Research Policy 44 (2015) 684–697

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Creativity in scientific teams: Unpacking novelty and impact☆

You-Na Lee[a], John P. Walsh[b,*], Jian Wang[c]

[a] Georgia Institute of Technology, USA
[b] Georgia Institute of Technology, USA, National Graduate Institute for Policy Studies (GRIPS), Japan
[c] University of Leuven, Belgium, Institute for Research Information and Quality Assurance (iFQ), Germany

ARTICLE INFO

Article history:
Received 4 January 2014
Received in revised form 20 October 2014
Accepted 20 October 2014
Available online 15 November 2014

Keywords:
Team science
Creativity
Interdisciplinarity
Division of labor
Size

ABSTRACT

The increasing dominance of team science highlights the importance of understanding the effects of team composition on the creativity of research results. In this paper, we analyze the effect of team size, and field and task variety on creativity. Furthermore, we unpack two facets of creativity in science: novelty and impact. We find that increasing team size has an inverted-U shaped relation with novelty. We also find that the size–novelty relationship is largely due to the relation between size and team field or task variety, consistent with the information processing perspective. On the other hand, team size has a continually increasing relation with the likelihood of a high-impact paper. Furthermore, variety does not have a direct effect on impact, net of novelty. This study develops our understanding of team science and highlights the need for a governance approach to scientific work. We also advance the creativity literature by providing an ex ante objective bibliometric measure that distinguishes novelty from impact, and illustrate the distinct team-level drivers of each. We conclude with a discussion of the policy implications of our findings.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

While traditionally science is seen as an individual endeavor, increasingly scientific projects are group activities (Hicks and Katz, 1996; Katz and Martin, 1997; Shrum et al., 2007), and the groups are growing larger (Adams et al., 2005; Wuchty et al., 2007). While high-energy experimental physics is the extreme example, even in other fields we can find research labs with dozens of members and research papers with 10 or more authors. For example, Wuchty et al. (2007) showed the rise in the number of authors per paper over the last 40 years, with mean group size in science and engineering nearly doubling over this period. Similarly, Adams et al. (2005) found an increase in co-authored papers, in the number of authors per paper, in papers spanning institutions, and in international collaborations. This increase in collaboration is driven by a variety of factors, including the importance of interdisciplinary research questions, growing specialization and the consequent gains from trade and division of labor, the diffusion of the Internet, and the need to develop and access large shared equipment and large databases (de Solla Price, 1986; Katz and Martin, 1997; Stephan, 2012). Jones (2009) argued that the burden of knowledge accumulation pushes scientists to specialize, increasing the need to work in teams that incorporate a variety of specialists to collectively solve a problem that spans narrow specializations.

The result of these changes is that increasingly scientific work takes place in a setting that more closely resembles a small factory, rather than an individual's lab bench (Etzkowitz, 1983; Hemlin et al., 2013; Latour and Woolgar, 1979; Shrum et al., 2007; Swatez, 1966). While scientific work has long taken place inside formal organizations such as universities, government labs and industrial R&D labs (Pelz and Andrews, 1976), the change we are focusing on here is the growth of the project team, which is taking on organization-like characteristics. This transformation of scientific

☆ We would to thank the Competence Center for Bibliometrics for the German Science System (KB) for providing access to their Web of Science database for constructing the novelty measure. The novelty measure is constructed from a database derived from the 1980 to 2012 Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (AHCI), Conference Proceedings Citation Index-Science (CPCI-S), and Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH) prepared by Thomson Reuters Inc. (TR®), Philadelphia, PA, USA: ©Copyright Thomson Reuters 2013. KB is funded by the German Federal Ministry of Education and Research (BMBF, project number: 01PQ08004A). The survey data were collected with funding from Hitotsubashi University and Japan's National Institute for Science and Technology Policy. J.P. Walsh was also partially supported by Georgia Tech Ivan Allen College research grant and NSF grant (#1262418). Earlier versions of this paper were presented at 4S Annual Meeting, San Diego; Workshop on the Organization, Economics and Policy of Scientific Research, Turin; Barcelona GSE Summer Forum; SciTS Conference, Austin; and ETH Zürich. We would like to thank the conference participants, and the anonymous reviewers, for helpful comments on the paper.

* Corresponding author at: School of Public Policy, Georgia Institute of Technology, 685 Cherry St, Atlanta, GA 30332, USA. Tel.: +1 404 385 0400.
E-mail address: jpwalsh@gatech.edu (J.P. Walsh).

http://dx.doi.org/10.1016/j.respol.2014.10.007
0048-7333/© 2014 Elsevier B.V. All rights reserved.

work suggests a need to bring organization and organizational behavior theories to the study of science (Antonelli et al., 2011; Barley and Bechky, 1994; Carayol and Matt, 2004; Cummings et al., 2013; Fiore, 2008; Shrum et al., 2007).

Organizing science into research teams implies a variety of changes in the structure of the work and the work group that might affect creativity. In particular, increasing size may be associated with increasing diversity (Fiore, 2008; Harrison and Klein, 2007). Diversity can be conceptualized along a variety of dimensions, including demographic characteristics, background, and specializations (Fiore, 2008; Harrison and Klein, 2007; Page, 2007; Pieterse et al., 2013; Taylor and Greve, 2006; Williams and O'Reilly III, 1998). Furthermore, the concept of "diverse" has a variety of meanings, including separation in attitudes or viewpoints; variety of positions, categories or backgrounds; and disparity in values on some resource or asset (Harrison and Klein, 2007). In this paper, we focus on the variety of scientific fields (interdisciplinarity) and the variety of tasks in the research team (division of labor). By "variety", we mean the number of different categories represented in the team and the distribution of team members across those categories. We argue that larger teams are associated with greater field and task variety, and that teams containing greater variety in fields or tasks should have access to broader knowledge and therefore should produce more creative outputs (Hong and Page, 2004; Page, 2007; Pieterse et al., 2013; Taylor and Greve, 2006). However, larger and more varied groups may suffer from declining marginal benefits as well as a variety of process losses that decrease creativity (Andrews, 1976; Hollingsworth, 2004; Kiesler, 1969; Nooteboom, 2008). These offsetting effects of growing team size suggest that the overall impact of larger teams on creativity is not clear. Furthermore, it suggests the possibility for managerial or policy interventions that would encourage the development of some aspects of scientific teams while attempting to limit the presence of less desirable characteristics (Falk-Krzesinski et al., 2010; Fiore, 2008; Stokols et al., 2008).

