441	1.825 0.060
Arab States	1.425	0.595	3.413 0.446
East Asia	1.346	0.575	3.155 0.435
Common Wealth Independent St.	0.951	0.135	6.684 0.995
Latin America	0.705	0.371	1.338 0.327
North America	4.509	3.469	5.860 0.134
South & West Asia	1.067	0.574	1.984 0.316
South-Central & Eastern Europe	1.230	0.704	2.151 0.285
Sub-Saharan Africa	0.373	0.073	1.916 0.835
Western Europe	1.348	1.033	1.760 0.136
Institutional Collaboration	1.819	1.609	2.058 0.063
International Collaboration	1.371	1.239	1.517 0.052
F_weight	0.920	0.794	1.065 0.075
Innovation	0.005	0.002	0.015 0.523
Dynamic Capabilities	4.743	2.199	10.231 0.392
Organizational Learning	0.410	0.138	1.214 0.554
Time	0.108	0.043	0.268 0.465
IOR	0.735	0.290	1.864 0.475
Strategic Mgmt	0.002	0.001	0.007 0.579
Leadership	0.000	0.000	0.001 0.865
Commitment	0.014	0.006	0.031 0.420
Decision making	0.047	0.019	0.118 0.466
HRM	0.004	0.001	0.013 0.643
Survey Studies	0.000	0.000	0.001 0.606
Multinational biz	0.781	0.352	1.735 0.407
Causal effects	0.468	0.207	1.056 0.416
Game theory	0.584	0.233	1.467 0.470
Structural Inequality	0.055	0.018	0.170 0.578
Behav. Mgmt	0.141	0.060	0.333 0.437
Knowledge transfer	0.157	0.067	0.369 0.437
Corporate Governance	0.488	0.213	1.115 0.422
Project Mgmt	0.020	0.007	0.055 0.530
Consumer econ.	0.542	0.269	1.092 0.357
Operations algorithms	1.282	0.661	2.485 0.338
Crisis Mgmt	0.024	0.009	0.060 0.468
Service oper.	0.011	0.003	0.041 0.652
Predictive Modeling	0.104	0.049	0.223 0.388
Lit reviews	0.045	0.022	0.094 0.368
Inventory mgmt	0.482	0.231	1.002 0.374
CSR	0.008	0.002	0.028 0.654
Tourism	0.000	0.000	0.000 1.138
Online Market	2.028	0.882	4.664 0.425
Supply chain mgmt.	0.041	0.015	0.108 0.501
Socio econ pol	0.001	0.000	0.003 0.633
Team mgmt	0.182	0.075	0.443 0.455
Employee appraisals	0.073	0.031	0.172 0.437
Constructionism	0.450	0.229	0.882 0.344
Healthcare mgmt	0.024	0.008	0.072 0.565
<b>Model summary</b>	-2 Log likelihood	13821.427	
	Cox & Snell R <sup>2</sup>	0.168	
	Nagelkerke R <sup>2</sup>	0.339	
	N	27,664	

Note: *Oceania* is the reference group for the geographical variables. *Financial governance* is the reference group for the topic variables.



**Table S1.22.** Binary logistic regression with ‘Financial times’ as the outcome variable (1=journal included in the FT journal list) (without the topic variables).

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.238	2.072	2.418 0.039
University Prestige	1.840	1.692	2.001 0.043
Arab States	1.207	0.646	2.256 0.319
East Asia	2.054	1.109	3.804 0.315
Common Wealth Independent St.	1.188	0.275	5.122 0.746
Latin America	0.953	0.618	1.469 0.221
North America	5.833	4.936	6.894 0.085
South & West Asia	0.856	0.549	1.336 0.227
South-Central & Eastern Europe	0.816	0.549	1.213 0.202
Sub-Saharan Africa	0.147	0.041	0.525 0.652
Western Europe	2.146	1.815	2.537 0.085
Institutional Collaboration	1.631	1.504	1.769 0.041
International Collaboration	1.409	1.309	1.516 0.037
F_weight	1.003	0.905	1.112 0.053
<b>Model summary</b>	-2 Log likelihood 25372.634		
	Cox & Snell R <sup>2</sup> 0.129		
	Nagelkerke R 0.198		
	N 27,664		

Note: *Oceania* is the reference group for the geographical variables.

**Table S1.23.** Binary logistic regression with ‘UT Dallas’ as the outcome variable (1=journal included in the UTD journal list) (without the topic variables)

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.413	2.184	2.666 0.051
University Prestige	1.905	1.713	2.119 0.054
Arab States	1.846	0.816	4.178 0.417
East Asia	1.952	0.882	4.324 0.406
Common Wealth Independent St.	1.828	0.278	12.007 0.960
Latin America	1.375	0.761	2.484 0.302
North America	6.930	5.413	8.871 0.126
South & West Asia	1.491	0.840	2.646 0.293
South-Central & Eastern Europe	1.337	0.789	2.265 0.269
Sub-Saharan Africa	0.261	0.052	1.326 0.828
Western Europe	1.745	1.356	2.247 0.129
Institutional Collaboration	2.039	1.822	2.281 0.057
International Collaboration	1.408	1.285	1.543 0.047
F_weight	0.890	0.778	1.019 0.069
<b>Model summary</b>	-2 Log likelihood 16562.554		
	Cox & Snell R <sup>2</sup> 0.105		
	Nagelkerke R 0.208		
	N 27,664		

Note: *Oceania* is the reference group for the geographical variables.

**Table S1.24.** Binary logistic regression with AJG top journals (4\*) as the outcome variable (1=AJG 4\*) (without the topic variables)

	<i>Exp(B)</i>	<i>CI (95%)</i>	<i>S.E.</i>
Business Prestige	2.282	2.061	2.527 0.052
University Prestige	1.826	1.637	2.036 0.056
Arab States	1.670	0.725	3.847 0.426
East Asia	1.980	0.877	4.470 0.415
Common Wealth Independent St.	1.737	0.265	11.409 0.960
Latin America	1.205	0.653	2.224 0.313
North America	6.351	4.960	8.133 0.126
South & West Asia	1.298	0.715	2.356 0.304
South-Central & Eastern Europe	1.209	0.705	2.071 0.275
Sub-Saharan Africa	0.251	0.049	1.271 0.828
Western Europe	1.665	1.293	2.144 0.129
Institutional Collaboration	2.001	1.784	2.245 0.059
International Collaboration	1.455	1.326	1.597 0.047
F_weight	0.920	0.802	1.054 0.070
<b>Model summary</b>			
	-2 Log likelihood	16154.532	
	Cox & Snell R <sup>2</sup>	0.095	
	Nagelkerke R <sup>2</sup>	0.192	
	N	27,664	

Note: *Oceania* is the reference group for the geographical variables.

## APPENDIX S2: Topic Modeling Parameters

### Preparing data:

The corpus has been preprocessed using the TM-package in R. Preprocessing has involved:

- Transform everything to Lower case
- Remove punctuation
- Remove numbers
- Removed standard English stopwords (e.g. “the”, “and”)
- Removed redundant whitespace
- Stemmed corpus to reduce number of unique tokens

**Table S2.1.** LDA Model Parameters.

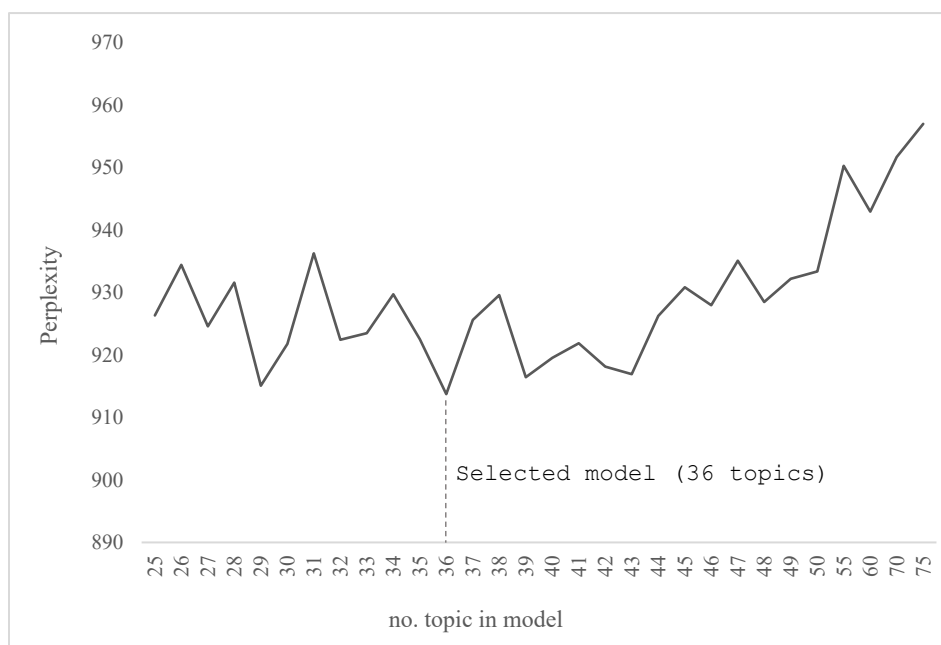
Number of topics	36
Topic Smoothing	0.01
Term Smoothing	0.01
Iterations:	1000 iterations using collapsed variational Bayes approximation to the LDA objective (CVB0LDA)
Minimum Word Length Filter	3
Term Minimum Document Count Filter	25
Term Stop List Filter	among, author, becom, can, design, methodology. approach, elsevier, even, find, howev, like, limitations, impl, make, may, must, new, often, one, originality, valu, paper, result, show, take, three, two, way, will

**Table S2.2.** Perplexity

No. of topics in model	Perplexity	No. of topics in model	Perplexity
25	926.32	40	919.54
26	934.41	41	921.85
27	924.57	42	918.12
28	931.55	43	916.91
29	915.07	44	926.23
30	921.76	45	930.82
31	936.26	46	927.97
32	922.43	47	935.06
33	923.48	48	928.46
34	929.69	49	932.17
35	922.50	50	933.35
36*	913.73	55	950.24
37	925.61	60	942.93
38	929.57	70	951.66
39	916.42	75	956.96

\*Selected model.

**Figure S2.1. Perplexity**



## APPENDIX S3: Topic summary

Table S3.1. Topic summary, including characteristic (exclusively high loading) abstract. Numbers indicate word instances of each topic/words that have occurred.

Topic 00	40454.1						
Name	Innovation						
Description	Characteristics and conditions for innovation and innovative practices in organizations						
Abstract	This paper analyzes four modes of innovation that differ in their scope of newness (innovation generation and adoption) and in their degree of change (radical and incremental). Building a theoretical model based on the market orientation (MO) and contingency theory literature and using a sample of innovative firms, we find that MO positively influences the numbers of incremental generation and adoption of innovations. We also find that environmental complexity moderates the relationship between MO and radical and incremental innovation generation and the number of incremental innovation adoption. That is, we have found that highly complex environments enhance the introduction of radical and incremental internally generated innovations and harm the adoption of incremental innovation for market-oriented firms. These findings add to the innovation and MO literatures. Our results also have important implications for both the commercial activities and R&D policies of firms.						
innov	6751.8	process	931.9	activ	437.6	compani	277.9
product	4742.9	open	631.3	idea	408.8	commerci	255.1
technolog	3150.1	industri	585.6	success	357.8	ict	230.7
develop	1631.8	npd	456.9	diffus	316.3	manufactur	215.4
market	981.7	adopt	450.3	radic	282.2	integr	203.4
Topic 01	49764.3						
Name	Dynamic capabilities						
Description	Dynamic, strategic capabilities of firms						
Abstract	In this research, we examine the dynamic capability of resource allocation to invest in operational capabilities. Using a computer simulation, we model a process of firms competing in factor markets for opportunities to invest in existing capabilities and acquire new ones. Based on the simulation results, we derive a set of propositions about the conditions under which there are and are not performance benefits from possessing a superior ability to search for new capabilities. Because the definition of what constitutes a new capability is based on a firm's preexisting capabilities, we also incorporate differences in initial endowments into the analysis. We find that endowment and search ability both matter, and that in many circumstances, the effects of possessing a superior endowment dominate the effects of superior search ability.						
firm	10493.7	market	1055.1	extern	610.7	empir	312.1
capabl	2016.1	strateg	885.8	dynam	574.7	level	311.2
competit	1610.2	outsourc	791.4	environ	346.6	intern	307.3
industri	1231.6	advantag	692.9	activ	322.0	suggest	302.2
resourc	1098.9	strategi	670.4	exploit	312.3	relat	296.3
Topic 02	28907.8						
Name	Organizational learning						
Description	Characteristics and conditions for organizational learning and adaptability						
Abstract	Organizational learning from failures is a key organizational process that can lead to improved outcomes. In this study, the authors address two key questions that have received only limited attention in the literature: (a) how learning leadership enables organizational learning from failures and (b) how these learning behaviors enhance organizational capacities for adaptation to environmental turbulence. Data from a sample of 121 organizations support a mediation model in which learning leadership is linked indirectly, through learning from failures, to perceived organizational capacity to adapt to environmental jolts. The authors discuss the theoretical and practical implications of these findings for the importance of learning leadership, organizational learning from failures, and organizational adaptability.						
organiz	3780.2	train	945.1	process	394.2	implic	194.2
organ	3225.2	compet	519.0	develop	345.0	context	188.3
learn	2866.4	practic	428.7	manag	276.9	languag	173.9
cultur	2272.3	environ	401.6	structur	260.2	enhanc	170.9
creativ	948.7	adapt	397.1	differ	209.5	valu	166.0
Topic 03	43641.2						
Name	Time						
Description	Temporal aspects of management, business environment and firm performance						
Abstracts	Based on an improved and extended database, the Establishment History Panel, we extend the analysis of Fritsch and Weyh [Small Business Economics 27(2), 245-260 (2006)] by investigating the development of employment in German start-up cohorts for the period 1976 to 2004. We confirm the typical pattern found in start-up cohorts: an initial increasing number of employees that is soon followed by a decreasing number. We also provide some of the first evidence for the "liability of aging" phenomena in Germany. Older firms face a relatively high risk of exit. We find that most entries with relatively large initial employment grow substantially in the first two decades. However, these businesses suffer from a considerable employment decline after 25 years.						
growth	1480.5	increas	856.0	data	570.6	higher	319.0
time	1084.2	bank	843.4	industri	544.3	panel	308.9
period	1073.0	year	790.6	impact	480.6	indic	306.1
chang	994.1	econom	732.4	signific	461.2	level	295.8
effect	901.4	rate	571.0	employ	352.7	evid	294.4
Topic 04	30422.3						

Name	Interorganizational relationships (IORs)						
Description	Inter-firm aspects of management, information sharing, efficiency, cooperation, competition and trust						
Abstract	The relational and dynamic aspects of interfirm trust and dependence produce a crucial, but insufficiently addressed, challenge for successful relationship coordination. In this paper we concentrate on this issue by examining how trust and dependence co-evolve in customer-supplier relationships. Building on a case study, we develop propositions and a model that illustrates how interorganizational trust and dependence coevolve through the different phases of customer-supplier relationships and how we may distinguish cooperative and trustworthy actors from those who will behave opportunistically. Theoretical and practical implications are offered.						
inform	2559.5	communic	1019.6	secur	474.5	partnership	304.5
collabor	1704.2	partner	995.3	interact	469.3	differ	297.5
trust	1627.8	allianc	724.8	coordin	454.2	develop	293.1
share	1332.1	formal	662.5	integr	382.7	interorganiz	256.8
relationship	1074.9	exchang	530.5	relat	305.0	joint	251.9
Topic 05	54040.5						
Name	Strategic management						
Description	Strategic aspects of management, management practices, roles and decisions						
Abstract	Competitive advantage has become the primary imperative for firms operating in an increasingly dynamic global marketplace, and may best be viewed through the dual prisms of strategic management and strategic human resource management (SHRM) theories. The adoption of a strategic business partner role by human resources (HR) has considerable potential to contribute to the strategic agility that firms require to successfully compete globally. This paper reports on a qualitative study of senior Indian managers' perspectives of these issues. It reinforces the importance of strategic agility, the need for close alignment between business and HRM strategies, and the growing emergence of a strategic business partner role for HR professionals in the Indian context. These findings have implications for researchers, senior managers and their HR specialists.						
manag	9207.3	practic	1091.0	organ	520.7	approach	443.6
busi	3850.0	oper	873.1	implement	510.4	align	434.0
strategi	3413.6	process	806.7	top	489.2	execut	411.6
compani	2311.3	case	629.3	role	469.8	success	409.1
strateg	1891.8	valu	601.2	identifi	445.6	adopt	400.6
Topic 06	77568.0						
Name	Leadership						
Description	Leadership and effects of different leadership styles (including teaching-related leadership)						
Abstract:	We propose a theoretical model to examine how authentic transformational leadership influences follower individual and group ethical decision making. We investigate how follower moral identity and moral emotions mediate the effect of authentic transformational leadership on follower authentic moral action. Furthermore, we explore how authentic transformational leadership develops group ethical climate, which in turn contributes to enhancing group ethics and to developing follower moral identity and moral emotions. Future research and practical implications are discussed.						
leadership	3005.8	transform	620.5	develop	286.9	cours	175.1
leader	1927.8	follow	518.3	teach	207.8	examin	175.1
student	896.8	univers	505.8	academ	187.8	moral	164.6
educ	805.3	school	501.3	graduat	181.1	authent	163.1
ethic	664.8	style	475.6	vision	176.7	faculti	160.8
Topic 07	52880.1						
Name	Commitment						
Description	Levels, characteristics and effects of employee commitment to organization and/or supervisors						
Abstract	We examined the relationships of affective organizational commitment and affective commitment to supervisors with turnover intentions and actual turnover, using three independent samples of employees. In Sample 1 (N = 172) and Sample 2 (N = 186), affective organizational commitment and affective commitment to supervisors were found to exert independent negative effects on turnover intentions. Moreover, in both samples, affective commitment to supervisors was more strongly related to turnover intentions when affective organizational commitment was low. In Sample 3 (N = 431), affective commitment to supervisors was the single significant predictor of actual turnover and interacted with affective organizational commitment such that its effect was stronger when affective organizational commitment was low. The implications of these findings for the understanding of the commitment-turnover relationship are discussed.						
employe	3885.9	satisfact	1318.6	turnov	701.5	examin	573.3
job	2659.1	organiz	1093.5	psycholog	672.3	supervisor	570.8
work	1699.9	support	848.6	posit	665.6	affect	539.9
relationship	1555.3	mediat	814.5	perceiv	660.8	climat	512.2
commit	1418.8	relat	794.4	intent	612.0	workplac	509.8
Topic 08	56698.1						
Name	Decision-making						
Description	Development and assessment of decision-making models						
Abstract	Many real-world decision problems involve conflicting systems of criteria, uncertainty and imprecise information. Some also involve a group of decision makers (DMs) where a reduction of different individual preferences on a given set to a single collective preference is required. Multi-criteria decision analysis (MCDA) is a widely used decision methodology that can improve the quality of group multiple criteria decisions by making the process more explicit, rational and efficient. One family of MCDA models uses what is known as "outranking relations" to rank a set of actions. The Electre method and its derivatives are prominent outranking methods in MCDA. In this study, we propose an alternative fuzzy outranking method by extending the Electre I method to take into account the uncertain, imprecise and linguistic assessments provided						

by a group of DMs. The contribution of this paper is fivefold: (1) we address the gap in the Electre literature for problems involving conflicting systems of criteria, uncertainty and imprecise information; (2) we extend the Electre I method to take into account the uncertain, imprecise and linguistic assessments; (3) we define outranking relations by pairwise comparisons and use decision graphs to determine which action is preferable, incomparable or indifferent in the fuzzy environment; (4) we show that contrary to the TOPSIS rankings, the Electre approach reveals more useful information including the incomparability among the actions; and (5) we provide a numerical example to elucidate the details of the proposed method.