Thus, the growth of team science leads to calls for application of organization and management theories of creativity to scientific work (Cummings et al., 2013; Fiore, 2008; Hackett, 1990; Vogel et al., 2013), in order to answer a key research question: how does the increasingly organized nature of scientific work affect the creativity of the research results? Since a key goal of investment in science is to produce creative outcomes, it is critical to study how the organization of scientific teams affects creativity. Much prior work on creativity has focused on individual characteristics, but in an era of team science, it is critical to study the drivers of team creativity (Harvey, 2014). In addition, we distinguish novelty from impact, which have often been conflated as proxies for creativity in the existing literature. The process of generating novel outcome and the process of those outcomes generating impact may be driven by different mechanisms, and we will analyze these processes separately in order to distinguish these components of creativity.

This paper is organized as follows. First, we discuss the two components of scientific creativity: novelty and impact. Second, we develop our theories of how novelty and impact are affected differently by team size and variety. To test these effects of team size and variety on creativity, we make use of a large-scale survey of scientific projects that collected data on team size and field and task variety. We combine these data with a new bibliometric measure of novelty based on the rarity of the reference combinations cited in the focal paper (Uzzi et al., 2013). Finally, we show how team characteristics affect novelty and how these, in turn, affect scientific impact (becoming a top-cited paper). We find that increasing team size has an inverted-U shaped relation with novelty. We also find that the size–novelty relationship is due to the relation between size and variety. On the other hand, team size has a continually increasing relation with the likelihood of a high-impact paper. In addition, while variety has a significant impact on novelty, it does not have a direct effect on impact, net of novelty. We discuss the implications of these findings in the conclusion.

2. Creativity in science

Following the definition of creativity proposed by Amabile (1983), we emphasize two aspects of a creation: its novelty and its usefulness. Correspondingly, we can discuss research teams as producing research outputs that are novel and/or useful. Psychologists have proposed diverse definitions of creativity in terms of the creative process, creative person, and creative product, but here we focus on the product definition, with novelty and appropriateness/value as the criteria for defining creative products (Amabile, 1983; Ford, 1996; Woodman et al., 1993). As Amabile (1982) pointed out, a major obstacle to creativity studies is translating the conceptual definition of creativity into an operational one in order to allow empirical assessment of creativity. Prior work has suggested a variety of indicators to categorize the creativity of artistic or scientific output: Nobel laureates as an indicator of eminent scientist (Zuckerman, 1967), prestigious prizes to identify path-breaking discoveries in biomedical research (Hollingsworth, 2004), surveying experts to nominate highly creative accomplishments (Heinze et al., 2009), financial success and critics' reviews for Broadway musicals (Guimera et al., 2005; Uzzi and Spiro, 2005), resale value of comic books (Taylor and Greve, 2006), citation counts for patents (Fleming, 2001; Fleming et al., 2007; Singh and Fleming, 2010), journal impact factor for collaboration teams (Guimera et al., 2005), and publications and citations to measure creativity of scientists (Simonton, 1999, 2004). These methods share the characteristic that creativity is assessed by experts, consumers, users, or peers, which leans towards an ex post measure of impact.

However, it is important to find indicators that allow us to unpack the concepts of novelty and impact, in order to better understand the drivers of creativity. Although the large-scale evaluation of relative creativity is generally based on the impact of those outcomes, much of the theory of creativity is built on trying to understand what leads to novel outcomes. For example, one stream of research views creativity as an evolutionary search process across a combinatorial space and sees creativity as the novel recombination of elements (Franzoni, 2010; Nelson and Winter, 1982; Schumpeter, 1939; Simonton, 2003). For example, Fleming (2001) argued that patents that combine patent subclasses that have not been combined before can be thought of as creative combinations. Similarly, Uzzi et al. (2013) argued that scientific papers that draw on unusual combinations of journals in their references can be thought of as representing relatively more novel knowledge. This work focuses the measure of creativity on the novelty of the research output.

However, at the same time, there is also substantial work focusing on the impact of the research. From Merton's perspective, citation serves as an elementary building block of the science reward system, and therefore can be viewed as a good proxy for scientific creativity. For a paper, acceptance for publication indicates an acknowledgment of its original contributions to science from peers in the field, but being cited further indicates the peer-recognition of its value and its impact on the scientific community (De Bellis, 2009; Merton, 1973; Simonton, 2004). Prior studies showed that the majority of Nobel laureates were amongst the top 0.1% most-cited authors (Garfield, 1973), and the number of citations was more significant than the number of publications in predicting receipt of awards, appointment to prestigious academic departments, and being widely known in the scientific community (Cole and Cole, 1967).

We argue that although creativity is composed of the novelty of the outcome and its impact, those are distinct concepts having their own causal relationships (Yong et al., 2014). More novel papers and patents which explore new combinations of knowledge space, carry a research stream into more unknown territory, or use a large number of knowledge domains may have high impact (Newman and Cooper, 1993; Schoenmakers and Duysters, 2010), while some contrary studies showed that patents that use unusual reference profiles or draw from distinct knowledge sources do not have high impact (Dahlin and Behrens, 2005; Nemet and Johnson, 2012). Some work in psychology even suggests a bias against novelty, such that novel ideas may have difficulty being recognized and taken up by others (Mueller et al., 2011). These prior studies suggest that impact is an important component of creativity, and also may be related to, but separate from, novelty.

Although the concepts and indicators of novelty and impact can be seen to have different meanings, novelty and impact have often been used as interchangeable proxies of creativity or have been discussed by the same theory, generally that an individual or team creates outputs with high impact because they are more likely to produce novel outputs or that high impact represents novelty (Fleming et al., 2007; Singh and Fleming, 2010; Wuchty et al., 2007). But, it can also be the case that very novel outputs are not so useful, if they cannot be integrated into existing paradigms and techniques. Sometimes, more conventional outputs are more popular, resulting in high impact. Therefore, creating novel outputs and producing outputs with high impact may have different mechanisms in team science (Fiore, 2008).

Prior work in the information processing model (and related work in the categorization–elaboration model (CEM)) argues that variety of knowledge available in the team may increase the novelty of the ideas (Page, 2007; Simonton, 2003; Taylor and Greve, 2006; Van Knippenberg et al., 2004). Thus, one advantage of teams over individuals is that larger groups may have access to a broader information set (on average). However, based on this prior work, we argue that it is not size per se, but rather the variety in the team that generates the information advantages that produce more novel outputs (Page, 2007; Simonton, 2003; Taylor and Greve, 2006; Van Knippenberg et al., 2004). Building on this information processing perspective, we review the literature on team size, variety and creative performance, and develop testable hypotheses on the drivers of novelty and impact. To distinguish novelty and impact in science, we will test the theories of creativity for novelty and impact and show how they may reflect different underlying mechanisms.

3. Size, variety and creativity

Collaboration is argued to play a critical role in scientific creativity (Vogel et al., 2013; Wuchty et al., 2007). By pooling together different expertise and perspectives, collaboration contributes to cross-fertilization of ideas and enables combining different pieces of knowledge to make something novel and useful (de Solla Price, 1986; Katz and Martin, 1997; Melin, 2000). Furthermore, collaboration allows members to build off of others' ideas to create new knowledge, so that collaborative team knowledge is greater than the sum of knowledge possessed by each individual (Hargadon and Bechky, 2006; Hudson, 1996). In other words, teamwork in science should contribute to higher creativity. However, the relative performance of teams also depends on the particular team characteristics (Heinze et al., 2009; Hulsheger et al., 2009; Stewart, 2006). We will focus on three team characteristics: team size, field variety (interdisciplinarity), and task variety (division of labor) (Taylor and Greve, 2006; Van Knippenberg et al., 2004).

3.1. Team size and novelty

The effects of team size on creativity and performance have been extensively studied (Cohen and Bailey, 1997; Hulsheger et al., 2009; Stewart, 2006). On the one hand, scholars emphasize the positive effect of team size. Larger team size is associated with a larger pool of resources and correspondingly higher ability to manage external uncertainties (Aldrich and Pfeffer, 1976; Pfeffer and Salancik, 1978). This resource advantage is particularly important for groups working on complex tasks in very uncertain environments (Cohen and Bailey, 1997; Stewart, 2006), and one feature that distinguishes science from other types of working environment is its high task uncertainty and continual novel production (Whitley, 1984). In terms of the effect of team size on creativity specifically, as team size increases, the amount of information and knowledge available to the team increases, so the team is more likely to generate new and useful ideas (Gooding and Wagner III, 1985; Hulsheger et al., 2009; Manners, 1975; Taylor and Greve, 2006). On the other hand, many social psychology studies point out the "process loss" associated with large teams, which leads to the underperformance of teams (Cummings et al., 2013). For example, larger teams have lower consensus (Hare, 1952; Manners, 1975), higher coordination costs (Ven De Ven et al., 1976), more free-riding (Fleishman, 1980), more emotional conflicts (Amason and Sapienza, 1997), and lower quality of group experience (Aube et al., 2011). All these effects can prevent a team from capitalizing on its resource advantages. These positive and negative team size effects on creativity suggest that larger teams may contribute to team productivity (defined as performance measured in absolute terms) but not to team efficiency (defined as performance measured in relative terms, e.g., output per capita) (Gooding and Wagner III, 1985; Hulsheger et al., 2009). These prior studies suggest that larger teams may do better at generating a larger set of potentially interesting ideas, but be less able to evaluate and select among these ideas to find the best to pursue. In addition, Jordan (2006) argued that as team size increases, there is greater need for control mechanisms, which may include more bureaucratic structures, and these may dampen creativity. Therefore, larger teams may contribute to high novelty, but the effect may diminish as the size increases and even decline above a certain threshold due to the various process losses and conflicts described above.

H1. Increase in team size will have an inverted-U shaped relation with novelty.

3.2. Variety and novelty

The diversity literature has studied various dimensions of diversity, such as race/ethnicity, gender, tenure, nationality and function/education (Ancona and Caldwell, 1992; Jehn et al., 1999; Milliken and Martins, 1996; Page, 2007; Pelled et al., 1999; Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998), and different conceptualizations of diversity, such as separation, variety, and disparity (Harrison and Klein, 2007). This paper focuses on knowledge variety that grows out of different disciplinary backgrounds and different work tasks (Taylor and Greve, 2006). Prior work on the information processing or CEM perspective suggests that these task-related dimension of diversity (including knowledge background and task specialization) should increase the task-related information and perspectives available to the group and hence may be especially salient for predicting performance, especially in domains that require high levels of information processing and originality, such as R&D or scientific teams (Hulsheger et al., 2009; Stewart, 2006; Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). Diversity of information can be proxied by the variety in educational background (fields) or specialty (tasks) (Williams and

O'Reilly III, 1998). In this paper, we examine both field variety and task variety of the team to test the effects of variety on creativity.