model	5611.0	method	1220.6	framework	842.8	decisionmak	673.4
decis	3076.2	evalu	1159.4	assess	750.4	choic	662.6
process	2610.5	select	1042.8	analysi	750.0	appli	634.1
approach	1648.5	develop	962.0	base	732.1	provid	594.2
propos	1223.9	applic	937.3	methodolog	687.1	tool	582.6
Topic 09	33549.8						

Name Human Recourse Management (HRM)

Description HRM practices, their consequences in different contexts and under different conditions

Abstract

Purpose - The purpose of this paper is to explore the implications for HRM of employers' use of migrants in low-skilled work in a UK-based firm. Is the use of migrant workers for low skilled work associated with "soft" or "hard" approaches to HRM? How do employers recruit migrant workers? What career progression paths are available to these workers in firms? What are the expectations and aspirations of migrant workers? Design/methodology/approach - The paper examines these issues through a case study of a UK-based employer using large numbers of migrant workers. The paper draws on data from a survey of migrant workers in the firm conducted in 2006, and from interviews with managers and migrant workers within this firm, conducted between 2005 and 2006. Findings - The paper highlights the "hard" HRM strategy pursued by the company in order to maintain a competitive advantage based on low labour costs and substitutability of workers. A contradiction is noted between the desire of the firm to retain migrant workers with a strong work ethic and gain high commitment, on the one hand, and their continued attempt to compete on the basis on minimal labour costs and follow a "hard" approach to HRM, on the other. Practical implications - The paper points to the importance of analysis of employers' use of migrants and the strategies they are adopting towards using these workers. Developing an understanding of these strategies is critical to understanding the social and economic experiences of migrant workers. Originality/value - The paper combines qualitative and quantitative research through an intensive case study to illuminate the implications for HRM of employers' use of migrants in low-skilled jobs.

work	1822.7	worker	1062.7	skill	523.3	individu	310.0
human	1671.6	practic	960.3	manag	473.8	job	288.5
resourc	1667.0	employe	737.6	organ	399.2	workforc	287.3
career	1152.9	hrm	681.9	recruit	349.0	older	284.9
employ	1124.7	profession	632.0	capit	336.6	age	274.1

Topic 10 74729.0

Name Survey studies

Description Design and results of survey studies

Abstract

The paper presents the results of the research on certain aspects of PR function in Serbian companies. 70 PR managers were polled, and it is estimated that they represent a third of the total number of PR managers in Serbia. 29 questions were asked, and they were divided into three characteristic categories: a) Questions relevant to the analysis of the position of PR managers and their profession; b) Questions relevant to the analysis of characteristic profession and education of PR managers; c) Questions relevant to the analysis of the most frequent and most important activities and media for the performance of PR function. In addition to the presentation of the results of the survey, the paper also presents their detailed analysis.

factor	3500.5	survey	1443.9	characterist	919.4	determin	864.8
differ	2860.6	practic	1140.2	collect	902.6	level	805.7
data	2183.8	dimens	989.8	type	891.7	empir	790.2
purpos	2165.4	identifi	968.8	aim	891.2	import	787.1
analysi	1734.9	analys	928.4	sampl	881.1	signific	727.5

Topic 11 35998.4

Name Multinational business

Description Aspects and issues related to trans- and/or multinational business

Abstract

This study examines how the institutional distance between a host country and a home country influences foreign subsidiary staffing, and how overseas business experience moderates the effect of institutional distance. Hypotheses regarding the effect of institutional distance on foreign subsidiary staffing are empirically tested using a sample of 2,980 foreign subsidiaries of Japanese firms. This study shows that although the ratio of parent country nationals to subsidiary employees decreases when firms face greater institutional distance, the absolute number of parent country nationals assigned to the subsidiary increases. This study also shows that firms with more overseas business experience replace host country nationals with parent country nationals when there is greater institutional distance.

intern	1707.4	busi	832.1	countri	647.7	local	486.7
global	901.5	foreign	808.8	acquisit	623.5	economi	451.2
chines	853.7	ventur	803.8	multin	598.9	develop	403.7
china	852.6	entrepreneurship	748.7	market	564.1	expatri	390.9
entrepreneuri	849.3	entrepreneur	678.8	subsidiari	541.9	emerg	331.3

Topic 12 59576.1

Name Causal effects

Description Causal relations between management/business related variables

Abstract

Research on the behavioural intentions (BI) of business-to-business customers has focused on the influence and the interactions among constructs driving BI, giving rise to two perspectives for the structural equation modeling of constructs influencing EI. In the first perspective, BI is impacted directly by relationship quality (RQ) and its antecedents such as relationship benefits (RB) and relationship sacrifices (RS). The second perspective suggests that antecedent dimensions of RB and RS may act indirectly through the construct of relationship value (RV), which drives RQ but also directly influences BI. The current study was undertaken to resolve the differences between these two perspectives. The findings indicate that RQ has a strong direct influence on BI while RV has a weak direct influence but a significant indirect influence on EI through the RQ construct. In addition, the constructs of RB and RS influence RQ directly, as well as through the RV construct.

perform	8207.6	test	1434.5	model	1034.4	support	748.1
relationship	4261.0	moder	1172.7	relat	1025.2	signific	704.5
effect	4108.7	influenc	1147.6	hypothes	904.7	data	667.4
posit	2474.0	orient	1080.5	direct	840.0	empir	627.4
impact	1531.5	examin	1047.1	negat	824.0	mediat	582.0

Topic 13 25015.9

Name Game theory

Description Game theoretical applications to management

Abstract

In information security outsourcing, it is the norm that the outsourcing firms and the outsourcers (commonly called managed security service providers, MSSPs) need to coordinate their efforts for better security. Nevertheless, efforts are often private and thus both firms and MSSPs can suffer from double moral hazard. Furthermore, the double moral hazard problem in security outsourcing is complicated by the existence of strong externality and the multiclient nature of MSSP services. In this prescriptive research, we first show that the prevailing contract structure in security outsourcing, bilateral refund contract, cannot solve double moral hazard. Adding breach-contingent sunk cost or external payment cannot solve double moral hazard either. Furthermore, positive externality can worsen double moral hazard. We then propose a new contract structure termed multilateral contract and show that it can solve double moral hazard and induce first-best efforts from all contractual parties when an MSSP serves two or more client firms, regardless of the externality. Firm-side externality significantly affects how payments flow under a multilateral contract when a security breach happens. When the number of client firms for an MSSP increases, we show that the contingent payments under multilateral contracts for any security breach scenario can be easily calculated using an additive method, and thus are computationally simple to implement.

contract	1127.9	mechan	565.7	agreement	299.4	relat	212.4
motiv	958.3	incent	565.6	reward	297.8	cost	202.8
power	892.8	transact	452.7	parti	285.0	outcom	199.7
negoti	814.7	agent	425.6	econom	279.5	ration	192.3
cooper	742.5	effort	307.9	benefit	235.7	effect	183.8

Topic 14 21799.1

Name Structural inequality

Description Prevailing inequalities in organizations related to gender, race, ethnicity and age

Abstract

Many contemporary studies of 'work-life balance' either ignore gender or take it for granted. We conducted semi-structured interviews with men and women in mid-life (aged 50 to 52 years) in order to compare their experiences of work-life balance. Our data suggest that gender remains embedded in the ways that respondents negotiate home and work life. The women discussed their current problems juggling a variety of roles (despite having no young children at home), while men confined their discussion of such conflicts to the past, when their children were young. However, diversity among men (some of whom 'worked to live' while others 'lived to work') and women (some of whom constructed themselves in relation to their families, while others positioned themselves as 'independent women') was apparent, as were some commonalities between men and women (both men and women constructed themselves as 'pragmatic workers'). We suggest ways in which gender-neutral theories of work-life balance may be extended.

women	909.0	age	359.0	coach	214.9	minor	161.2
gender	812.9	men	357.0	ethnic	201.9	profil	157.7
differ	618.4	individu	315.6	discrimin	184.8	stereotyp	153.9
status	406.6	male	270.9	generat	172.0	work	151.3
women	909.0	age	359.0	coach	214.9	minor	161.2

Topic 15 37646.5

Name Behavioral management

Description Behavioristic management methods and practices

Abstract

The current research reconciles two contradicting sets of findings on the role of cognitive control in socially desirable behaviors. One set of findings suggests that people are tempted by self-serving impulses and have to rely on cognitive control overriding such impulses to act in socially desirable ways. Another set of findings suggests people are guided by other-regarding impulses and cognitive control is not necessary to motivate socially desirable behaviors. We theorize that the dominant impulse is to behave in a socially desirable manner when the interpersonal impact of an action is salient, and that the dominant impulse is to behave in a self-serving manner when the interpersonal impact of an action is not salient. Studies 1-3 found that impairing participants' cognitive control led to less socially desirable behavior when interpersonal impact was not salient, but more socially desirable behavior when interpersonal impact was salient. Study 4 demonstrates that behaving in a socially desirable manner causes cognitive control impairment when interpersonal impact is not salient. But, when interpersonal impact is salient, behaving in a self-serving manner impairs cognitive control. We discuss the implications of our findings for understanding and managing socially desirable behaviors.

behavior	3040.9	cognit	775.0	toward	443.8	affect	376.4
person	1386.7	attitud	704.0	motiv	433.8	relat	371.4
individu	1285.3	social	671.4	belief	422.9	experi	368.7
emot	1118.8	predict	520.1	theori	415.1	negat	323.9
influenc	914.3	intent	504.5	posit	394.9	examin	321.0

Topic 16 36518.7



Name	Knowledge transfer						
Description	Conditions and methods for knowledge transfer between firms, academia and industry and between countries						
Abstract	This research introduces a framework for selecting efficient interunit structures in facilitating the transfer of knowledge with different levels of complexity. We argue that while the boundary spanner structure is efficient for transferring discrete knowledge, it is inadequate for transferring collectively held complex knowledge. We propose that the transfer of such knowledge requires a more decentralized interunit structure collective bridge, which is a set of direct interunit ties connecting the members of the source and the recipient units, with the configuration of the interunit ties matching the complexity of knowledge to be transferred. We suggest that while a collective bridge is inefficient in transferring discrete knowledge relative to a boundary spanner structure, it is more efficient for transferring collective knowledge.						
knowledg	6046.6	univers	625.7	creation	388.9	scientist	277.7
network	3293.4	capit	563.6	structur	362.5	properti	222.4
transfer	1131.3	tie	453.4	technolog	344.6	scienc	217.3
patent	911.9	academ	422.1	intellectu	305.1	share	213.9
social	825.4	activ	398.4	sourc	279.1	scientif	211.0
Topic 17	29321.8						
Name	Corporate governance						
Description	Corporate governance design and effects on board, CEO and firm performance						
Abstract	Each year shareholders, via exercise of their proxy votes, have the opportunity to voice their support or displeasure with firms and director nominees. Examining over 2,000 Fortune 500 director nominees, we explore those indicators available to shareholders at the time of directors' (re) election to provide insight into shareholder discontent with director monitoring. By studying actual voting behaviors, we provide new perspective to understanding director elections as a governance process. Employing a multilevel approach, we find support for agency-theoretic relationships between several firm and director characteristics and shareholder opposition to directors seeking (re) election to the board. At the firm level, we find that CEO compensation level and board size are positively related to the withholding of shareholder votes in director elections, a behavior indicative of shareholder discontent. Complementing these findings, at the director level, we find that affiliated director status, tenure, and number of outside directorships are positively related, and director block ownership is negatively related to shareholder discontent with director monitoring.						
govern	1906.4	ceo	642.8	director	439.3	audit	291.9
control	1259.1	ownership	563.1	execut	404.5	institut	253.6
corpor	1071.4	agenc	558.5	reput	363.4	implic	247.3
board	769.9	pay	495.7	sharehold	327.0	structur	241.9
firm	754.2	regul	475.6	compens	302.5	monitor	225.7
Topic 18	43016.9						
Name	Project management						
Description	Efficient management of projects, frameworks for project management						
Abstract	In this article, the process of developing the project plan and the project planning and control (PP&C) is analyzed and compared to the process of developing the project team to identify opportunities for integrating these actions to produce more successful projects. Results of structured research across some 137 different organizations and representing a wide range of approaches to establishing projects are reported. The results of this cross-organizational research strongly support the proposed integration of project planning, PP&C development, and project team building. A recommended process for accomplishing this integration is proposed.						
project	4260.7	success	1163.8	process	452.0	engin	412.5
system	4239.7	manag	1135.4	program	445.1	phase	376.5
design	1503.0	complex	817.6	support	432.9	construct	362.7
develop	1494.9	softwar	695.7	requir	417.0	integr	309.4
implement	1395.2	plan	485.3	improv	413.3	case	309.4
Topic 19	37891.3						
Name	Consumer economics						
Description	Consumer market assessment using economics						
Abstract	This paper shows that the effect of different distribution channel structures on product quality depends on the type of consumer heterogeneity and its distribution in a market. When consumer heterogeneity is uniformly distributed either vertically on willingness to pay or horizontally on transaction costs, a manufacturer may provide the same or lower product quality in a decentralized channel than in a centralized channel. In contrast, when consumer heterogeneity follows a more general distribution on willingness to pay, under certain conditions, the manufacturer may provide higher product quality in a decentralized channel than in a centralized channel. Decentralization also may lead to a higher product quality if consumer heterogeneity is uniformly distributed both vertically and horizontally, but not if consumer heterogeneity is uniformly distributed vertically on each of two product-quality attributes. Additionally, competition at the retail level may amplify these findings.						
price	2268.8	sale	796.7	auction	392.9	equilibrium	361.9
market	1584.7	consum	598.2	competit	391.8	strategi	329.8
product	1411.1	game	574.5	purchas	375.6	offer	325.0
retail	1063.1	demand	468.2	cost	367.0	increas	311.8
profit	1017.0	model	430.2	revenu	364.1	store	306.4
Topic 20	47696.9						
Name	Operations algorithms						
Description	Development and evaluation of algorithms and programming approaches for operation management						
Abstract	The problem of equipment selection for a production line is considered. Each piece of equipment, also called unit or block, performs a set of operations. All necessary operations of the line and all available blocks with their costs are known. The difficulty is to choose the most appropriate blocks and group them into (work)stations. There are some constraints that restrict the assignment of different blocks to the same						

station. Two combinatorial approaches for solving this problem are suggested. Both are based on a novel concept of locally feasible stations. The first approach combinatorially enumerates all feasible solutions, and the second reduces the problem to search for a maximum weight clique. A boolean linear program based on a set packing formulation is presented. Computer experiments with benchmark data are described. Their results show that the set packing model is competitive and can be used to solve real-life problems.

problem	2707.2	comput	700.1	propos	540.5	present	356.0
optim	1271.5	program	631.8	approach	534.8	stochast	346.3
solut	1142.6	solv	629.8	method	498.0	bound	320.7
effici	1075.8	algorithm	593.9	heurist	404.3	consid	314.8
function	786.9	set	587.5	rule	367.4	formul	307.6
Topic 21	48682.1						
Name	Crisis management						
Description	Strategies to mitigate risks and to respond accurately to crises						
Abstract	Both academics and practitioners have recently discovered resilience as a core topic of interest. Resilience is widely viewed as a potential solution to the challenges posed by crises and disasters. The promise of resilience is an organization or society that absorbs shocks and bounces back' after a disturbance. While the idea of resilience is increasingly popular, empirical research on resilient organizations is actually quite rare. This article explores whether a relation exists between organizational characteristics, processes and resilience. Building on the insights of high reliability theory and crisis research, it probes this relation in two organizations that experienced deep crises: the California Independent System Operator (CAISO) and National Aeronautics and Space Agency (NASA).						
challeng	571.1	question	385.7	problem	339.4	deal	256.1
failur	536.5	help	367.6	caus	337.6	success	255.5
peopl	485.4	face	364.5	world	337.2	potenti	243.3
crisi	469.3	situat	357.3	time	305.7	much	240.8
need	442.7	respons	340.0	disast	282.9	reason	233.0
Topic 22	27419.1						
Name	Service operations						
Description	Design, methods and management of service operations						
Abstract	Although the quality of the customer contact centre is pivotal for services, a thorough conceptualization and operationalization of perceived customer contact centre quality does not exist. The extensive scale development process moves from focus group sessions for item generation to exploratory and confirmatory factor analyses. Perceived customer contact centre quality consists of seven dimensions: reliability, empathy, customer knowledge, customer focus, waiting cost, user friendliness of the voice response unit, and accessibility. Compared with existing conceptualizations of service quality, perceived customer contact centre quality offers additional dimensions.						
servic	5346.4	valu	562.5	improv	216.0	model	176.9
custom	3810.5	loyalti	341.3	recoveri	213.1	relationship	175.0
qualiti	2770.3	deliveri	255.1	perceiv	186.1	cocreat	163.8
satisfact	703.6	offer	234.7	industri	182.6	market	155.8
provid	699.4	restaur	232.0	import	179.4	need	148.1
Topic 23	49332.4						
Name	Predictive modeling						
Description	Development and evaluation of predictive models for management						
Abstract	This paper investigates the forecasting ability of four different GARCH models and the Kalman filter method. The four GARCH models applied are the bivariate GARCH, BEKK GARCH, GARCH-GJR and the GARCH-X model. The paper also compares the forecasting ability of the non-GARCH model: the Kalman method. Forecast errors based on 20 UK company daily stock return (based on estimated time-varying beta) forecasts are employed to evaluate out-of-sample forecasting ability of both GARCH models and Kalman method. Measures of forecast errors overwhelmingly support the Kalman filter approach. Among the GARCH models the GJR model appears to provide somewhat more accurate forecasts than the other bivariate GARCH models						
measur	2318.6	valid	860.8	perform	574.6	compar	397.4
model	2062.6	predict	827.2	variabl	484.7	error	396.0
forecast	1449.9	method	811.9	score	432.6	indic	381.3
estim	982.5	data	801.7	construct	409.3	statist	369.7
test	966.8	scale	747.4	reliabl	406.0	rate	361.8
Topic 24	83066.3						
Name	Literature reviews						
Description	Assessment and description of previous literature on management						
Abstract	This paper provides a comprehensive literature review of the phenomenon of spinouts from academic institutions. We systematically identified spinout papers in key management journals, categorised the literature and critically synthesised the findings. We present the findings of each literature stream in turn and also identify inconsistencies and directions for further research. We conclude that while the early literature has been mainly atheoretical and focused on describing the phenomenon, a core group of recent studies were theory-driven.						
literatur	2214.3	articl	1246.6	theoret	1132.6	perspect	904.0
theori	1567.2	framework	1191.1	practic	1088.5	conceptu	875.1
review	1558.9	issu	1175.9	concept	1069.6	contribut	758.2
develop	1461.8	field	1156.5	provid	1064.1	empir	752.9
discuss	1295.1	futur	1147.5	manag	1021.6	understand	739.2
Topic 25	38215.9						
Name	Inventory management						
Description	Methods and strategies for efficient inventory management						