There has been a strong policy push to encourage interdisciplinary teams and to understand their impacts and how to manage them better (Falk-Krzesinski et al., 2010; Fiore, 2008; Stokols et al., 2008; Vogel et al., 2013). Teams made of those from different disciplines (field variety) are seen to have information processing advantages. Each discipline is grounded in a set of knowledge and heuristics, and these can provide the foundation for novel perspectives, problems and solutions (Page, 2007; Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). For example, in measuring the creativity of comic books, Taylor and Greve (2006) showed that the greater the number of genre backgrounds represented in comic book creator teams, the higher the innovativeness of the team, even net of size. Ancona and Caldwell (1992) found that teams with different functional backgrounds engage in more outside communication, and are rated as more innovative. In science, there is evidence that field variety improves scientific creativity (Hollingsworth, 2006; Rafols and Meyer, 2010; Yong et al., 2014). For example, Hollingsworth argued that scientific organizations that generate many breakthrough discoveries in biomedical research are characterized by moderately high levels of field variety, as well as organizational structures that facilitate cross-field communication and integration of ideas. In a study of environmental sciences, Steele and Stier (2000) found that articles that draw information from a diverse set of journals are cited with greater frequency than articles having smaller or more narrowly focused bibliographies. However, as field variety increases, we would expect declining marginal benefit from adding additional variety, suggesting a curvilinear relation between field variety and creativity (Williams and O'Reilly III, 1998). Larivière and Gingras (2010) found that articles with very low and very high levels of interdisciplinarity have lower citations, suggesting that medium levels of field variety may be optimal. While the information processing perspective suggests a benefit from interdisciplinary teams, there is some counter evidence. For example, Cummings et al. (2013) found that field variety is associated with lower productivity. At the same time, as cognitive diversity of the project increases, participants may have a more difficult time in communicating and absorbing the potential contributions of cognitively distinct members of the team (Cohen and Levinthal, 1990). For example, Nooteboom's cognitive-based theory of inter-firm alliances argued that cognitive distance between alliance partners should have an inverted-U shaped relation to, or at least diminish the marginal effect on, learning and innovation (Nooteboom, 2008). As variety becomes very high, it becomes more difficult to ensure that members can communicate easily and convert specialized expertise into a novel final product (Cummings et al., 2013; Fiore, 2008; National Academy of Sciences, 2004).

H2a. Field variety increases novelty but at a decreasing rate.

In addition, within scientific teams, we might find variety not only in fields, but also in tasks. Prior work in the study of organizations suggests that as work group size increases, the division of labor becomes more elaborate (Blau, 1970; Child, 1973; Meyer, 1972), although this effect may be weaker for very small organizations (Blau et al., 1966; Evers et al., 1976). We might expect a similar relation for scientific teams, as collaborators specialize in particular aspects of the research task, for example, having some concentrate on data collection (with further subdivision into specific aspects of data collection), others on statistical analysis, and still others on designing experiments, or integrating findings into a research report (Latour and Woolgar, 1979; Shibayama et al., 2013). In larger teams, this division of labor could be quite elaborate, as we see in large experimental physics collaborations (Brown and Ross, 2013; Traweek, 1988). Division of labor offers many opportunities for gains from specialization, and hence may be associated with more innovation (Smith, 1776). Task variety in the group should produce diversity in information available for idea generation and problem solving in the team, leading to greater novelty.

At the same time there are concerns about ossification into specialized roles, making creativity difficult. High task variety teams may face a problem of integrating or synthesizing multiple specialized tasks and hence have difficulty creating a final novel output (Hong, 2008). Simple accretion of specialized tasks can disperse the focus of the work and the project may have difficulty synthesizing these disparate components into an integrated work (Clark et al., 1987; Merton, 1940). This will be also the case for field variety. In addition, increasing specialization raises concerns about the creation of overly specialized "sub-scientists" who are not able to integrate diverse information across roles (Hackett, 1990; Walsh and Lee, 2013). Moreover, task variety creates within-group boundaries between different functions, which is an obstacle for a more integrated and iterative idea generation and evaluation process, which may be needed to nurture creativity (Harvey and Kou, 2013; Van Knippenberg et al., 2004). For example, Hong (2008) described the conflicts in a geophysics lab between those who specialize in analyzing samples, with deep tacit knowledge of the equipment, and those who interpret results, with greater command of the theoretical literature. Task variety and focus on the problems of the specialty can even lead to goal displacement, such that the demands of the task supersede the goals of the project (Merton, 1940). Thus, although research teams can benefit from field and task varieties, these can cause problems of integration across the diverse team members (Blau, 1970; Cummings et al., 2013; Jordan, 2006). Thus, like field variety, this combination of information benefits with possible integration problems suggests that task variety increases novelty, but with diminishing marginal returns.

H2b. Task variety increases novelty but at a decreasing rate.

The arguments above suggest that the primary effect of size on novelty is that larger groups tend to be more varied. Thus, the size effects noted above may be primarily due to the relation between size and knowledge variety (i.e., field and task variety), so that if we control for knowledge variety, there may be little added effect from size (Taylor and Greve, 2006). Therefore, knowledge (field and task) variety should more directly affect novel combination of knowledge, or novelty, and thereby, having a more dominant effect than team size on novelty.

H3a. Field variety will mediate the effect of team size on novelty.

H3b. Task variety will mediate the effect of team size on novelty.

3.3. Team size, variety, and impact

In addition to its effect on novelty, team size should have additional benefits that raise the impact of the paper produced. It takes several steps for a novel idea to have an impact, including idea generation, selection, retention, implementation, and knowledge diffusion, and previous studies suggest that these different steps may respond differently to particular group characteristics (Baer, 2012; Fleming et al., 2007; Lavie and Drori, 2012; Obstfeld, 2005; Singh, 2005; Singh and Fleming, 2010). First, as argued above, larger teams may be better at producing greater and more novel collective knowledge, thereby producing a more influential output (Wuchty et al., 2007). Furthermore, the resource advantages associated with large teams contribute to not only generating and successfully publishing novel ideas, but may also raise the quality of the paper on dimensions other than novelty, because of within-team peer review and filtering (Singh and Fleming, 2010). While novelty is related to the collective activity in a team, impact may depend on accretion of team members' individual networks or reputation

as well as the novelty of the output (Fleming et al., 2007; Singh and Fleming, 2010). In addition to more self-citations, larger teams may also bring more citations by having a larger network of colleagues and through more presentations at conferences (Bentley, 2007; Valderas, 2007). Katz and Martin (1997) found a correlation between the number of authors and the citation count of papers and argued that more authors increases the chance that the paper will be found by those searching for work related to any given author, increasing the impact of the paper through the aggregations of these reputational networks. Therefore, impact may be more related to how many networks are aggregated by members and established as the distribution routes for the finding rather than to managing the internal dynamics in a team. Larger team size implies more networks or more opportunities for disseminating the team's output. In addition, due to the Matthew effect in science, that is, the cumulative advantage or self-reinforcement in the accumulation of recognitions, team size may raise impact in a super-linear fashion (Katz, 1999, 2000; Merton, 1968; van Raan, 2006a,b). Therefore, unlike the case for novelty, larger team size should have a continuously positive effect on impact (in other words, that the square term would also be non-negative). Based on these arguments, we hypothesize that:

H4. Increase in team size will have a continually increasing effect on the likelihood of a high impact paper, net of novelty.