## Abstract

We use exponential lead times to demonstrate that reducing mean lead time has a secondary reduction of the variance due to order crossover. The net effect is that of reducing the inventory cost, and if the reduction in inventory cost overrides the investment in lead time reduction, then the lead time reduction strategy would be tenable. We define lead time reduction as the process of decreasing lead time at an increased cost. To date, decreasing lead times has been confined to deterministic instances. We examine the case where lead times are exponential, for when lead times are stochastic, deliveries are subject to order crossover, so that we must consider effective lead times rather than the actual lead times. The result is that the variance of these lead times is less than the variance of the original replenishment lead times. Here we present a two-stage procedure for reducing the mean and variance for exponentially distributed lead times. We assume that the lead time is made of one or several components and is the time between when the need of a replenishment order is determined to the time of receipt.

cost	2223.9	oper	595.1	order	449.7	reduc	298.0
time	1311.1	inventori	586.4	plan	415.8	schedul	295.0
demand	1126.9	system	584.9	transport	414.8	increas	291.2
capac	825.9	model	485.0	level	392.7	delay	289.7
polici	729.0	optim	457.2	alloc	363.7	facil	277.7

Topic 26 28826.7

Name Corporate social responsibility (CSR)

Description Firms' CSR considerations, external demands/expectations of CSR and CSR management

## Abstract

This paper analyses how large Danish companies are responding to new governmental regulation which requires them to report on corporate social responsibility (CSR). The paper is based on an analysis of 142 company annual reports required by the new Danish regulation regarding CSR reporting, plus 10 interviews with first-time reporting companies and six interviews with companies that failed to comply with the new law. It is concluded that coercive pressures from government have an impact on CSR reporting practices. Further, the analysis finds traces of mimetic isomorphism which inspires a homogenisation in CSR reporting practices. Finally, it is argued that non-conformance with the new regulatory requirements is not solely about conscious resistance but may also be caused by, for example, lack of awareness, resource limitations, misinterpretations, and practical difficulties.

environment	1577.7	small	802.3	csr	506.0	adopt	282.1
sustain	1422.9	smes	750.9	compani	390.2	pressur	279.9
social	1026.6	corpor	723.2	report	372.1	mediums	273.7
respons	951.2	enterpris	685.8	activ	314.6	green	263.9
stakehold	806.3	standard	579.1	develop	299.4	issu	257.5

Topic 27 26858.9

Name Tourism management

Description Tourism and event management, destination area development

## Abstract

Destinations use sport events to attract participants and spectators, who then hold perceptions of both the sport event and destination. This research aimed to a) understand how active sport tourists perceive the meaning of a sport event experience and b) develop a scale for that meaning. Both aims are studied in a post trip context as evaluative research. Two focus groups were used to understand the meaning of the sport event experience among active sport tourists. Results from the focus groups suggest participants attribute meanings related to organizational, environmental, physical, social, and emotional aspects of the sport event experience. Next, semantic differential items were developed to measure the meaning of a sport event experience in the post trip phase. The items were tested with two different sport event participant samples using surveys. A uni-dimensional scale of 11 semantic differential items emerged. These items provide a measure for the evaluative meaning of a sport event experience.

organis	2016.6	sport	630.9	citi	293.5	activ	226.2
communiti	1077.3	hotel	548.1	particip	268.6	develop	221.0
behaviour	777.4	travel	429.0	area	237.7	interview	214.1
tourism	730.0	tourist	392.0	experi	233.4	australian	187.4
event	640.8	destin	345.3	centr	228.6	impact	186.4

Topic 28 29685.7

Name Online marketing

Description Issues , practices and strategies of online marketing

## Abstract

Although online retailers detail their privacy practices in online privacy policies, this information often remains invisible to consumers, who seldom make the effort to read and understand those policies. This paper reports on research undertaken to determine whether a more prominent display of privacy information will cause consumers to incorporate privacy considerations into their online purchasing decisions. We designed an experiment in which a shopping search engine interface clearly and compactly displays privacy policy information. When such information is made available, consumers tend to purchase from online retailers who better protect their privacy. In fact, our study indicates that when privacy information is made more salient and accessible, some consumers are willing to pay a premium to purchase from privacy protective websites. This result suggests that businesses may be able to leverage privacy protection as a selling point.

user	1055.7	search	776.4	market	378.7	site	273.3
onlin	948.6	internet	458.9	media	364.1	data	264.8
consum	912.6	technolog	421.6	content	361.1	electron	257.0
inform	829.5	mobil	397.2	channel	332.6	usag	248.2
brand	818.9	web	389.9	adopt	293.6	increas	247.5

Topic 29 28427.2

Name Supply chain management

Description Integration and management of supply and value chains

## Abstract

Purpose - The purpose of this paper is to empirically explore supply chain integration (SCI) enabling practices, their benefits and barriers in a retail product returns process context. Design/methodology/approach - The study adopts a case study research strategy. It draws on a single

case, comprised of an original equipment manufacturer (OEM) and its two retailers. It utilizes an in-depth semi structured interviewing approach, combined with walk-through observations. Findings - The study finds that management of retail product returns can significantly benefit both an OEM and its customers when appropriate SCI enabling practices are deployed. While these practices are similar to those in forward supply chain processes, barriers are driven by the characteristics of product returns processes. Research limitations/implications - The limitations of this study stem primarily from its methodological design. A single case research strategy provides a limited opportunity for external generalization of the research findings. Practical implications - This study illustrates the value of SCI initiatives in product returns processes and informs managers' decision making in the planning and execution of similar SCI implementations in product returns processes. Originality/value - This research claims to be one of the first works that systematically and empirically explores SCI in reverse supply chain processes, as opposed to forward supply chain processes.

suppli	2804.9	integr	552.2	improv	302.6	plant	254.0
chain	2581.9	product	533.9	purpos	290.8	sourc	240.5
supplier	1609.9	flexibl	474.3	case	275.7	buyer	236.8
manufactur	1477.0	industri	404.5	practic	273.8	lean	232.9
logist	797.9	oper	399.7	purchas	256.4	cost	222.2

Topic 30 43267.2

Name Socio-economic policies

Description Design of socio economic policies affecting business and industry conditions

Abstract

During the 1990's there was considerable debate in Australia about the desirability or otherwise of changing the nation's official trade focus away from traditional trading partners in Western Europe and North America to Asian countries located within Australia's own East Asia/Pacific region. This paper analyses Australia's trade patterns to better understand whether the economic opportunities that have emerged with East Asia's growth have trumped the nation's close historical, cultural and political relationships with Western Europe and North America. An analysis of cultural differences and trade indicates that culture plays little if any part in Australia's national trade outcomes, and that Australia's international trade interests are much more closely aligned with East Asia than cultural argument might have predicted.

polic	1945.2	nation	1025.8	privat	625.7	economi	487.4
public	1857.6	institut	901.1	europ	540.0	industri	400.1
countri	1827.3	develop	811.4	state	526.5	differ	353.3
sector	1233.9	econom	652.2	govern	497.1	level	348.8
region	1137.5	polit	637.8	local	495.6	union	347.7

Topic 31 25131.0

Name Team management

Description Team management, team performance and team composition

Abstract

This study revisits the commonplace research conclusion that greater team member collectivism, as opposed to individualism, is associated with higher levels of individual-level performance in teams. Whereas this conclusion is based on the assumption that work in teams consists exclusively of tasks that are shared, typical teamwork also includes tasks that are individualized. Results of a laboratory study of 206 participants performing a mix of individualized and shared tasks in four-person teams indicate that heterogeneous combinations of individualism and collectivism are associated with higher levels of team member performance, measured as quantity of output, when loose structural interdependence enables individual differences in individualism/collectivism to exert meaningful effects. These results support the modified conclusion that a combination of individualism and collectivism is associated with higher levels of member performance in teams under typical work conditions; that is, conditions in which the tasks of individual members are both individualized and shared.

team	3734.9	famili	704.4	interact	287.3	meet	198.6
member	1553.3	individu	396.5	perform	283.8	function	162.5
conflict	1440.4	work	388.1	virtual	267.4	time	157.7
task	1141.2	level	328.3	process	264.2	composit	153.3
divers	947.0	effect	306.6	group	223.8	identif	151.2

Topic 32 40601.2

Name Employee appraisals

Description Methods for evaluation of employees, effects on performance and commitment from various evaluation techniques

Abstract

Three studies were conducted investigating the effects of irrelevant anchors on performance judgments. Both a lab and field study demonstrated that an alternative anchoring manipulation that did not involve an explicit comparative question had effects on performance judgments similar to a traditional anchoring manipulation. The final study examined whether the anchoring effects were more likely when the anchor was highly applicable to the final judgment. The results indicated that both highly applicable and low applicable anchors produced an anchoring effect, but the highly applicable anchors had a larger effect on performance judgments. Evidence was also found for asymmetrical anchoring effects. In two of the three studies, high anchors increased performance judgments relative to the control group, whereas low anchors were not significantly different from the control group.

effect	1270.0	goal	834.3	fair	517.3	level	422.7
particip	1183.2	attribut	743.3	low	482.1	expect	417.8
percept	1028.9	rate	730.2	evalu	473.8	fit	411.6
experi	989.2	high	706.2	feedback	455.6	bias	355.8
perceiv	859.6	prefer	547.3	outcom	435.3	procedur	349.4

Topic 33 72057.3

Name Constructionism

Description Organizations organization members and organizational practices as socially constructed

Abstract

This article aims to explore critically the role of an action research team in the social construction of interorganizational collaboration aimed at transgressing organizational and professional boundaries. We argue that the new relationships, actor conceptions and in some cases forms of work organization arising from the change process have been socially constructed through the discursive interventions of the researchers. This

has largely occurred through informal interaction with and between the actors engaged in the development process. The action researcher, rather than being a neutral discursive gatekeeper in collaborative development projects, is an active constructor of the discourse shaping the collaboration. A case is presented showing how the researcher role is thus better seen as being an active boundary subject mediating across various professional and organizational perspectives rather than a passive boundary object. Accordingly, by focusing on the discursive role of active researchers as boundary subjects, we can reflect more critically on the roles we adopt in our intervention endeavours and their inevitably political nature.

chang	2816.3	action	849.6	context	715.5	actor	619.4
process	1346.8	theori	789.0	organiz	664.2	draw	604.6
social	1043.5	ident	782.2	work	656.0	role	602.6
practic	941.7	articl	738.3	argu	622.1	explor	602.2
organ	894.3	understand	735.2	dynam	620.2	within	577.1
Topic 34	31003.4						
Name	Health care management						
Description	Management aspects of health care						
Abstract	<p>Aims To compare staff opinions about standardized care plans and self-reported habits with regard to documentation, and their perceived knowledge about the evidence-based guidelines in stroke care before and after implementation of an evidence-based-standardized care plan (EB-SCP) and quality standard for stroke care. The aim was also to describe staff opinions about, and their use of, the implemented EB-SCP. Background To facilitate evidence-based practice (EBP), a multi-professional EB-SCP and quality standard for stroke care was implemented in the electronic health record (EHR). Method Quantitative, descriptive and comparative, based on questionnaires completed before and after implementation. Results Perceived knowledge about evidence-based guidelines in stroke care increased after implementation of the EB-SCP. The majority agreed that the EB-SCP is useful and facilitates their work. There was no change between before and after implementation with regard to opinions about standardized care plans, self-reported documentation habits or time spent on documentation. Conclusions An evidence-based SCP seems to be useful in patient care and improves perceived knowledge about evidence-based guidelines in stroke care. Implications for nursing management For nursing managers, introduction of evidence-based SCP in the EHR may improve the prerequisites for promoting high-quality EBP in multi-professional care.</p>						
nurs	2310.0	patient	650.0	work	376.5	implic	339.0
care	1164.7	aim	467.7	programm	374.7	profession	337.2
health	966.1	staff	416.7	conclus	355.9	clinic	327.0
manag	888.0	background	397.9	improv	353.7	medic	324.8
hospit	825.8	healthcar	376.9	method	343.8	need	317.0
Topic 35	32597.0						
Name	Corporate finance						
Description	Firms' financing, risk/return optimization and investments						
Abstract	<p>This paper develops empirical proxy measures of information technology (IT) risk and incorporates them into the usual empirical models for analyzing IT returns: production function and market value specifications. The results suggest that IT capital investments make a substantially larger contribution to overall firm risk than non-IT capital investments. Further, firms with higher IT risk have a higher marginal product of IT relative to firms with low IT risk. In the market value specification, the impact of IT risk is positive and significant, and inclusion of the IT risk term substantially reduces the coefficient on IT capital. We estimate that about 30% of the gross return on IT investment corresponds to the risk premium associated with IT risk. Taken together, our results show that IT risk provides part of the explanation for the unusually high valuations of IT capital investment in recent research.</p>						
risk	2180.8	capit	886.4	asset	499.9	real	361.6
valu	2140.7	return	854.5	investor	415.0	flow	333.3
invest	1527.5	portfolio	846.0	uncertainti	395.1	inform	322.8
financi	1473.5	stock	680.9	trade	393.5	financ	304.0
market	1226.0	option	579.0	fund	390.6	equiti	278.0

## APPENDIX S4: Robustness and quality estimates for the correspondence analysis

**Table S4.1.** Summary statistics for the correspondence analysis

Dimensions	1	2	3	4
Eigenvalue	0.039	0.004	0.001	0.000
Inertia (%)	87.942	8.918	2.384	0.756
Cumulative %	87.942	96.860	99.244	100.000

**Table S4.2.** Quality of the model

Trace	0.045
Correlation coefficient for active variables (topics and gender categories)	0.211
Average inertia per dimension	25%
Average rule for inclusion of dimensions	
No. of dimensions with larger inertia than average:	1

**Table S4.3.** Chi-square test of significance of interdependence between variables (i.e. topics and gender categories)

Chi-square (Observed value)	1237.801
Chi-square (Critical value, at $\alpha=0.01$ )	181.840
Degrees of freedom	140
p-value	< 0.0001

**Table S4.4.** Permutation test of significance of interdependence between variables (i.e. topics and gender categories)

Observed total inertia	break-off value for the 999th percentile of the permuted total inertia	p-value
0.045	0.007	< 0.0001

Note: Since the observed inertia is larger than the 999th percentile, the observed inertia is significant at (at least) alpha 0.001

**Table S4.5.** Permutation test of significance of dimensions

Dimension	Observed inertia	break-off value for the 99th percentile of the permuted distribution	p-value
1	0.0381	0.0024	< 0.0001
2	0.0041	0.0018	< 0.0001
3	0.0008	0.0014	0.453
4	0.0003	0.0011	0.999

**Table S4.6.** Malinvaud's test of significance of dimensions

Dimension	Eigen value	Chi-square	Degrees of freedom	p-value
1	0.039	1237.801	140	< 0.0001
2	0.004	149.260	102	0.002
3	0.001	38.868	66	0.997
4	0.000	9.362	32	1.000

**Table S4.7.** Contribution biplots of each gender category

Gender categories	Score in D1	Sqr. cos. D1.	Contribution D1	Marginal value*
Female(d)	0.461	0.955	0.558	2858
Female(m)	0.260	0.819	0.118	1910
Male(d)	-0.155	0.934	0.283	12852
Male(m)	-0.048	0.204	0.010	4853
Mixed	0.079	0.544	0.030	5205

\*Variable marginal values equal the corresponding total loading in the topic model.

**Table S4.8.** Contribution biplots for each topic

Topics	Score in D1	Sqr. cos. D1	Contribution D1	Marginal value*
Innovation	-0.094	0.968	0.006	740.448
Dynamic capabilities	-0.137	0.985	0.016	946.926
Organizational learning	0.140	0.974	0.010	542.902
Time	-0.140	0.946	0.015	856.128
IORs	0.019	0.159	0.000	560.839
Strategic mgmt	-0.009	0.078	0.000	962.202
Leadership	0.222	0.992	0.019	424.509
Commitment	0.279	0.803	0.072	1006.207
Decision-making	-0.134	0.949	0.018	1054.605
HRM	0.347	0.999	0.067	605.324
Survey studies	0.051	0.399	0.003	1322.190
Multinational biz	-0.021	0.282	0.000	721.849
Causal effects	-0.064	0.276	0.004	1132.815
Game theory	-0.051	0.366	0.001	482.839
Structural inequality	0.439	0.961	0.074	417.246
Behavioral mgmt	0.138	0.608	0.013	739.791
Knowledge transfer	-0.016	0.071	0.000	681.591
Corporate governance	-0.099	0.937	0.005	526.120
Project mgmt	-0.085	0.822	0.005	792.956
Consumer economics	-0.268	0.971	0.046	690.067
Operations algorithms	-0.318	0.953	0.090	972.005
Crisis mgmt	-0.039	0.142	0.001	928.694
Service operations	-0.033	0.164	0.000	474.821
Predictive modeling	-0.249	0.963	0.057	997.012
Literature reviews	-0.038	0.514	0.002	1660.703
Inventory mgmt	-0.260	0.960	0.043	693.949
CSR	0.095	0.831	0.004	535.645
Tourism mgmt	0.146	0.777	0.010	508.877
Online marketing	-0.033	0.321	0.001	540.423
Supply chain mgmt	-0.202	0.911	0.018	470.988
Socio-economic policies	-0.013	0.340	0.000	849.228
Team mgmt	0.207	0.896	0.018	454.376

Employee appraisals	0.103	0.551	0.008	777.555
Constructionism	0.181	0.649	0.045	1492.814
Health care mgmt	0.816	0.940	0.304	498.286
Corporate finance	-0.204	0.979	0.024	615.070

\*Variable marginal values equals the corresponding total loading in the topic model.



## **APPENDIX S5: Inspecting the validity of the topic model using co-citation networks and co-word mapping**

The objective of this Appendix is to inspect how the thematic structure identified in our topic solution corresponds with the outcomes of alternative techniques for mapping the cognitive content of the management literature. Further, we seek to verify the plausibility of our interpretations of the 36 topics in the model. Specifically, we inspect differences and similarities between three techniques: topic modeling, co-citation analysis and co-word mapping. The comparison of the latent themes extracted from each of these techniques allows us to make more robust and transparent conclusions on the strengths and weaknesses of the selected topic solution and the plausibility of our interpretation of the model results. Researchers have already documented how a combination of different techniques (e.g. co-word and co-citation analysis) can assist the researcher in obtaining a clearer view of the cognitive structure of a given scientific discipline (Braam et al., 1991a, 1991b).

The appendix will be divided into three sections. First, we briefly describe the two alternative techniques employed in the comparison: co-citation networks and co-word mapping. Second, we present the outcomes of the co-citation and co-word analysis. Third, we highlight overlaps and variations in the outcomes of the three techniques, and use this information to reflect on the overall robustness of the topic model and the plausibility of our interpretation of model results.

### **1. Alternative techniques for mapping cognitive structures of scientific texts**

#### ***1.1. Co-citation analysis***

Co-citation analysis, unlike top-modeling, takes authors or documents as its basic unit of analysis. The technique establishes relationships between documents or authors based on citing documents. Schneider and Borlund (2004, p. 536) summarizes two key assumptions of co-citation analysis:

- (1) When two documents [or authors] are cited together by a third document, then a cognitive relationship exists between them.
- (2) The strength of this relationship is proportional to the frequency of the co-citation linkage, i.e. the number of documents that co-cite the two documents [or authors].

Put differently, if documents or authors are frequently cited together by other documents, a similarity in content is assumed to exist between them. Specified co-citation thresholds can be used to construct clusters of related documents/authors in large corpora of scientific texts. Further, the relationships and proximity/distance between these clusters can be visually displayed in a co-citation network. Clusters are here conceived as topics or specialties (Schneider and Borlund, 2004). Since co-citation analysis relies on references to prior literature, its clustering networks may not accurately depict the current state of the field. Further the technique has a bias towards established scientific frontiers with high citation-rates and is less reliable for clustering niche specialties comprised by documents/authors with relatively low citation rates (Zupic and Čater, 2015).

In this study we focus on author co-citations. The basic assumption underlying this technique is that references for authors that frequently co-occur in citing documents will share a similarity in terms of content (e.g. research topic or area of focus).

To identify author based co-citation clusters, we employ VOSviewer's science mapping framework. Using co-citation techniques to analyze large data-sets in VosViewer requires huge computer power. Hence, for this validation exercise we have limited our sample to management articles published within the period 2012-2013. We set the citation threshold value to 40, meaning that a researcher should have received at minimum 40 citations in the given period to be included in our analysis. 2180 cited authors meet this criteria.

The user can adjust several parameters of the co-citation network to construct a more or less fine grained cluster solutions in VosViewer (for specifications on VosViewers clustering method, see Waltman et al., 2010). In this study we aim for a cluster resolution corresponding with the number of topics emerging in our topic model. We have decided on a solution with 38 clusters (specifications on the parameter setting for our solution are provided in section 2).

## ***1.2. Co-word analysis***

Like topic modeling, co-word analysis is a useful quantitative technique for analyzing the content of large corpora of scientific text. Co-word analysis establishes links between concepts and ideas of a given scientific discipline or domain based on co-occurrences of pairs of words derived from paper-titles, key-word lists, abstracts or full texts. Words and noun-phrases that frequently co-occur across documents are conceived as closely related and used to extract broader concepts and ideas (or topics) that represent the conceptual space of the selected domain. Semantic maps can be employed to visualize the proximity/distance and network relations between these concepts and ideas, hence providing an overview of the structure of the scientific frontiers of the management literature (He, 1999; Zupic and Čater, 2015).

For purposes of co-word analysis we make use of the VOSviewer science mapping framework's term-map solution (Van Eck and Waltman, 2011). VosViewer employs the Apache Open NLP toolkit and a linguistic filter to identify all noun phrases and adjectives that end with a noun in the selected corpora of text (in our case scientific abstracts). The user is asked to select a term-occurrence threshold specifying how frequent a term (i.e. a noun-phrase) should occur in the text corpora to be considered eligible for mapping. VosViewer's threshold is set to 10 occurrences by default, but due to the large size of our text corpora, we apply a term-occurrence threshold of 40 resulting in 2127 relevant terms. Further, VosViewer includes a technique for selecting the most relevant of these terms. Van Eck and Waltman (2011, p. 2) describes the logic behind this technique thus:

Intuitively, the idea is that noun phrases with a low relevance (or noun phrases with a general meaning) (.....) have a more or less equal distribution of their (second-order) co-occurrences. On the other hand, noun phrases with a high relevance (or noun phrases with a specific meaning) (...) have a distribution of their (second-order) co-occurrences that is significantly biased towards certain other noun phrases.

VosViewer's relevance score is set to 60% by default. This means that the 60% percent most relevant terms will be included in the semantic map and cluster analysis. In this study we apply a relevance score of 50%, meaning that only the 1064 most relevant terms will be included in our map. General and irrelevant nouns (e.g. Elsevier", "John Wiley & Sons", "key contribution", "existing literature", "conclusion", "implications for practice", "research design") can be manually removed by the researcher to reduce noise in the subsequent processing of co-occurrences over noun-phrases. In this study we ended up removing 146 terms. Our final list of nouns consists of 947 terms.

Next, VosViewer groups frequently co-occurring terms into clusters (or topics) and provides a semantic map visualizing the clustering of results and the proximity/distance between clusters (for further specifications see Van Eck and Waltman 2011). The user can adjust several parameters of the semantic map to construct a more or less fine grained cluster solution. In this study we aim for a cluster-solution corresponding with the number of topics emerging in our topic model. We have decided on a solution with 36 clusters (specifications on the parameter setting for our solution are provided in section 2).

An important distinction should be made between the LDA topic modeling technique and the co-word analysis. In the LDA topic model, words are treated as multinomial distributions over topics, meaning that a given term can enter into several topics in the same model. Co-word analysis, in contrast, typically attributes each word to a single cluster (Leydesdorff and Nerghe, 2016). As we shall return to, this has implications for the outcomes of the two models and how we compare them, since the co-word analysis will be more likely to conflate thematic areas with overlapping terminology. Further, this means that the two techniques to some extent may rely on different terms to capture the same research frontiers.

## **2. Outcomes of the co-citation and co-word analysis**

### ***2.1. Co-citation cluster solution and network***

Figure A1 specifies the parameter settings for the co-citation network. The clustering resolution has been set to 2.60 and the minimum number of authors per cluster to 5, resulting in 38 clusters. Author groups range from 163 authors in Cluster 1 to 5 authors in Cluster 38.

**Figure A1.** Parameter settings for the co-citation network.

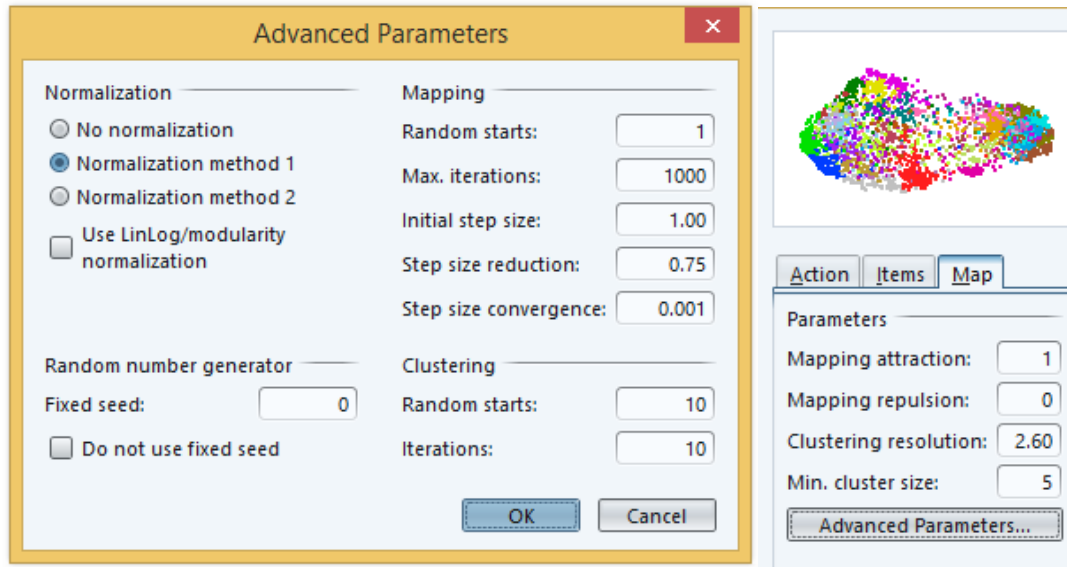
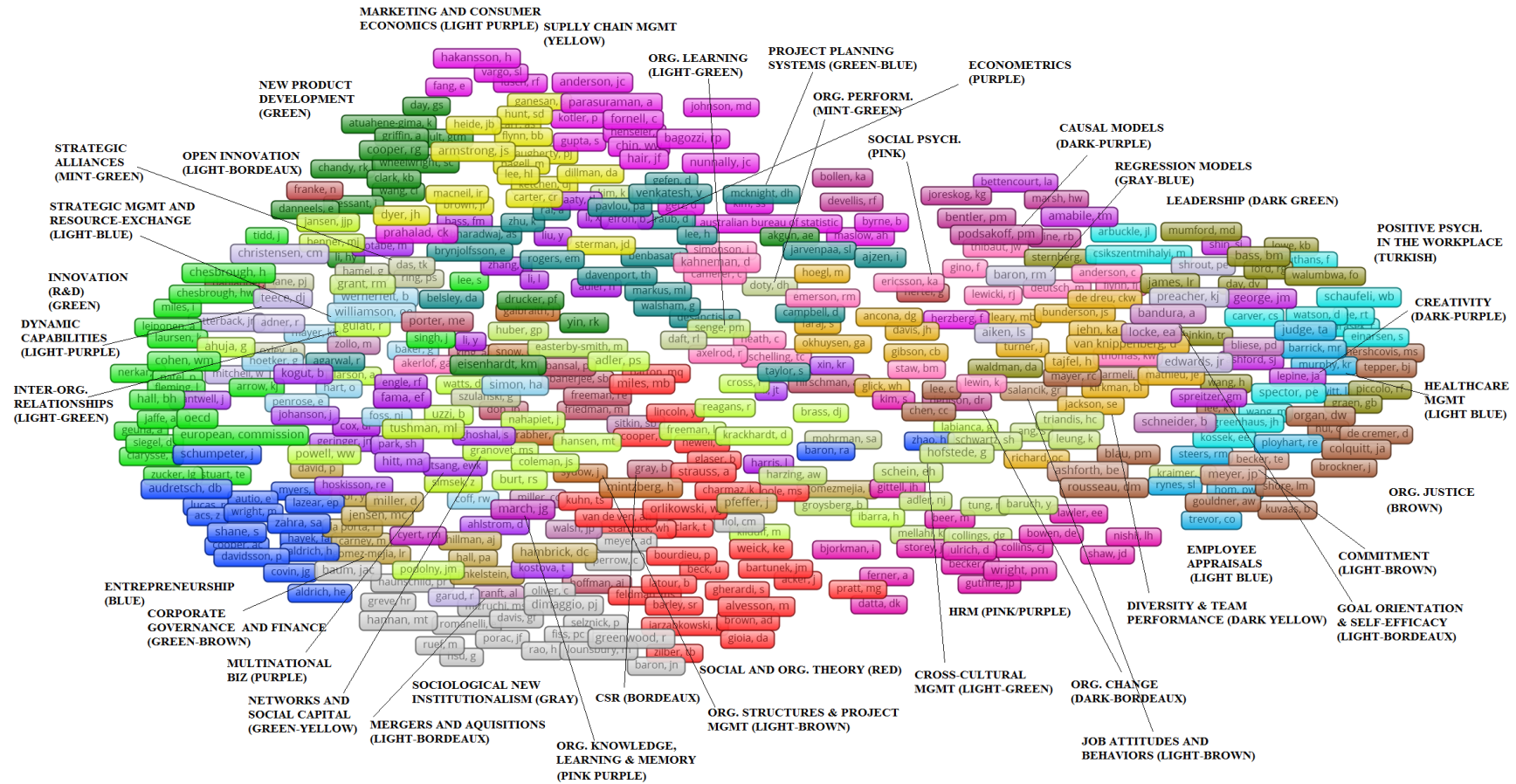


Figure A2 visualizes the co-citation network. Frames represent authors, and the distance/proximity between them, their relative co-citation frequencies. Frequently co-cited authors are closely located, while rarely co-cited authors are situated far from each other. Colors and shades are used to distinguish the 38 clusters in the network. Due to the many clusters, some clusters have fairly similar color-codings, but a search-function in the software helps us specify which authors pertain to each cluster. Cluster tags have been added manually for purposes of intelligibility.

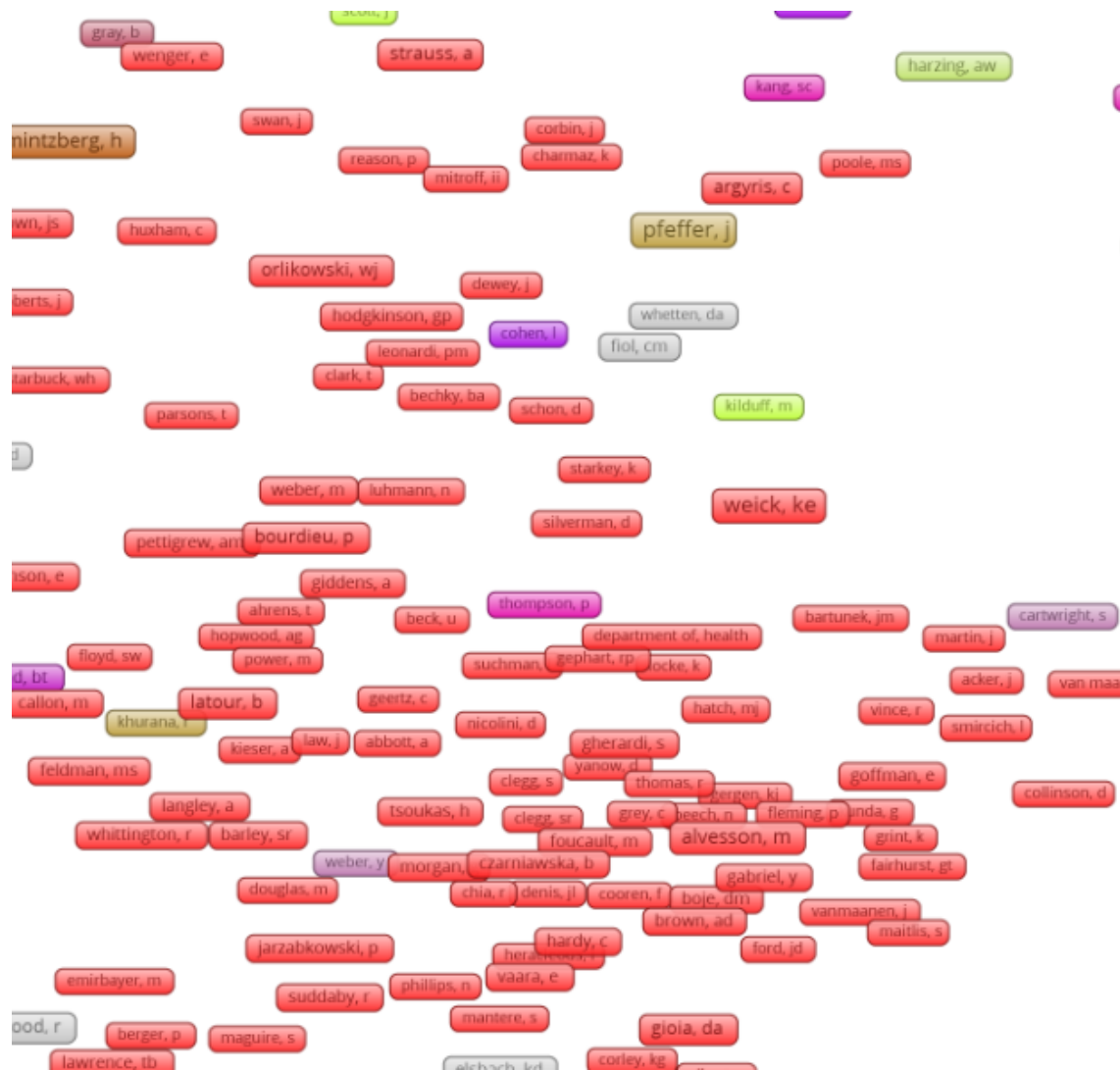
As mentioned earlier, clusters represent topics or research areas in the management literature. The shared content linking authors to particular clusters has been more obvious for some clusters than others. In cases of doubt, we have carefully familiarized ourselves with the focus areas of each author pertaining to a given cluster to detect the shared content explaining their connection. Yet, researchers can span over several research areas and the co-citation network only allow them to populate one cluster. This setup has made it particularly demanding to tease out content similarities between authors in sparsely populated clusters. Consequently, our interpretations of the shared content in the smallest clusters may in a few cases be too general.

A more fine-grained outline of the authors pertaining to each cluster is provided with screenshots below the co-citation network.