Since size is also a proxy for the span of the aggregated network and resources for disseminating the finding, even if we control for novelty and variety in the team, we still expect a direct effect of team size on impact from this network effect. This leads to our fifth hypothesis, about size, variety and impact.

H5. The effect of size on impact dominates over variety, net of novelty.

Put differently, we can think of size as a complex indicator of multiple constructs, representing both the variety of knowledge and the breadth of network. The information processing advantage of team size is mediated by field or task variety, while the network advantage of team size is independent of field and task variety.

4. Data and methods

We use a survey of scientists in the US to test these hypotheses. The survey provides information from a large sample of projects spanning fields and institution types. The population of interest is research-active faculty in the fields of science covered by Thomson Reuters Web of Science (WoS). Here, a field is defined by the field of the journal where the paper is published, as defined by the WoS classification. However, we can collect information about researchers' individual fields of expertise from the survey. This allows us to measure field variety of project teams distinctly from the discipline of the publication (as the field of the project is defined by the journal in which the results were published, while field variety is measured by the composition of the team that produced the result). The survey began with a sample of 9428 publications, covering publication years 2001–2006, stratified by 22 WoS journal fields (see Appendix) and by forward citations, with an oversampling of the papers in the top 1% of citations in each field in each year (citation counts retrieved December 31, 2006). About 3000 of the sampled papers were in the top-cited papers and about 6000 were from other random papers.

The list of publications was searched for an appropriate contact author, beginning with the reprint author, followed by the first or last author (depending on the name ordering conventions in that field), and then going through the list of authors to find a US author for whom a current address (email or, if none available, post mail) could be found. In about 80% of cases, the respondent was either the contact author or the first or last author. In cases where no valid contact was available (for example, the author was deceased, or had moved out of the country), we excluded those cases from the sample. Furthermore, to reduce respondent burden, for those scientists that appeared more than once in our sample, we randomly sampled one paper, giving priority to the top-cited papers. This process led to a total of 8864 papers. We received at least partial responses from 3742 scientists (42%), with 2327 completed responses (26% response rate).[1] For this analysis, we limit responses to those in universities and hospitals and with at least two authors on the paper, including all of science, engineering and social science ($N = 1493$). We categorize papers into 20 fields, based on the WoS journal field classifications (with "Economics & Business" merged into "Social science, general" and "Multidisciplinary" journal field papers assigned to one of the 20 according to the main field of the references in the paper). We use survey estimation methods to control for the differential sampling and response rates between top and random papers and across fields. The survey weights used are based on the overall population of publications, so that weighted means account for the underlying population distributions on field and top vs. random papers (Kalton, 1983). All statistics in this paper are calculated taking into account the sampling structure and weights (Lee et al., 1989).

The survey asked the respondent to describe the research project that produced the sampled paper (which was named on the cover of the survey). This strategy allows us to link bibliometric, survey and institutional data. We also collected additional data (number of authors and references) from WoS. We used data from the 2007 AUTM survey to get university-level data on size of research effort in the respondent's organization (total R&D expenditure) (Association of University Technology Managers, 2007) and the National Research Council (NRC) graduate school ranking report to get a department ranking score (National Research Council, 1995). We use the survey data, WoS, AUTM, and NRC data to generate the following measures of our dependent, independent and control variables.

4.1. Dependent variables

4.1.1. Novelty

Novelty in scientific work is operationalized by a measure, adapted from Uzzi et al. (2013), that treats the novelty of a paper as the rareness of its pairwise combinations of prior work, i.e., references. This method is built from the observations that combining certain pairs of knowledge domains is rarer or more novel than combining other pairs, and that a given paper draws from a variety of prior knowledge, producing a knowledge combination distribution profile for each paper. This profile can then be characterized by the relative rareness or commonness of various cut-points on the distribution, such as the median or the 10th percentile. For

[1] A detailed non-response bias analysis shows that respondents and non-respondents are similar on most observable indicators, including citation count and likelihood of being a highly-cited paper, number of authors, multiple-organization, and publication year. We find a small bias for novelty, which, though statistically significant, shows that respondents' papers are, on average about one-tenth of a standard deviation less novel. The only major difference between respondents and non-respondents is that clinical medicine researchers are less likely to respond, although even clinical medicine had a response rate over 20%. As a further test for response bias, we tested a simple regression model on all 9428 cases in our sample using only bibliometric data, to test H1 (the novelty effect of size and size square) and H4 (the impact effect of size and size square), controlling for field and the number of references (and novelty for H4). We find that, in the full sample, size has an inverted-U relation with novelty (H1) and a continuously increasing relation with impact (H4), consistent with our survey results (results available from the contact author). This multivariate test for response bias gives us further confidence in the representativeness of our survey sample.

example, Uzzi et al. (2013) characterized the novelty and conventionality of a paper by the two cut-points of the 10th percentile and the median, respectively. They use the relative rarity of the 10th percentile to measure the "novelty" of the paper (and refer to the commonness of the median combination as the "conventionality" of the paper).

Building from Uzzi et al. (2013), we use the following two-step process to construct our novelty measure: (1) calculating the commonness of co-cited journal pairs for the whole WoS database and (2) calculating the novelty of papers based on their references for our sampled papers. This method has the advantage over Uzzi et al.'s measure in being substantially less computationally intensive.