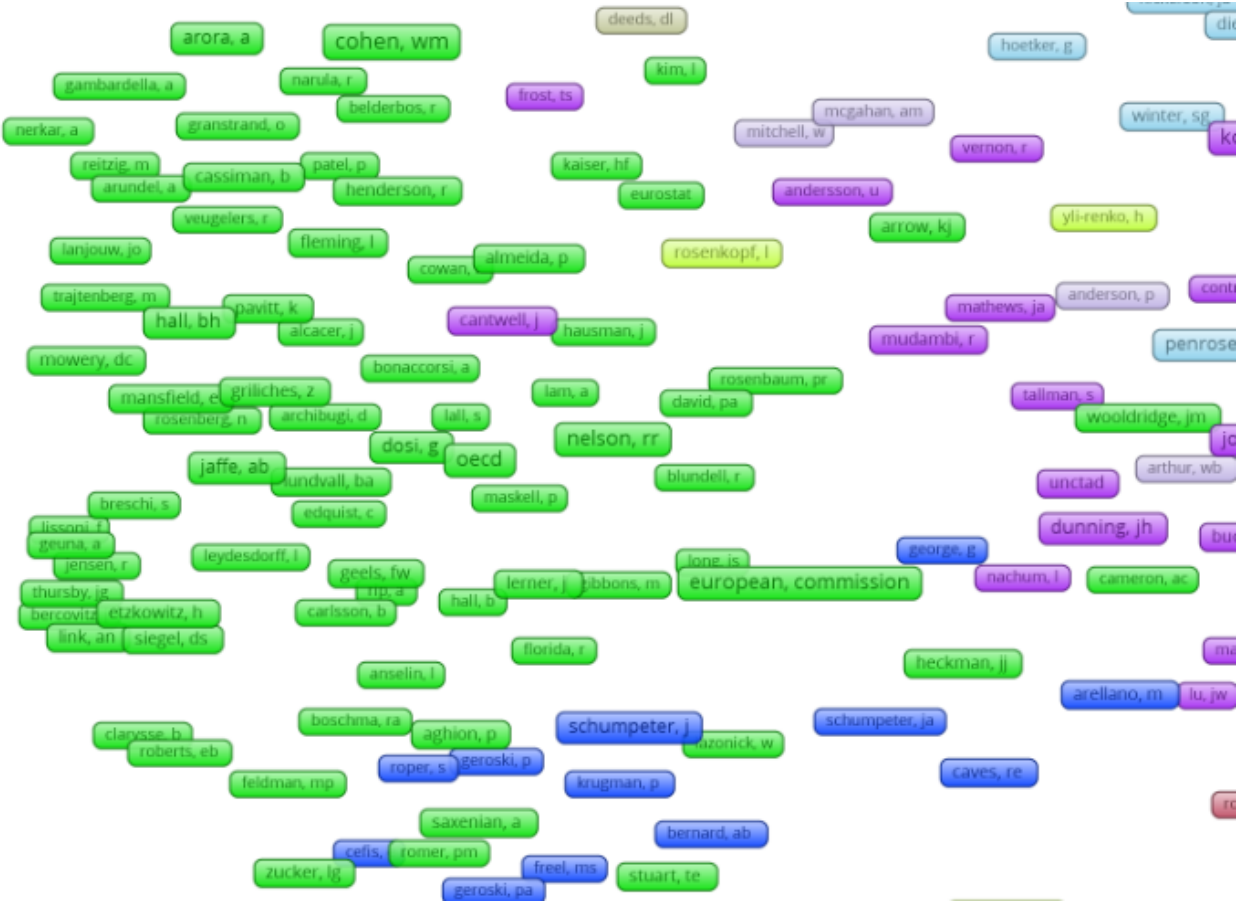
**Figure A2.** Co-citation network w. Cluster Tags (2180 authors, 38 Clusters)



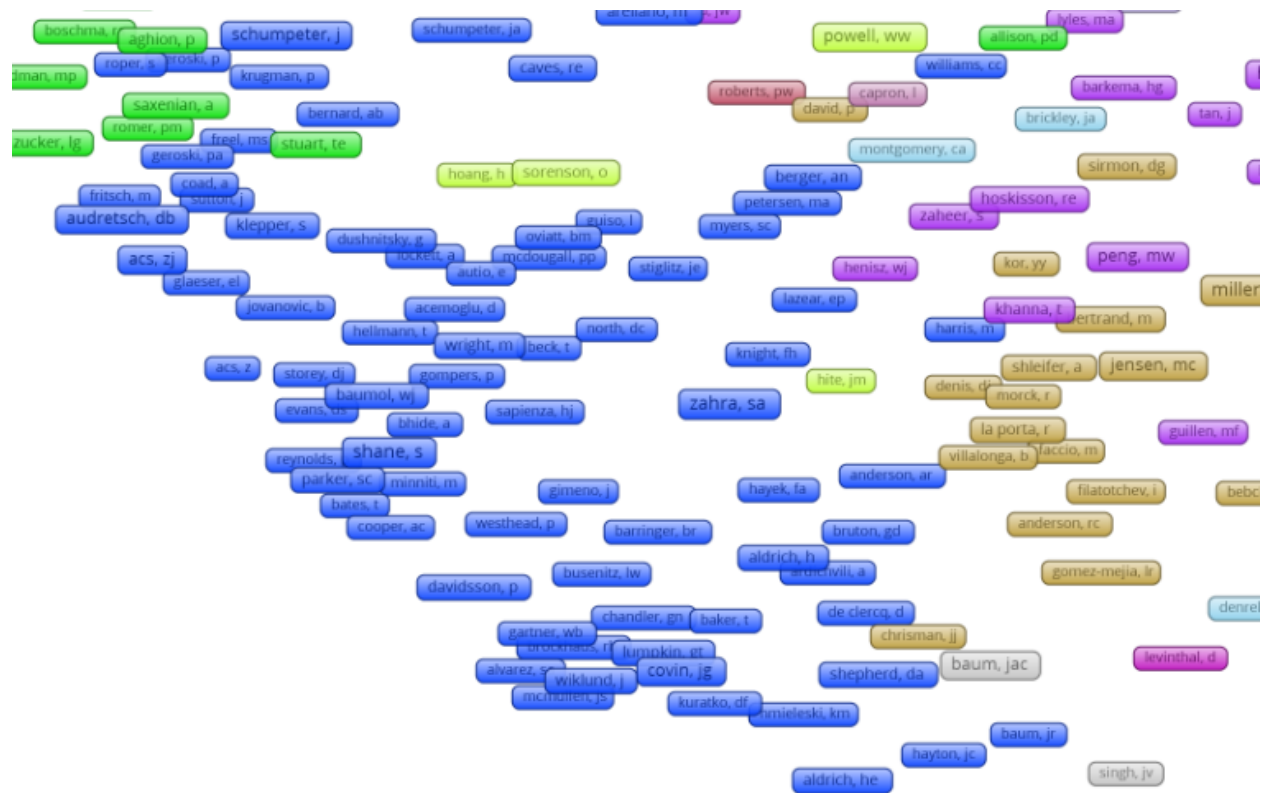
## 90



## 91



### Cluster 3. Entrepreneurship (Blue) (126 authors)

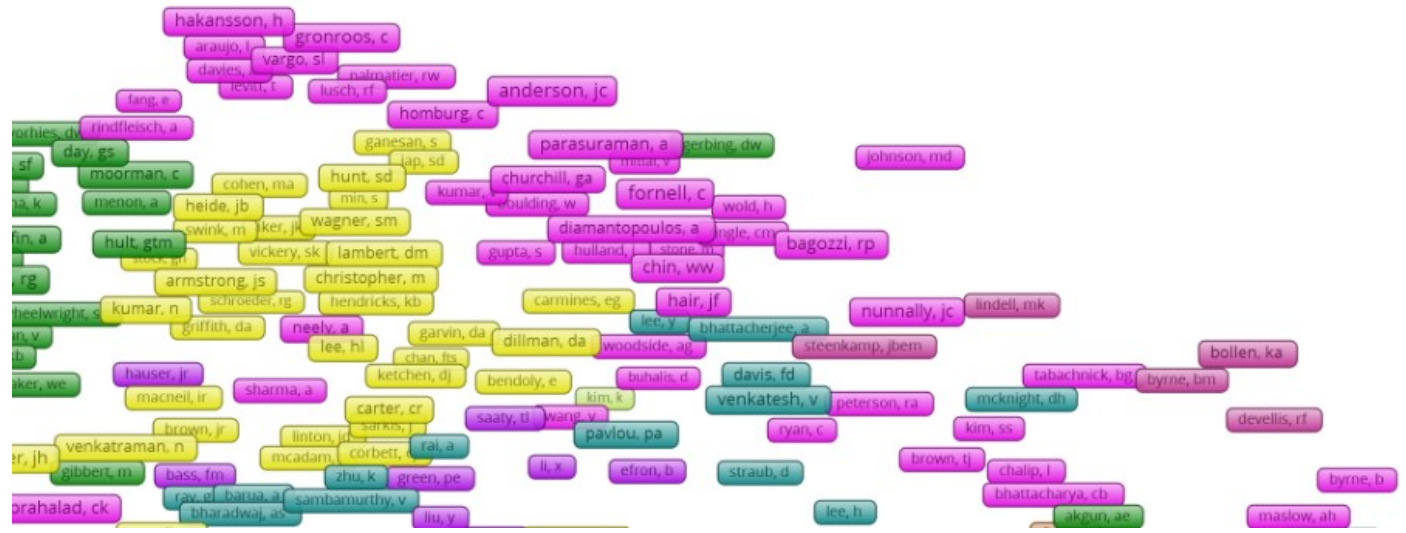




## 93



## Cluster 5. Marketing and Consumer Economics (Light Purple) (111 authors)



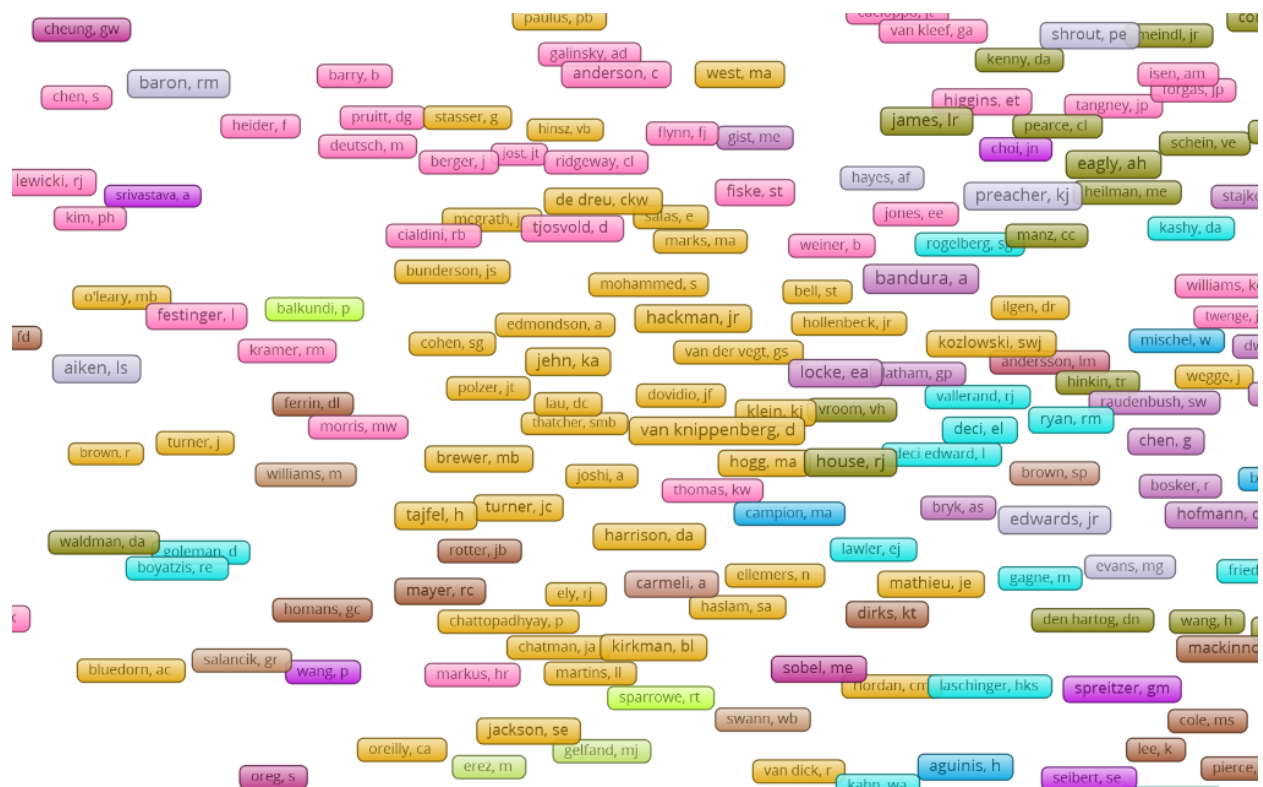
## Cluster 6. Positive Psychology in the work place (Turkish) (97 authors)



## Cluster 7. Employee Appraisals (Light-Blue) (89 authors)



## Cluster 8. Diversity and Team Performance (Dark-Yellow) (87 authors)



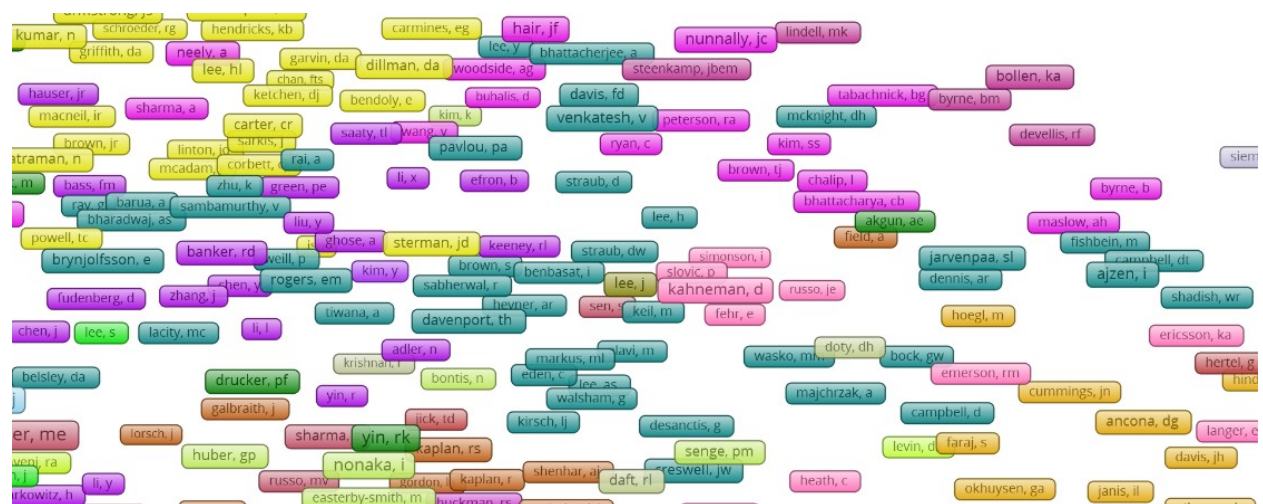
A word cloud visualization of author names, where the size of each name corresponds to the number of papers they have published. The names are arranged in a dense, overlapping pattern, with larger names like 'cohen, j', 'baron, rm', 'bandura, a', and 'lorenz, g' being more prominent. The names are color-coded by their frequency, with a gradient from purple (high frequency) to yellow (low frequency).



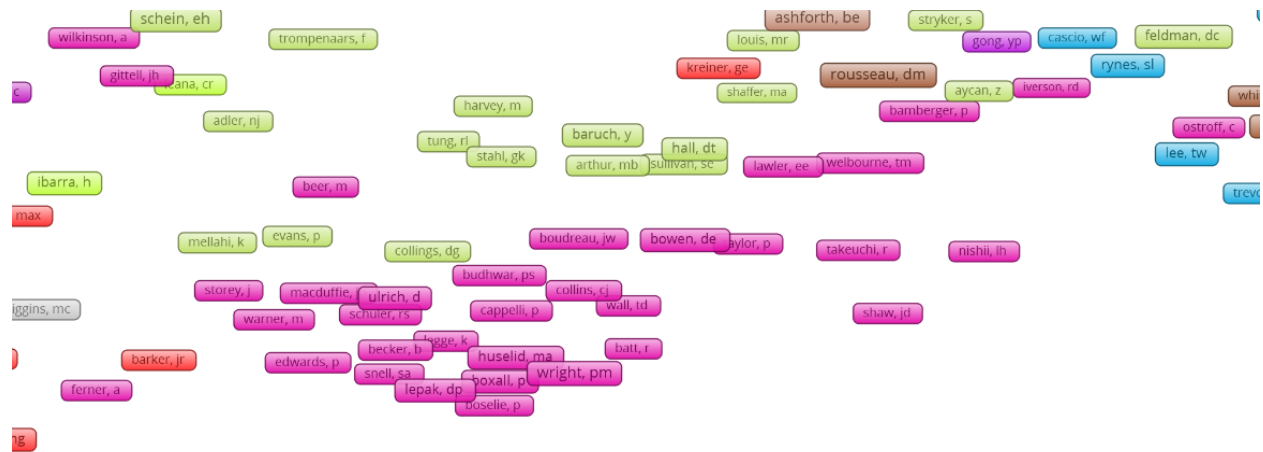
### Cluster 11. Organizational Justice (Brown) (83 authors)



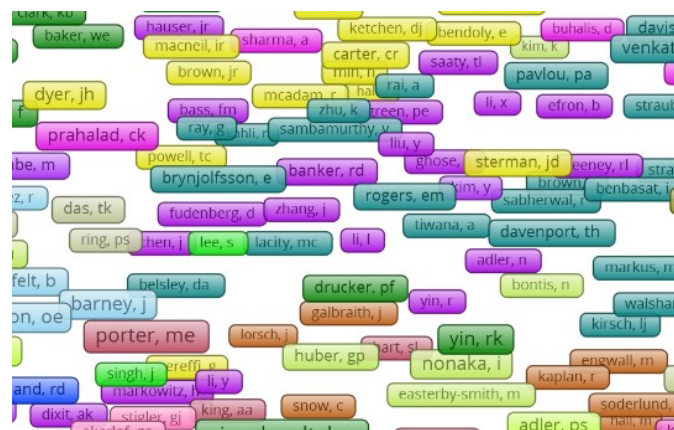
### Cluster 12. Project Planning Systems, Information Tech (Green-Blue) (82 authors)



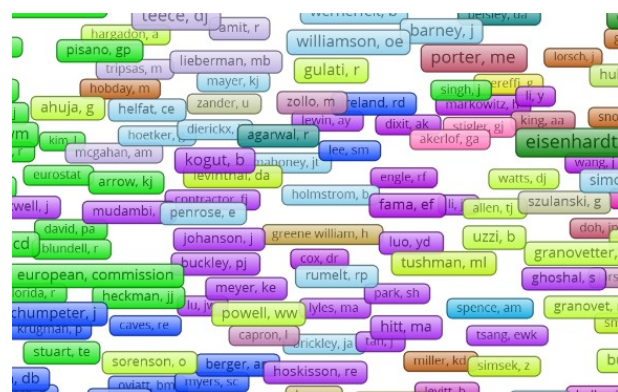
### Cluster 13. HRM (Purple/Pink) (77 authors)



### Cluster 14. Econometrics (Dark Purple) (67 authors)



### Cluster 15. Multinational Business (Purple) (66 authors)



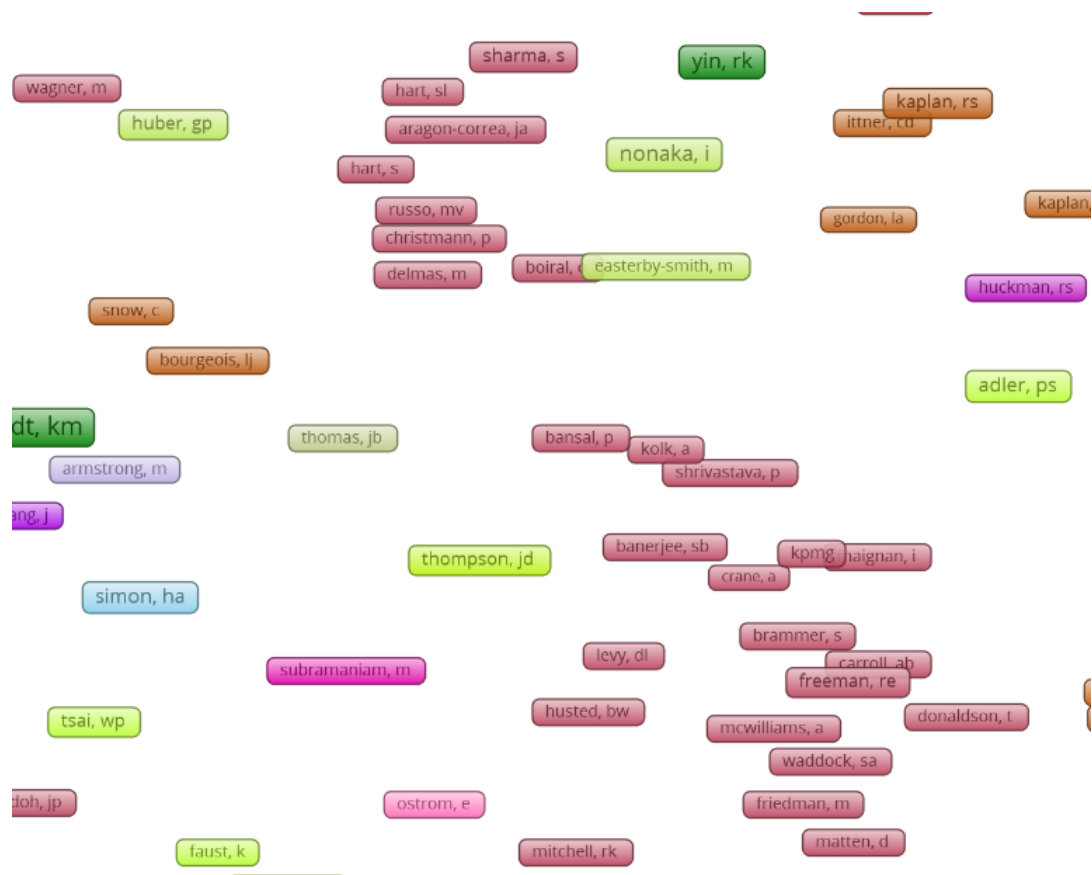




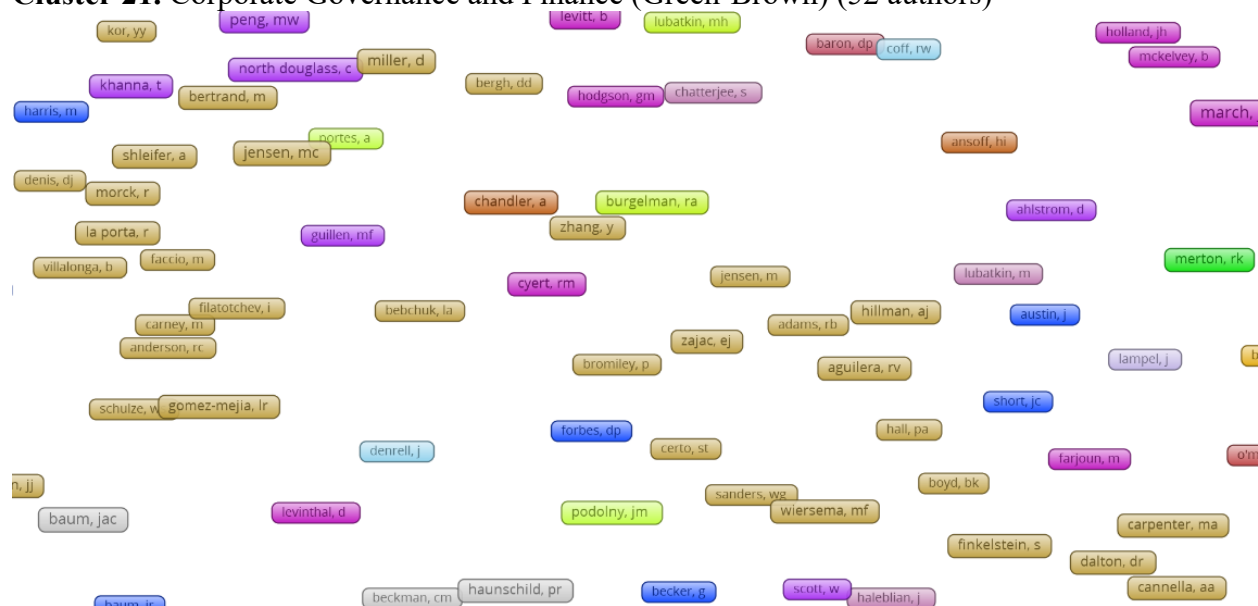




## Cluster 20. CSR (incl. Business Ethics and Environment) (Bordeaux) (53 authors)

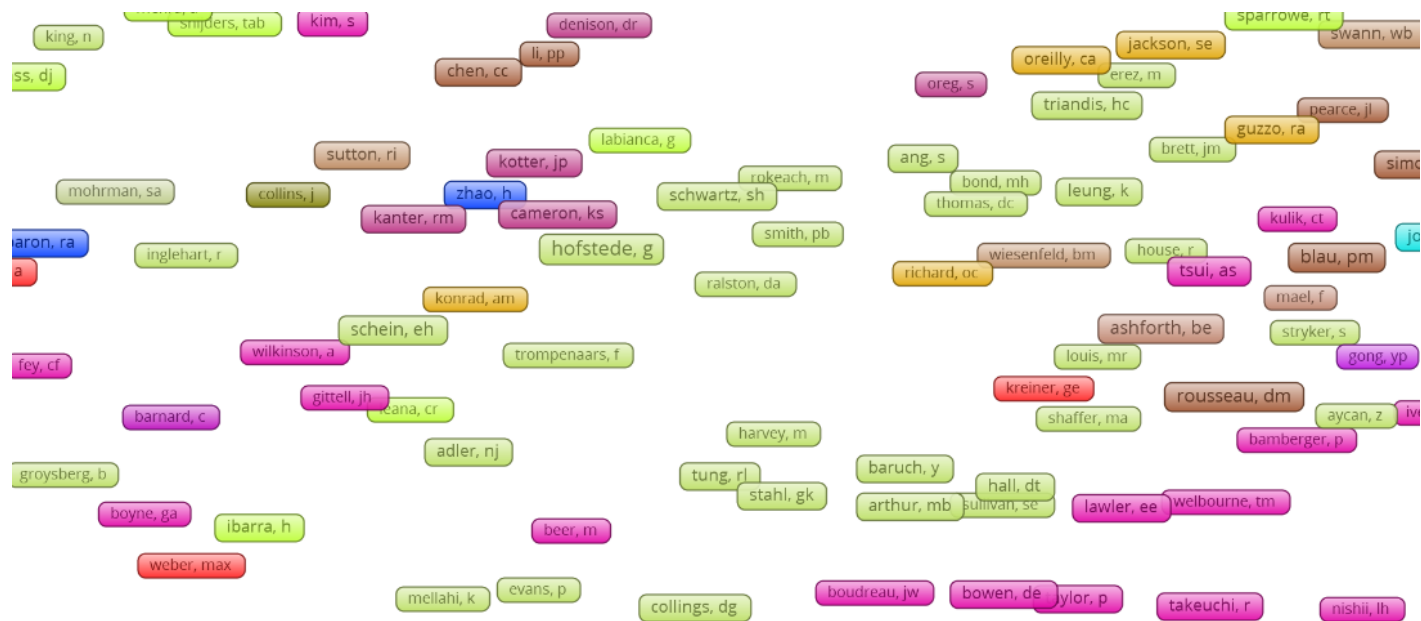


## Cluster 21. Corporate Governance and Finance (Green-Brown) (52 authors)





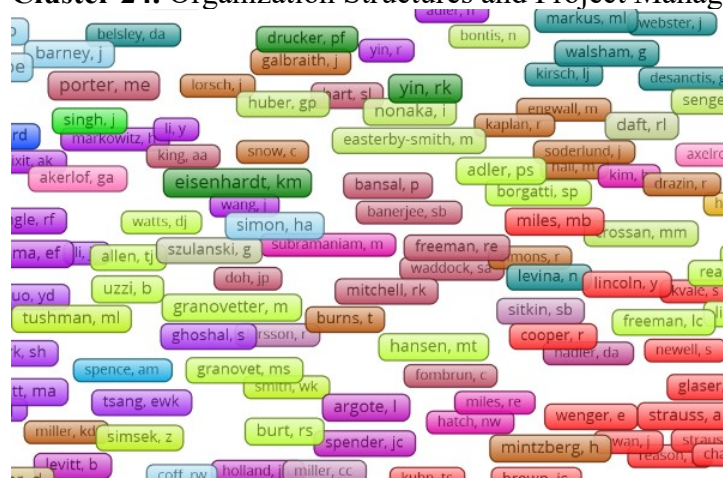
## Cluster 22. Cross-Cultural Management (Light green) (52 authors)



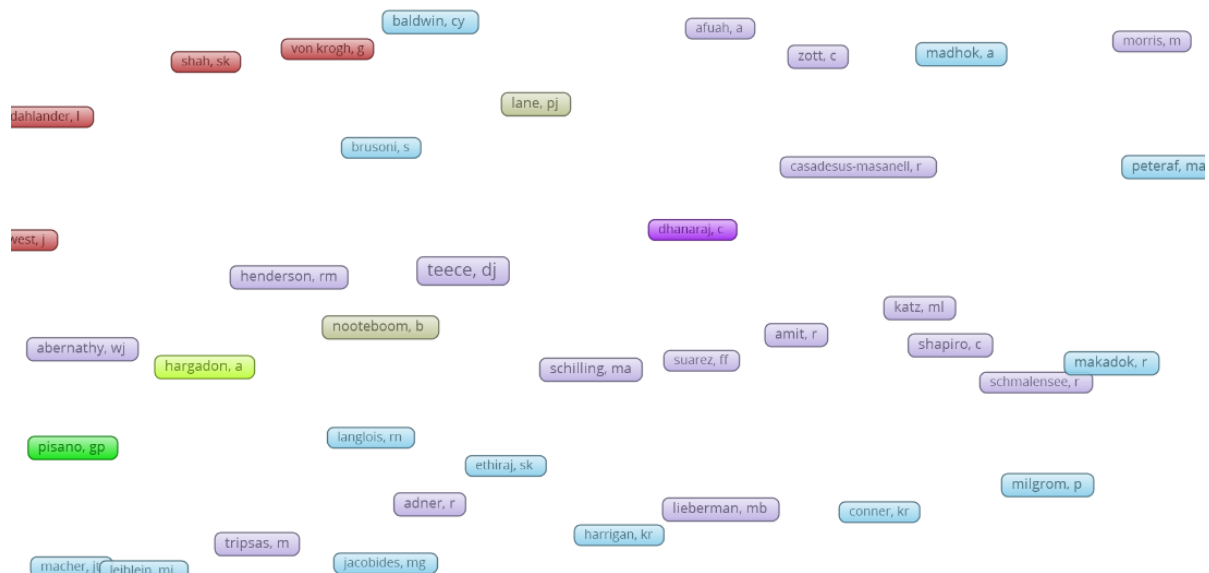
## Cluster 23. Creativity (Purple) (38 authors)



## Cluster 24. Organization Structures and Project Management (Light Brown) (35 authors)



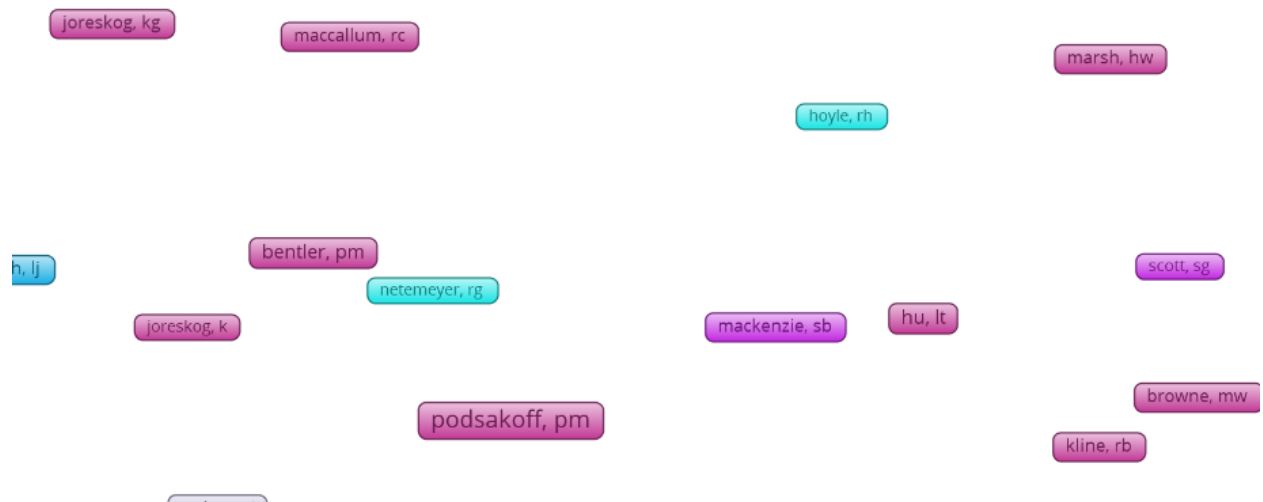
## Cluster 25. Dynamic Capabilities (Light Purple) (34 authors)



## Cluster 26. Interorganizational Relationships (Green-Yellow) (23 authors)

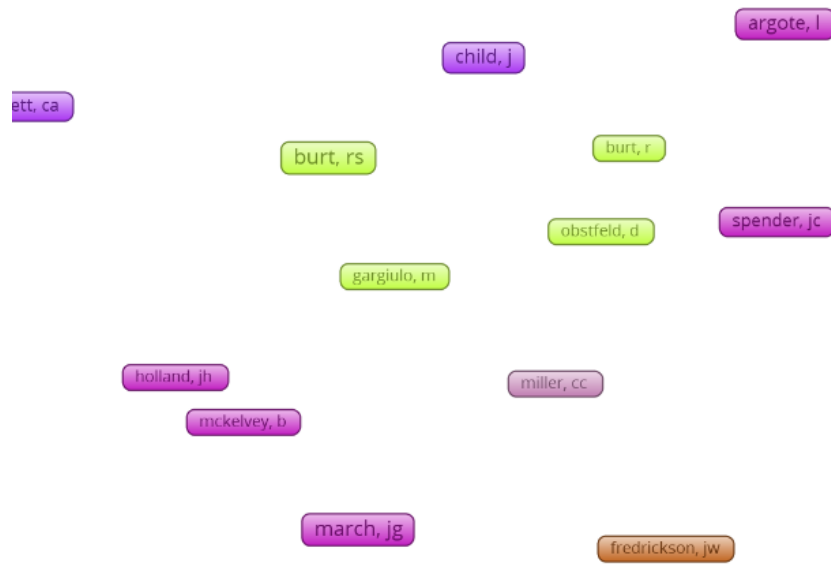


## Table 27. Causal models (Dark-Purple) (20 authors)



[illegible][illegible]