The first step implements the following procedure: (1) retrieve all papers indexed in WoS, (2) retrieve all references for each paper, (3) list all pairwise combinations of references for each paper, (4) record the two paired journals for each reference pair, (5) record the publication year t for each paper, and (6) pool together all journal pairs of papers published in the same year to construct a universe of journal pairs for each year. We refer to this universe of journal pairs in year t as U_t. Then we define the commonness of each journal pair (journal i and j) in year t as:

$$\text{Commonness}_{ijt} = \frac{\text{observed number of pairs}_{ijt}}{\text{expected number of pairs}_{ijt}} = \frac{N_{ijt}}{\frac{N_{it}}{N_t} \cdot \frac{N_{jt}}{N_t} \cdot N_t} = \frac{N_{ijt} \cdot N_t}{N_{it} \cdot N_{jt}}$$

where N_{ijt} is the number of i–j journal pairs in U_t, N_{it} is the number of journal pairs which include journal i in U_t, and N_t is the number of all journal pairs in U_t. Therefore, $\frac{N_{it}}{N_t}$ is the probability that journal i appears in U_t, $\frac{N_{it}}{N_t} \cdot \frac{N_{jt}}{N_t}$ is the joint probability for the co-appearance of journal i and j, and $\frac{N_{it}}{N_t} \cdot \frac{N_{jt}}{N_t} \cdot N_t$ is the expected number of i–j journal pair.

The second step calculates novelty at the paper level. For our sampled papers, we repeat procedure (2–5) in the first step. Then for a paper published in year t, we record the $commonness_{ijt}$ for each of its cited journal pairs. This produces a series of commonness values, and subsequently we sort these numbers and record the 10th percentile as an indication of the commonness at the paper level. In addition, taking the 10th percentile instead of the minimum reduces the noise and improves the reliability of this measure. Uzzi et al. (2013) also tested the 1st and the 5th percentiles and demonstrated that results are robust to different cutoffs. Furthermore, we take the natural logarithm transformation to get a roughly normally distributed variable for commonness at the paper level. Then we add a minus sign to this commonness variable to give the final measure for paper novelty, since novelty is the opposite of commonness.

To test the robustness of this measure, we also constructed the novelty variable using the following variations: (1) adopt a three-year time window instead of a single year, that is, U_t pools together journal pairs from year $t-2$ to t, instead of just year t; (2) exclude 25% of the least cited journals in case results are sensitive to rare pairings.

4.1.2. Top 1%

Following much prior work in this area, we measure impact in terms of forward citations (Martin and Irvine, 1983; Moed, 2005; Wang, 2014). To measure the impact of the paper, we use a dummy variable representing "high" impact, where high impact is operationalized as being in the top 1% most cited papers in that WoS field in that year, based on citation counts as of December 31, 2006. While there is some concern about the validity of citations as a measure of impact (De Bellis, 2009), we follow Uzzi et al. (2013) and argue that if a paper is in the very top of the citation distribution (in this case, the 99th percentile), it can be considered a high impact paper. In addition, considering the errors in using short citation time windows for identifying top cited papers (Rogers, 2010; Wang, 2014), we adopt alternative impact measures: (1) five-year moving time window citation counts and (2) citation counts up to April 2013, for robustness checks.

4.2. Independent variables

4.2.1. Team size

Our operational definition of team size is natural log of number of authors, collected from WoS: minimum in our sample is 2 authors and the maximum is over 100. We also test the robustness of our results to alternative measures of size, including: (1) the log of the total number of people on the project team including both co-authors from WoS and non-co-authors (e.g. post-doctoral researchers, technicians and graduate students) from the survey, and (2) the log of project research funds, collected from the survey.

4.2.2. Field variety

We use information from the survey asking the fields of expertise out of 29 different fields, roughly equivalent to the university department level, associated with co-authors (up to 7) for the paper[2] and create a Blau index as a measure of *field variety*, i.e. $1 - \sum p_k^2$ where p_k is the proportion of members in the kth field category (Harrison and Klein, 2007; Pieterse et al., 2013). We also create a coarser Blau index using 8 aggregate fields of expertise, roughly equivalent to the NSF directorate level, and test our models to compare between fine and coarse measures of field variety (see Appendix for the list of fields). As an additional robustness check, we use the count of the number of fields (out of 29) represented in the team (Taylor and Greve, 2006).

4.2.3. Task variety

The overall task of the project can be divided into specialized responsibilities that require specialized knowledge (Blau et al., 1966), which comprise differences in task-related knowledge for the project (Katz and Martin, 1997). We measure *Task variety* using a set of survey items that capture internal and external task variety. Internal task variety is on a 5-point scale (from "not at all" to "very much"), measuring if the project "involved a strict division of labor with each person responsible for a specific part of the research". External task variety (i.e. *outsourcing*) is on the same 5-point scale, measuring if "the project involved outsourcing parts of the work to other research groups". We create a task variety variable from the sum of internal task variety and outsourcing scores. We also test the effects of internal task variety and outsourcing separately.

4.2.4. Control variables

The overall levels of novelty and impact, as well as the tendency to engage in field- or task-varied research teams, may vary by size of the organization or by whether it is public or private (Cullen et al., 1986; Hollingsworth, 2004). We control for size of the respondent's

[2] In our sample 84% of cases have 7 or fewer authors and so we have complete information on our continuous measure of field variety. We also create a binary measure (multiple fields or not) for robustness. If we use the binary measure, then of those with over 7 authors, 63% are coded as a multi-field team, based on the answers to the 7 reported authors. Using this binary measure, only 6% of cases are possibly affected by missing information on the remaining authors (with some of those possibly miscoded as not multi-field when they are in fact). Our findings are robust to excluding these cases with potential measurement error.

Table 1
Descriptive statistics.