**Cluster 30.** Organizational knowledge, learning and memory (Pink-Purple) (17 authors)

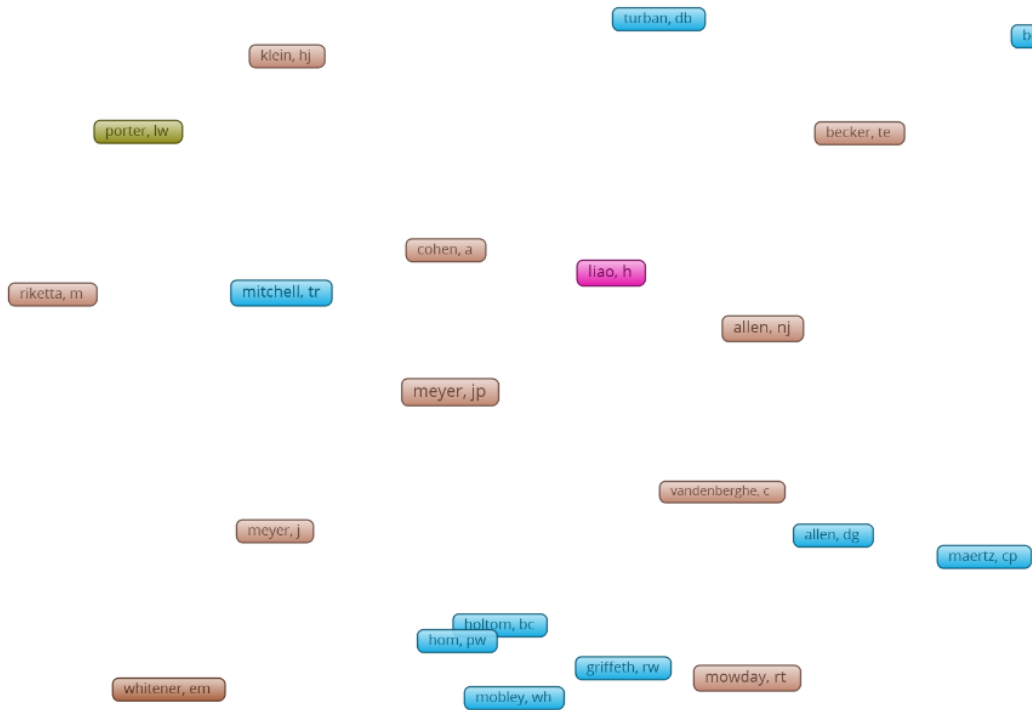


**Cluster 31.** Open Innovation, Online Communities (Bordeaux) (17 authors)

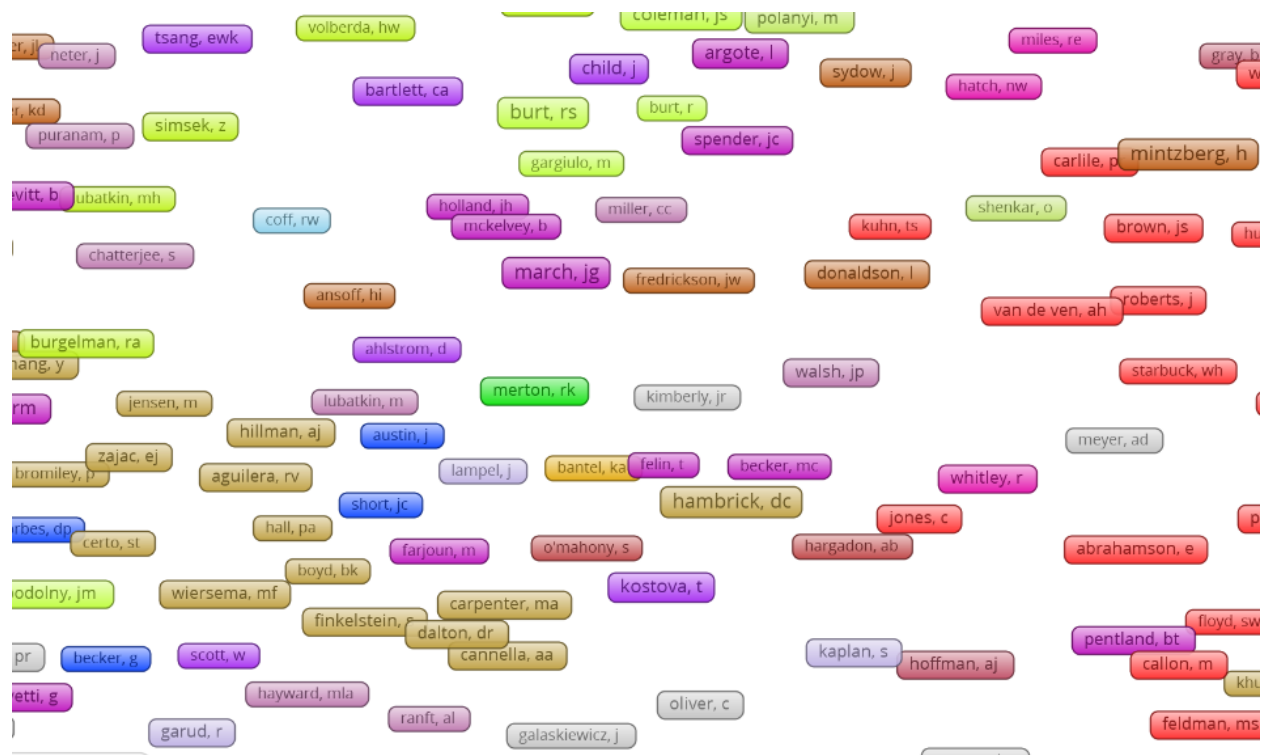




### Cluster 32. Commitment (Light Brown) (15 authors)



### Cluster 33. Mergers and Acquisitions (Light-Bordeaux) (15 authors)



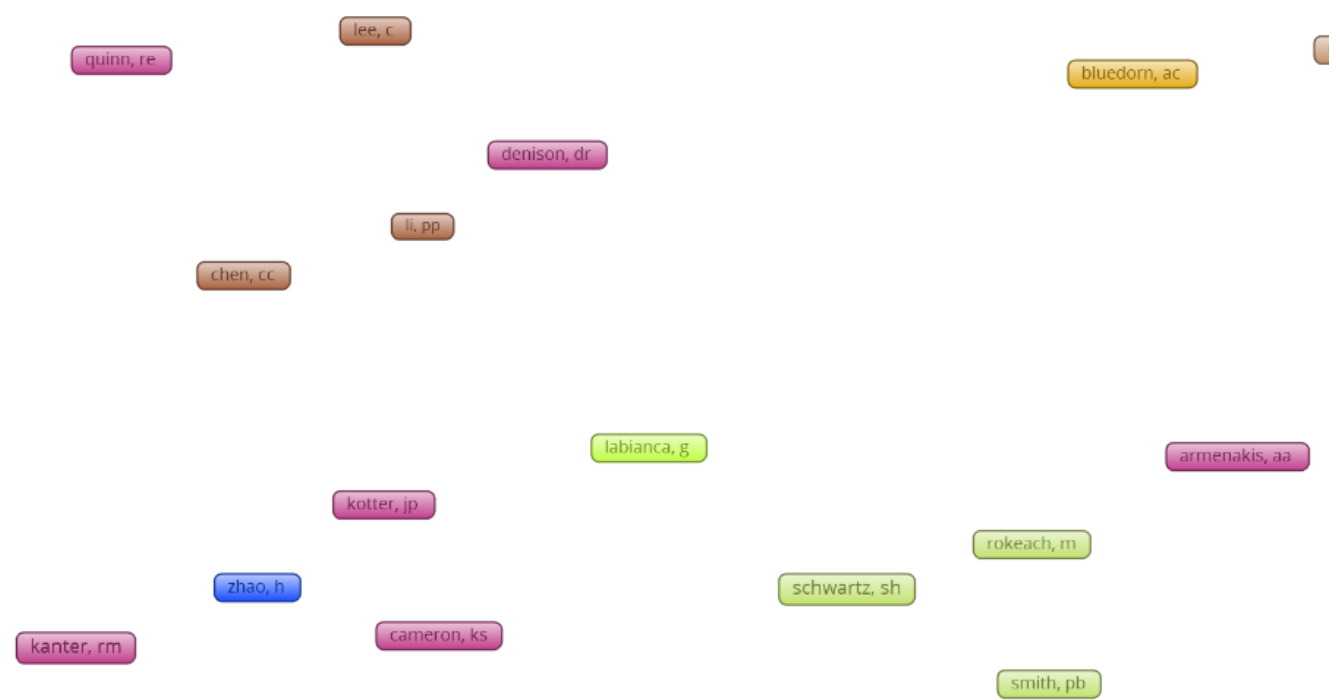
**Cluster 34. Regression models (Grey-Blue) (13 authors)**



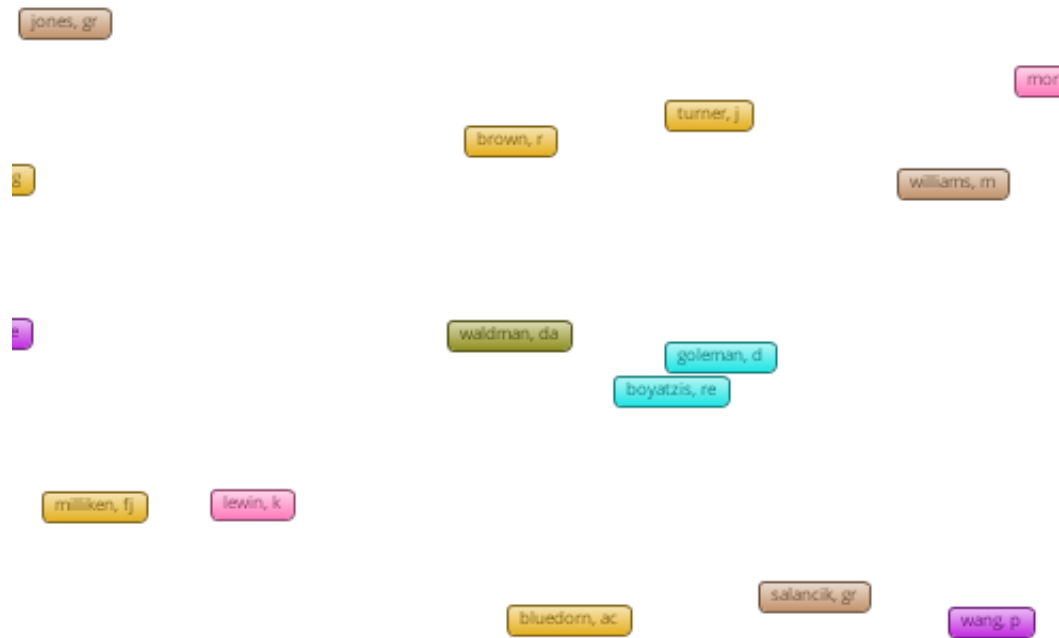
**Cluster 35. Organizational Learning (Light-Green) (8 authors)**



**Cluster 36. Organizational Change (Bordeaux) (8 authors)**



### Cluster 37. Job Attitudes and Behaviors (Brown) (7 authors)



#### Cluster 37 (7 items)

jones, gr
salancik, gr
snyder, m
sutton, ri
swann, wb
wiesenfeld, bm
williams, m

**Cluster 38.** Organizational Performance (Mint Green) (5 authors)



Cluster 38 (5 items)
daft, rl
dess, gg
doty, dh
mohrman, sa
thomas, jb

## 2.2. Co-word analysis and semantic network

Figure A3 presents the parameter settings for the semantic map. The clustering resolution has been set to 2.40 and the minimum cluster size to 10 terms, resulting in 36 clusters each comprising between 73 and 10 terms.

**Figure A3.** Parameter settings for the Semantic map

The screenshot shows a dialog box titled "Advanced Layout Paramet..." with a close button (X) in the top right corner. The dialog is divided into two main sections: "Layout" and "Clustering".

**Layout Section:**

- Random starts:** 1
- Max. iterations:** 1000
- Initial step size:** 1.00
- Step size reduction:** 0.75
- Step size convergence:** 0.001
- Random seed:** 0

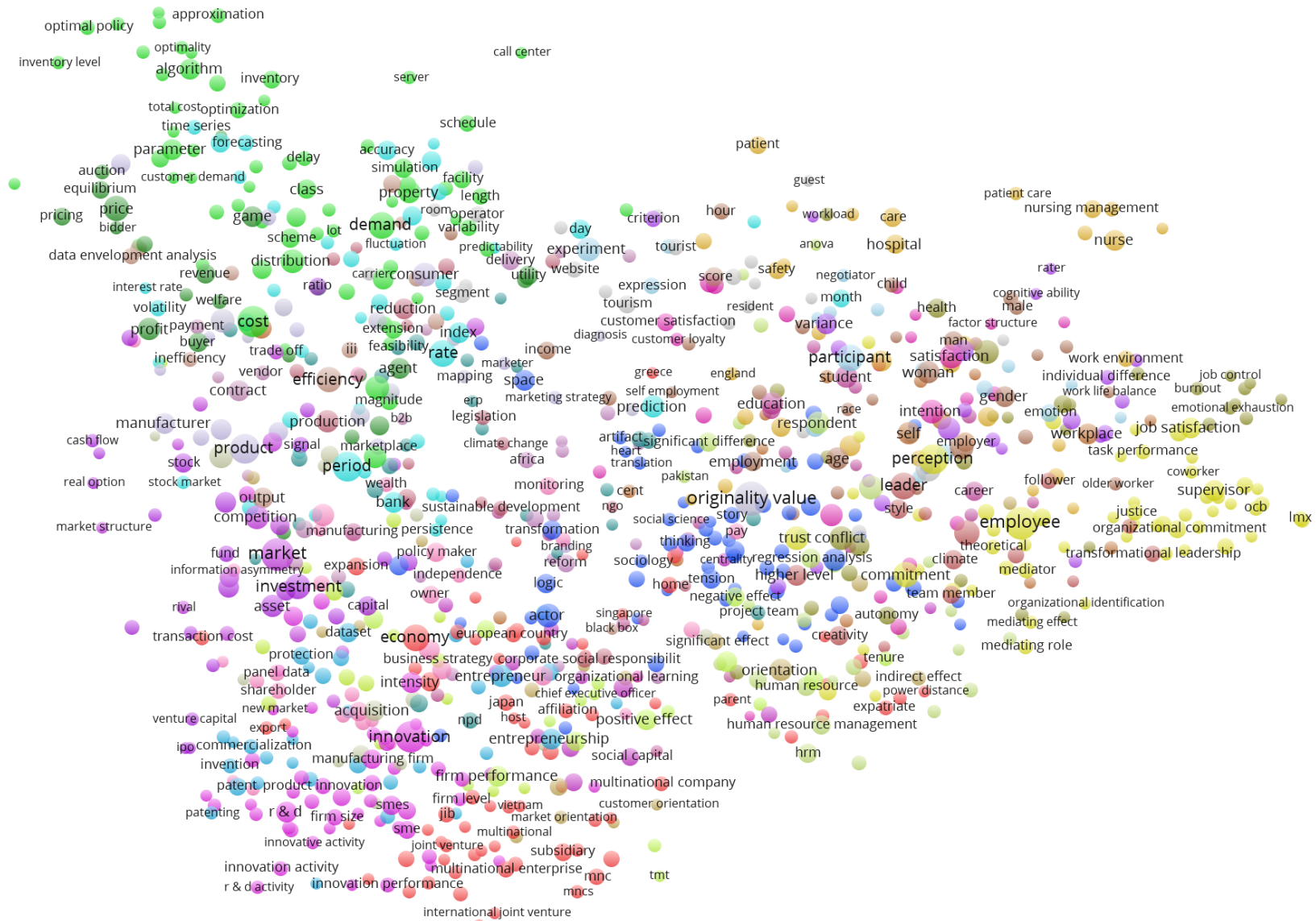
At the bottom of the Layout section are "OK" and "Cancel" buttons.

**Clustering Section:**

- Normalization Method:** Association ...
- Layout:**
  - Attraction:** 2
  - Repulsion:** 1
  - ☒ Use default values ?
  - Advanced Parameters...
  - Update Layout
- Clustering:**
  - Resolution:** 2.40
  - Min. cluster size:** 10
  - ☒ Merge small clusters ?
  - Advanced Parameters...
  - Update Clustering

Figure A4 visualizes the semantic map. Each circle represents a term, and the distance/proximity between them specifies their co-occurrence frequency. Colors and shades are used to distinguish the 36 clusters in the map. Some clusters have fairly similar color-codings, but the search-function in the software can be used to specify which terms pertain to each cluster. An overview of the terms included in each cluster is provided below the semantic network. The clusters can be interpreted as management topics or research areas. As mentioned earlier, the co-word analysis attributes each term to a single cluster. This setup results in a situation where some thematic areas with overlapping terminology are conflated into broader themes. In the cluster solution we therefore see a few fairly broad clusters (see Cluster 12. New Product Development and Project Management and Cluster 31. Leadership and team-performance). Further, this means that topic modeling and co-word analysis to some extent rely on different terms to capture research frontiers that are similar in content.

**Figure A4. Semantic map (Co-word analysis)**



### Cluster 1. Multi-national business and Cross cultural management (73 terms)

- affiliation
- asia
- brazil
- chinese firm
- competitiveness
- cultural context
- cultural difference
- cultural dimension
- cultural distance
- cultural value
- czech republic
- denmark
- economy
- european country
- expansion
- expatriate
- export
- fdi
- foreign direct investment
- foreign
- foreign firm
- foreign market
- france
- globalisation
- globalization
- globe
- greece
- headquarters
- home
- home country
- host
- host country
- ijjv
- institutional environment
- institutional pressure
- institutional theory
- international business
- international joint venture
- internationalisation
- internationalization
- japan
- Japanese firm
- Jib
- Joint venture
- knowledge acquisition
- korea
- liability
- local firm
- malaysia
- mnc
- mncs
- mne
- mnes
- multinational
- multinational company
- multinational corporation
- multinational enterprise
- multinational firm
- national culture
- nationality
- parent
- portugal
- power distance
- russia
- singapore
- south korea
- subsidiary
- taiwan
- technological capability
- transaction cost
- transaction cost economics
- transition economy
- vietnam

### Cluster 2. Operations algorithms (70 terms)

- algorithm
- allocation
- approximation
- branch
- call center
- capacity
- class
- computational experiment
- computational result
- cost
- cost saving
- customer demand
- decision maker
- decision process
- decision support system
- delay
- demand
- demand uncertainty
- deviation
- disruption
- distribution
- extension
- facility
- formulation
- heuristic
- horizon
- inventory
- inventory level
- length
- location
- lot
- magnitude
- mathematical model
- model parameter
- motivated
- node
- numerical experiment
- objective function
- operations research
- optimal policy
- optimality
- optimization
- optimization model
- parameter
- practical application
- property
- resource allocation
- robustness
- route
- saving
- schedule
- scheme
- second
- sensitive analysis
- sequence
- server
- service level
- simulation

- simulation model
- staffing
- subset
- sufficient condition
- system dynamic
- threshold
- total cost
- tradeoff
- transportation
- variability
- variant
- vehicle

### Cluster 3. Sociological theory, Organizational theory and Constructionism

- action research
- actor
- appropriation
- artefact
- business network
- capitalism
- change agent
- change process
- cognition
- conception
- conversation
- criticism
- critique
- dialogue
- discourse
- economic
- essay
- ethic
- experimentation
- heart
- identity
- institutional change
- institutional context
- intersection
- legitimacy
- logic
- longitudinal case study
- meaning
- metaphor
- moment
- narrative
- organization study
- organization theory
- organizational behaviour
- organizational change
- organizational context
- organizational identity
- organizational life
- organizational member
- organizational practice
- organizational research
- organizational setting
- organizational theory
- outline
- participant
- politic
- psychology
- qualitative method
- reflection
- resistance
- rhetoric
- scholarship
- sensemaking
- social context
- social science
- sociology
- space
- story
- strategic management
- tension
- thinking
- thought
- tradition
- transformation
- translation

### Cluster 4. Commitment (58 terms)

- absenteeism
- affective commitment
- affective organizational
- chinese context
- co worker
- commitment
- co-worker
- differential effect
- direct relationship
- distributive justice
- dyad
- employee
- employee attitude
- employee perception
- employees perception
- fairness
- group level
- integrity
- interactive effect
- job performance
- job satisfaction
- justice
- leader member exchange
- lmx
- mediating effect
- mediating role
- mediation
- mediator
- moderate
- moderator
- newcomer
- ocb
- organizational citizenship
- organizational commitment
- organizational identification
- organizational outcomes
- organizational support
- perception
- pos
- procedural justice
- reciprocity
- salience
- social exchange
- social exchange theory
- social identity
- social identity theory
- subordinate
- supervisor
- supervisor support
- task performance
- theoretical
- trust
- trustworthiness
- turnover intention



- way interaction
- work group

### Cluster 5. Dynamic capability (51 terms)

- |                          |   |                             |
|--------------------------|---|-----------------------------|
| • absorptive capacity    | • innovation system                     | • organizational innovation |
| • dynamic capability     | • innovative performance                | • panel                     |
| • expenditure            | • innovator                             | • process innovation        |
| • exploitation           | • intensity                             | • product innovation        |
| • external knowledge     | • knowledge base                        | • r & d                     |
| • firm growth            | • knowledge intensive business strategy | • r & d activity            |
| • firm level             | • large firm                            | • r & d investment          |
| • firm size              | • larger firm                           | • radical innovation        |
| • global market          | • manufacturing firm                    | • service innovation        |
| • imitation              | • manufacturing sector                  | • small firm                |
| • incremental innovation | • many firm                             | • smaller firm              |
| • innovation             | • medium enterprise                     | • sme                       |
| • innovation activity    | • medium sized enterprise               | • smes                      |
| • innovation management  | • new market                            | • spanish firm              |
| • innovation performance | • new product                           | • technological development |
| • innovation policy      | • open innovation                       | • technological innovation  |
| • innovation process     |   |                             |
| • innovation strategy    |   |                             |