	Variable	N	Mean	SD	Min	Max	Correlation											
							1	2	3	4	5	6	7	8	9	10	11	12
1	Novelty	1463	−0.18	2.06	−10.64	5.83	1.00											
2	Top 1%	1493	0.02	0.12	0.00	1.00	**0.02**	1.00										
3	Team size	1493	1.34	0.52	0.69	5.33	**0.19**	**0.08**	1.00									
4	Field variety (fine)	1489	0.24	0.27	0.00	0.88	**0.25**	**0.02**	**0.38**	1.00								
5	Field variety (coarse)	1489	0.17	0.24	0.00	0.83	**0.26**	0.01	**0.32**	**0.83**	1.00							
6	Task variety	1430	4.45	1.95	2.00	10.00	**0.18**	0.00	**0.29**	**0.21**	**0.19**	1.00						
7	Internal	1457	2.84	1.34	1.00	5.00	**0.14**	0.00	**0.14**	**0.15**	**0.15**	**0.78**	1.00					
8	Outsourcing	1442	1.62	1.23	1.00	5.00	**0.14**	0.01	**0.30**	**0.18**	**0.14**	**0.73**	**0.15**	1.00				
9	Multi org auth	1493	0.50	0.50	0.00	1.00	−0.05	**0.05**	**0.23**	**0.10**	**0.07**	**0.10**	0.02	**0.15**	1.00			
10	Univ. size (R&D)	1118	19.88	1.10	16.03	22.11	0.01	**0.03**	0.05	0.01	0.04	0.03	0.01	0.04	0.02	1.00		
11	Public	1215	0.67	0.47	0.00	1.00	**−0.16**	**−0.03**	**−0.08**	−0.07	**−0.08**	−0.01	0.00	−0.02	−0.02	**0.11**	1.00	
12	Dep. NRC rank	1493	1.75	1.82	0.00	4.96	**0.07**	**0.03**	−0.02	−0.02	−0.01	0.03	0.02	0.02	−0.06	**0.29**	**−0.11**	1.00
13	No. of WoS ref	1493	23.88	16.36	0.00	251.00	**0.36**	**0.05**	**0.24**	**0.10**	**0.10**	**0.10**	0.07	**0.10**	0.04	0.04	**−0.10**	**0.12**

Note: Correlation with bold numbers at $p < .05$.

organization using the log of total R&D expenditure, collected from the 2007 AUTM US licensing activity survey as a measure of the size of the research effort in that university (*university size*). We also classify universities into public or private and create a binary variable where it is 1 if the university is *public*. Additionally, in order to make our external division of labor measure a cleaner measure of task variety (as distinct from remote collaboration), we control for whether the authors are all in the same organization or more than one organization (*multi org auth*), as prior work has suggested that remote collaborations suffer from higher coordination costs (Cummings and Kiesler, 2005).[3] Because the rank of the respondent's department may affect novelty and impact (halo effect), we also control for these using data from the National Research Council (NRC)'s graduate school ranking report, matching our respondents' field and institutional affiliation to NRC's field and affiliation records. Novelty and citation characteristics can also vary by field. Therefore, we control for 20 journal fields where *chemistry* is the reference category (see Appendix). Lastly, we control for the number of WoS references. Because our *novelty* variable is measured from each paper's references registered in WoS, the number of WoS references may be related to the value of *novelty*.

5. Results

5.1. Descriptive statistics

Table 1 summarizes the measures and their correlations. We, first, see that mean *team size* is 1.34 (i.e. about 4 authors). The mean *field variety* is .23, which is a relatively low number and implies unequal spread of fields in a project team on average. We also find that the mean count of different fields is 1.7 and median is 1 (meaning less than half of the teams are multi-field). The mean *task variety* is 4.5, ranging from 2 to 10 (the full range possible). We also see that team size, field variety, and task variety are correlated with novelty, but only size and field variety are correlated with impact. Also, novelty and impact have a significant correlation.

5.2. Team size, knowledge variety and novelty

Given the importance of the size effect in science, our first research question is how novelty in scientific work responds to the size of the project team. Table 2 estimates the effect of team size (measured as log(author count)) on *novelty* using OLS.[4] Column 1 shows that team size has a positive effect on novelty, consistent with many prior studies. Although team size has a positive effect on novelty, we argue in Hypothesis 1 that the relation should be curvilinear (inverted-U), due to declining marginal benefit and to process losses in large teams. To test this hypothesis, we estimate the model with a squared term of team size in column 2 and find a significant negative effect for the square of team size, consistent with our inverted-U hypothesis.

We also argue that knowledge variety in teams can broaden the search across various knowledge domains and increase the chances of creating novel combinations. However, as argued above, these information processing benefits may have declining marginal effects (Blau, 1970; Cummings et al., 2013; Nooteboom, 2008). Column 3 in Table 2 shows that our fine-grained measure of field variety (Blau index based on 29 fields) has a direct positive effect on novelty. Its squared term in column 4 is not significant. Columns 5 and 6 show the same models using our coarse-grained measure of field variety (8 field Blau index), which gives similar results. Thus, we find some support for Hypothesis 2a, with the main effect of field variety having a positive relation with novelty, but not a significantly negative effect on the square term. Columns 7 and 8 show that task variety has a direct positive effect on novelty and a significantly negative square term, suggesting a declining marginal effect. As we described earlier, our task variety measures consist of internal task variety and outsourcing as external task variety. We test each component of our task variety measure separately in columns 9–12, and see consistent effects except the squared term of outsourcing losing significance in column 12, although still having a negative direction. Therefore, the results generally support Hypotheses 2b.

[3] One may think outsourcing (external task variety) and having authors in more than one organization are the same concept and that both measure the difficulty of coordination and communication. Although they are correlated, these are not identical concepts. For example, although all authors on the paper are in the same organization, they may outsource some assays in the experiment (in biology), or might outsource survey execution to a survey organization (in sociology). While such external contributors might be a critical part of task variety of the project, they often do not have authorship. In our data, *outsourcing* and *multi org auth* are correlated at .15 (See Table 1). Out of projects whose authors are all in the same organization, 17% does outsourcing, and out of projects whose authors are from more than one organization, 70% report they do not outsource parts of their work. Hence, we interpret our measures of outsourcing and mutli-org authorship as measuring distinct concepts, and control for multi-org authorship in our models.

[4] As we described in the data section, we limited responses to projects with at least two authors (i.e. subpop obs in Table 2) because team variety has meaning when there are at least two people. However, we estimate SEs considering the unconditional, full population (i.e. observations in Table 2) to avoid biased results, because subpopulation sizes within strata are random and the true subpopulation size is not known, and needs to be estimated in the full population (West et al., 2008).