### Cluster 6. Predictive modelling (markets and finance) (44 terms)

- |                   |                           |                   |
|-------------------|---------------------------|-------------------|
| • accuracy        | • financial crisis        | • persistence     |
| • bank            | • financial institution   | • predictability  |
| • beginning       | • financial market        | • prediction      |
| • calculation     | • fluctuation             | • rate            |
| • coefficient     | • forecast                | • real time       |
| • credit          | • forecasting             | • recession       |
| • day             | • gdp                     | • revision        |
| • duration        | • global financial crisis | • sensitivity     |
| • economic crisis | • index                   | • stock market    |
| • equation        | • inflation               | • strong evidence |
| • error           | • interest rate           | • time series     |
| • estimate        | • monte carlo simulation  | • timing          |
| • estimation      | • month                   | • volatility      |
| • factor model    | • online                  |                   |
| • finance         | • period                  |                   |

### Cluster 7. Knowledge transfer (33 terms)

- |                        |                            |                             |
|------------------------|----------------------------|-----------------------------|
| • academia             | • entrepreneur             | • knowledge spillover       |
| • citation             | • entrepreneurial activity | • licensing                 |
| • commercialization    | • founder                  | • nanotechnology            |
| • dataset              | • funding                  | • new firm                  |
| • discovery            | • intellectual property    | • new venture               |
| • economic development | • invention                | • organizational boundaries |
| • economic growth      | • knowledge flow           |                             |

- patent
- patenting
- policymaker
- protection
- public policy
- scientist
- spillover
- start up
- technology development
- technology transfer
- venture

#### **Cluster 8. Healthcare management (32 terms)**

- care
- coaching
- empowerment
- england
- focus group
- health care
- hospital
- leadership development
- mentoring
- nurse
- nurse manager
- nursing
- nursing management
- patient
- patient care
- physician
- pivotal role
- profession
- professional
- programme
- qualitative approach
- qualitative data
- recruitment
- registered nurse
- retention
- safety
- shortage
- staff
- work environment
- workforce

#### **Cluster 9. Workplace conditions and satisfaction (32 terms)**

- autonomy
- burnout
- causality
- colleague
- conflict
- conservation
- cross sectional design
- emotional exhaustion
- engagement
- facilitation
- family conflict
- health
- hierarchical regression analysis
- intrinsic motivation
- job
- job control
- job demand
- resource
- longitudinal data
- longitudinal study
- multiple regression analysis
- negative outcome
- self report
- social support
- strain
- stress
- work engagement
- work family conflict
- workload

#### **Cluster 10. Corporate governance (terms)**

- agency
- analyst
- board
- board member
- code
- compensation
- concentration
- corporate governance
- director
- disclosure
- family firm
- firm characteristic
- firm value
- fraction
- governance
- governance mechanism
- governance structure
- incentive
- independence
- information asymmetry
- monitoring
- owner
- ownership
- ownership structure
- panel data
- pay
- policy maker
- regulator
- shareholder
- wealth

#### **Cluster 11. Structural inequality (20 terms)**

- age
- american
- child
- demographic characteristic
- demographic variable
- discrimination
- employment
- ethnicity
- female
- gender
- gender difference
- hour
- household
- income
- labour market
- male
- man

- minority
- occupation
- older worker
- personal characteristic
- race
- self employment
- sex
- unemployment
- wage
- woman
- work life balance
- worker

#### **Cluster 12. New product development & Project management (28 terms)**

- architecture
- budget
- business process
- complex system
- construction project
- contractor
- critical success factor
- designer
- erp
- feasibility
- functionality
- life cycle
- modularity
- new product development
- npd
- open source software
- procurement
- product design
- product development
- project management
- project manager
- project performance
- project success
- project team
- questionnaire survey
- realization
- risk management
- successful implementation

#### **Cluster 13. Service Quality and Customer Satisfaction (25 terms)**

- behavioural intention
- construct
- customer loyalty
- customer relationship
- customer satisfaction
- discriminant validity
- dissatisfaction
- empathy
- exploratory factor analysis
- factor analysis
- intention
- loyalty
- online survey
- partial least square
- reliability
- satisfaction
- sem
- service quality
- structural equation model
- structural equation modelling
- structural equation modelling
- structural model

#### **Cluster 14. Corporate finance (25 terms)**

- asset
- capital
- cash
- cash flow
- debt
- earning
- endogeneity
- equity
- financing
- fund
- investment
- investment decision
- investor
- ipo
- leverage
- portfolio
- proxy
- ratio
- real option
- share
- signal
- stock
- valuation
- venture capital
- venture capitalist

#### **Cluster 15. Employee appraisals (22 terms)**

- adverse impact
- applicant
- appraisal
- cognitive
- cognitive ability
- conscientiousness
- criterion
- emotional intelligence
- individual difference
- locus
- meta analysis
- openness
- optimism
- performance appraisal
- personality
- personality trait
- predictor
- rater
- rating
- self efficacy
- trait
- variance

#### **Cluster 16. Tourism (22 terms)**

- airline
- attitude
- attraction
- carrier
- friend
- guest
- hospitality
- hospitality industry
- hotel
- intent
- operator
- positive attitude
- resident
- restaurant
- room
- segment
- tourism
- tourism industry
- tourist
- travel
- visitor
- website

#### **Cluster 17. Game theory (21 terms)**

- agent
- auction
- bidder
- buyer
- coalition
- equilibrium
- expense
- game
- offer
- payoff
- player
- price
- pricing
- profit
- revenue
- risk aversion
- seller
- substitute
- utility
- welfare
- winner

#### **Cluster 18. Behavioral management (20 terms)**

- anger
- anxiety
- attribution
- control group
- emotion
- experiment
- experimental design
- experimental study
- expression
- field study
- group member
- laboratory experiment
- negative consequence
- negative emotion
- participant
- social influence
- surface

#### **Cluster 19. Top management teams and performance (19 items)**

- archival data
- ceo
- chief executive officer
- environmental uncertainty
- financial performance
- firm performance
- heterogeneity
- inverted u
- moderating effect
- negative effect
- performance implications
- positive association
- positive effect
- strategic change
- survey data
- tenure
- tmt
- top management team
- top manager

#### **Cluster 20. Corporate Social Responsibility (18 items)**

- climate change
- consumption
- corporate social responsibility
- csr
- disaster
- dollar
- environmental impact
- environmental issue
- environmental performance
- equipment
- law
- legislation
- million
- ngo
- production
- reduction
- social responsibility
- sustainable development

#### **Cluster 21. 18 Strategic Orientations (18 terms)**

- business performance
- customer need
- customer orientation
- customer value
- direct effect
- direct impact
- entrepreneurial orientation
- indirect effect
- innovativeness
- market condition
- market orientation
- market performance
- organizational learning
- orientation

- path analysis
- significant effect
- strategic orientation
- structural equation

### Cluster 22. Human Resource Management (18 terms)

- hr manger
- hr practice
- hrm
- hrm
- human resource
- human resource management
- human resource practice
- interview data
- line manager
- middle manager
- organizational culture
- organizational effectiveness
- organizational factor
- organizational level
- organizational performance
- public sector organization
- strategic human resource management
- training

### Cluster 23. (17 terms)

- accumulation
- competition
- competitive environment
- competitor
- decrease
- entrant
- incumbent
- market
- market share
- market structure
- network effect
- new technology
- profitability
- rival
- technological change
- trade
- trade off

### Cluster 24. Workplace bullying (17 terms)

- bullying
- cent
- employer
- employment relationship
- exclusion
- important determinant
- probability
- psychological contract
- representative sample
- self
- turnover
- victim
- violation
- volunteer
- workplace

### Cluster 25. Marketing (16 terms)

- advertising
- branding
- channel
- consumer
- distributor
- manufacturer
- marketer
- marketing strategy
- marketplace
- producer
- product
- product category
- purchase
- retailer
- sale
- store

### Cluster 26. Strategic management (15 terms)

- business model
- business strategy
- business value
- capture
- commodity
- complementarity
- corporate strategy
- diversification
- organizational capability
- rbv
- strategic choice
- superior performance
- sustainable competitive advantage

### Cluster 27. Management learning and education (15 terms)

- business leader
- business school
- career
- career development
- education
- educator
- faculty
- graduate
- higher education
- management education
- school
- teacher
- teaching
- work experience

**Cluster 28. Development management (15 terms)**

- africa
- bottom
- citizen
- corruption
- delivery
- developed country
- latin america
- lesson
- local community
- negative influence
- poverty
- public service
- reform
- significant influence

**Cluster 29. Mergers and Acquisitions (14 terms)**

- acquirer
- acquisition
- asymmetry
- buyer supplier relations
- focal firm
- m & a
- merger
- performance outcome
- regime
- short run
- strategic alliance
- supplier
- theoretical argument

**Cluster 30. Knowledge management and networks (14 terms)**

- centrality
- communication technology
- human capital
- ict
- intellectual capital
- knowledge creating
- knowledge management
- knowledge sharing
- knowledge transfer
- knowledge worker
- network structure
- social capital
- social network analysis
- tacit knowledge

**Cluster 31. Leadership and team performance (12 terms)**

- climate
- creativity
- follower
- higher level
- leadership
- leadership style
- style
- team level
- team performance
- team work
- transformational leaders
- virtual team

**Cluster 32. Decision making (13 terms)**

- ahp
- benchmark
- black box
- data envelopment analysis
- dea
- Efficiency
- iii
- inefficiency
- output
- performance measure
- ranking score
- weight

**Cluster 33. B2B outsourcing (12 terms)**

- b2b
- cost reduction
- customer service
- e commerce
- intermediary
- lower cost
- outsourcing
- payment
- privacy
- transaction
- vendor

**Cluster 34. Supply chain management (12 terms)**

- diagnosis
- flow
- logistic
- mapping
- originality value
- scm
- single case study
- supply
- supply chain
- supply chain management
- supply chain performance
- workshop

### Cluster 35. Survey methods (10 terms)

- |                        |                          |                            |
|------------------------|--------------------------|----------------------------|
| • anova                | • regression analysis    | • significant relationship |
| • independent variable | • respondent             | • t test                   |
| • pakistan             | • sample size            |                            |
| • quantitative study   | • significant difference |                            |

### Cluster 36. TQM and Company Performance (10 terms)

- |                       |                   |                      |
|-----------------------|-------------------|----------------------|
| • company performance | • operational     | • quality management |
| • iso                 | performance       | • total quality      |
| • manufacturing       | • organizational  | management           |
| • multiple case study | performance       | • tqm                |
|                       | • product quality |                      |

## 3. Comparing outcomes of the three techniques

Table A1 displays similarities and differences in the outcomes of the topic model, co-citation analysis and co-word analysis. Each row in the table represents a given research area, topic or theme. The topic model is the baseline, meaning that the clusters derived from the co-citation and co-word analysis are assigned to one or several rows in the table based on overlaps with content in the topic model (henceforth TM). Since clusters in the co-citation and co-word analysis tend to vary considerably with respect to level of specificity, some topics (or themes) are assigned to several rows in the table (see e.g. Cluster 12. New Product Development and Project Management in column 3).

26 of the co-citation clusters have clear overlaps with content in the TM. The co-citation analysis, however, fails to capture the most generic TM-topics (“literature reviews”, “time”) and topics representing smaller or recently emerging areas in the management literature (“decision-making”, “game theory”, “operations algorithms”, “crisis management”, “service operations”, “tourism”, “socio-economic policies”). This may partly be explained by limitations of the method. Since, co-citation analysis relies on references to prior literature, its clustering networks are not suitable for capturing the most recent developments in a field. Further co-citation analysis has a bias towards established scientific frontiers with high citation-rates and is less reliable for clustering niche specialties comprised by documents/authors with relatively low citation rates.

In comparison, the co-word analysis has clear overlaps with 28 topics in the TM. Five of these cover TM-topics not captured by the co-citation analysis (“decision-making”, “game theory”, “operations algorithms”, “service operations”, and “health management”). Four topics in the TM

share no clear affinities with topics identified in neither the co-citation nor co-word analysis (i.e. “socio-economic policies”, “crisis management”, “time” and “literature reviews”). These four topics may be viewed as the least convincing representations of autonomous “knowledge areas” identified in our model.

Further, of the 38 clusters in the co-citation analysis, four clusters do not share any clear affinity with a specific TM-topic (i.e. “goal orientation and self-efficacy”, “creativity”, “social psychology”). All of these, however, indirectly taps into one or more TM-topics: The literature on creativity (cluster X) has links to research discussions on innovation, organizational learning, and leadership, the literature on goal orientation and self-efficacy is linked to management research on performance appraisals, commitment and behavioral Mgmt, and the social psychological literature strongly influences the scholarly work on structural inequalities, team management and commitment.

Overlaps between the outcomes of the co-citation and co-word analysis exist for 22 clusters. A closer look at the outcomes of the co-word analysis reveals the existence of a few fairly broad clusters tapping into to several clusters/topics in the co-citation analysis and TM (i.e. “New Product Development and Project Management”, “Team management and performance”). The co-word clusters without overlaps (“development mgmt.”, “workplace bullying” and “TQM and Company performance”) are all intuitively meaningful, though highly specific.

Summing up, this comparison has confirmed the TM’s aptitude in capturing key research frontiers in the management literature. Further, it offers converging evidence for the plausibility of our interpretations of model results. A few topics in our model are, too general (“time” and “literature reviews”) while others tend to be too specific or “niche” (“socio-economic policies” and “crisis management”) to be captured by alternative science mapping techniques. The uncertainties associated with these topics should off course be taken into account in the next steps of the analysis.



**Table A1.** Overlap and variation between Topic model, Co-citation and co-word analysis

Topic Model (36 topics)	Co-citation clusters (38 clusters)	Co-word clusters
1. Innovation	2. Innovation (R&D), 17. New Product Develop., 31. Open Innovation	12. New Product Development & Project Mgmt. 23. Technology Development and Market Structures
2. Dynamic Capabilities	3. Entrepreneurship, 25. Dynamic Capabilities	5. Dynamic Capabilities
3. Organizational Learning	30. Org. Knowledge, Learning & Memory 35. Org. Learning	
4. Time		
5. Interorg. Relationships	26. Interorg. Relationships (IOR), 29. Strategic Alliances, 32. Mergers & Acquisitions	29. Mergers & Acquisitions 32. B2B outsourcing
6. Strategic Mgmt	18. Strategic Mgmt. and Resource exchange 35. Strategic Planning and Mgmt. Systems 38. Org. Performance	26. Strategic Mgmt
7. Commitment	6. Positive Psych. in the workplace, 32. Commitment, 11. Org. Justice	4. Commitment, 9. Workplace Conditions and Satisfaction
8. Leadership (education and training)	9. Leadership, 37. Organizational Change	27. Management education and learning 31. Leadership & Team Performance
9. Decision-making		32. Decision-making
10. HRM	13. HRM	22. HRM
11. Survey Studies	33. Regression models	35. Survey Studies
12. Multinational Biz	15. Multinational Business, 22. Cross-cultural Mgmt. & Psych. 27. Causal models	1. Multinational Business. & Cross-cultural Mgmt.
13. Causal effects		
14. Game theory		17. Game Theory
15. Structural Inequality	10. Social Psychology	11. Structural Inequality
16. Behavioral Mgmt	37. Job Attitudes and Behaviors	18. Behavioral Mgmt.
17. Knowledge transfer	19. Networks & Social Capital	7. Knowledge Transfer
18. Corporate Governance	21. Corporate Governance & Finance	10. Corporate Governance
19. Project Mgmt	12. Project Planning Systems & Information Tech., 24. Org. Structures & Project Mgmt. 5. Marketing and Consumer econ.	12. New Product Development & Project Mgmt
20. Consumer Econ.		
21. Operations Algorithms		2. Operations Algorithms
22. Crisis Mgmt		
23. Service Operations		13. Service Quality & Customer Satisfaction
24. Predictive Modeling	14. Econometrics	6. Predictive Modeling (Markets and Finance)
25. Literature Reviews		
26. Inventory Mgmt.	4. Supply Chain Mgmt.	34. Supply Chain Mgmt.
27. CSR	20. CSR	20. CSR
28. Tourism		16. Tourism
29. Online Marketing	5. Marketing and Consumer economics	25. Marketing, 21. Strategic Orientations
30. Supply Chain Mgmt	4. Supply Chain Mgmt	34. Supply Chain Mgmt.
31. Socio-economic Pol.		
32. Team Mgmt	8. Diversity and Team Performance	19. Top Mgmt Teams & Performance
33. Employee Appraisals	7. Employee Appraisals 23. Organizational Citizenship 1. Social and Org. Theory, Construct. 16. Sociological New Institutionalism	15. Employee Appraisals 3. Sociological Theory, Org. theory and Construct.
34. Constructionism		
35. Healthcare Mgmt		8. Healthcare Mgmt
36. Corporate Finance	21. Corporate Governance & Finance 28. Goal Orientation and Self-Efficacy 23. Creativity 10. Social Psychology	14. Corporate Finance
		28. Development Mgmt 24. Workplace bullying 36. TQM & Company Performance

## References

- Braam, R. R., Moed, H. F., van Raan, A. F. J., 1991a. Mapping of science by combined co-citation and word analysis. I. Structural aspects. *Journal of the American Society for Information Science*, 42(4), 233-251.
- Braam, R. R., Moed, H. F., van Raan, A. F. J., 1991b. Mapping of science by combined co-citation and word analysis. II: Dynamical aspects. *Journal of the American Society for Information Science*, 42(4), 252-266.
- González-Álvarez, J., & Cervera-Crespo, T. (2017a). Research production in high-impact journals of contemporary neuroscience: A gender analysis. *Journal of Informetrics*, 11(1), 232-243.
- González-Álvarez, J., & Cervera-Crespo, T. (2017b). Contemporary psychology and women: A gender analysis of the scientific production. *International Journal of Psychology*.
- He, Q., 1999. Knowledge discovery through co-word analysis. *Library trends*, 48(1), 133-133.
- Leydesdorff, L., & Nerghe, A. (2017). Co-word maps and topic modeling: A comparison using small and medium-sized corpora ( $N < 1,000$ ). *Journal of the Association for Information Science and Technology*, 68(4), 1024-1035.
- Nielsen, M. W. (2017). Gender and citation impact in management research. *Journal of Informetrics*, 11(4), 1213-1228.
- Schneider, J. W., Borlund, P., 2004. Introduction to bibliometrics for construction and maintenance of thesauri: methodical considerations. *Journal of Documentation*, 60(5), 524-549. <https://doi.org/10.1108/00220410410560609>
- Van Eck, N. J., Waltman, L., 2011. Text mining and visualization using VOSviewer. *arXiv preprint arXiv:1109.2058*
- van Eck, N. J., Waltman, L., 2014. Visualizing bibliometric networks. In Ding Y., Rousseau R., Wolfram D. (eds) *Measuring Scholarly Impact*. Springer International Publishing, pp. 285-320.
- Waltman, L., Van Eck, N.J., Noyons, E.C.M., 2010. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635. <https://doi.org/10.1016/j.joi.2010.07.002>
- Zupic, I., Čater, T., 2015. Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429-472. DOI: 10.1177/1094428114562629