Table 2
OLS regression of novelty on size, field variety and task variety.

Variables	Novelty OLS														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Team size	0.27*	1.08**											0.05	0.21	0.02
	(0.15)	(0.53)											(0.14)	(0.15)	(0.15)
Team size squared		−0.28*													
		(0.15)													
Field variety (fine)			1.25***	1.21									1.22***		1.14***
			(0.26)	(0.85)									(0.26)		(0.26)
Field var. sq (fine)				0.07											
				(1.29)											
Field variety (coarse)					1.47***	2.08**									
					(0.28)	(0.85)									
Field var. sq (coarse)						−1.07									
						(1.30)									
Task variety							0.11***	0.42**						0.10**	0.08*
							(0.04)	(0.18)						(0.04)	(0.04)
Task variety squared								−0.03*							
								(0.02)							
Internal									0.13**	0.83***					
									(0.06)	(0.27)					
Internal sq										−0.12***					
										(0.04)					
Outsourcing											0.13**	0.67**			
											(0.06)	(0.34)			
Outsourcing sq												−0.10			
												(0.06)			
Multi org auth	0.02	0.01	0.00	0.00	0.02	0.01	0.03	0.05	0.10	0.09	0.02	0.01	−0.01	−0.02	−0.03
	(0.15)	(0.15)	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)	(0.14)	(0.14)	(0.15)	(0.15)	(0.14)	(0.15)	(0.14)
Univ. size (R&D)	−0.02	−0.02	−0.01	−0.01	−0.02	−0.02	−0.03	−0.02	−0.01	−0.01	−0.03	−0.03	−0.01	−0.03	−0.03
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Public university	−0.43***	−0.41***	−0.39***	−0.39***	−0.40***	−0.39***	−0.41***	−0.42***	−0.43***	−0.40***	−0.41***	−0.42***	−0.39***	−0.40***	−0.36**
	(0.15)	(0.15)	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.14)	(0.15)	(0.14)
Dep. NRC rank	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
No. of WoS ref	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1248	1248	1245	1245	1245	1245	1204	1204	1227	1227	1213	1213	1245	1204	1201
Subpop obs	1098	1098	1095	1095	1095	1095	1054	1054	1077	1077	1063	1063	1095	1054	1051
R^2	0.26	0.26	0.28	0.28	0.28	0.28	0.28	0.28	0.27	0.28	0.27	0.27	0.28	0.28	0.30

Team size = log (authors); team size squared = [log (authors)]^2.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

On the other hand, although team size and knowledge variety have direct positive effects on novelty individually, the effect of team size on novelty should be mediated by knowledge variety, which is directly associated with information processing across diverse sets of knowledge. To analyze the mediation effect, first, we examine whether size predicts variety in field and task, regressing field and task variety, respectively, on size and all other controls. Team size has a positive and significant relationship with both field and task variety (results not shown). Next, if field and task varieties have mediating effects, the magnitude of size coefficients on novelty should be dampened significantly after including each of the variety variables as a predictor (Baron and Kenny, 1986). Comparing the coefficient on team size in column 1 with those in columns 13 and 14, we can see that the comparable coefficients all decrease after adding variety variables. To test the significance of these effects, we conducted seemingly unrelated regression estimation and then compared coefficients of team size using Wald tests (Singh and Fleming, 2010; Wooldridge, 2010). We find that field variety mediates the effect of team size on novelty with significant difference in coefficients on team size ($p < 01$), but task variety does not significantly mediate the effect of team size ($p > .10$), though still having a direct effect on novelty, which supports Hypothesis 3a and weakly supports 3b.[5] In addition, adding both variety variables together in column 15, our results still hold and both variety variables jointly mediate the effect of team size (i.e., .27 on team size in column 1 vs. 02 in column 13, $p < 01$). Column 15 also shows that the two measures of knowledge variety (field and task) are significantly associated with novelty, net of size and the other knowledge variety measure, suggesting that field variety and task variety provide distinct mechanisms for increasing knowledge variety available to the project.

5.3. Team size, knowledge variety and research impact

Novelty and impact are separate components of creativity, but also should be related to each other (Uzzi et al., 2013). To validate our measure of novelty, we examine the relationship between novelty and impact, using a probit specification, and find a positive and

[5] We also tested the variance inflation factor (VIF) for the two variety measures (based on models 13 and 14), and find that field variety has a VIF of 1.27 and task variety has a VIF of 1.14. These scores are low enough that we do not have major concerns about multicollinearity affecting our estimates of the variety coefficients.

Table 3
Probit regression of impact on novelty, size, field variety and task variety.

Variables	Top 1% Probit									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Novelty	0.06***	0.05***	0.06***	0.05***	0.05***	0.06***	0.06***	0.06***	0.06***	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Team size		0.53***	−0.35						0.61***	−0.24
		(0.07)	(0.28)						(0.08)	(0.30)
Team size squared			0.26***							0.25**
			(0.09)							(0.10)
Field variety (fine)				0.15					−0.06	0.01
				(0.12)					(0.14)	(0.14)
Field variety (coarse)					0.19					
					(0.14)					
Task variety						0.00			−0.04**	−0.04**
						(0.02)			(0.02)	(0.02)
Internal							0.01			
							(0.02)			
Outsourcing								−0.02		
								(0.03)		
Multi org auth	0.42***	0.29***	0.30***	0.41***	0.41***	0.40***	0.42***	0.40***	0.28***	0.28***
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Univ. size (R&D)	0.08***	0.07**	0.07**	0.07***	0.08***	0.08**	0.07**	0.08**	0.07*	0.07*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Public university	−0.16**	−0.17**	−0.19***	−0.16**	−0.16**	−0.16**	−0.17***	−0.15**	−0.16**	−0.18**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Dep. NRC rank	0.03*	0.04**	0.04**	0.03	0.03	0.03*	0.03*	0.04*	0.05**	0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
No. of WoS ref	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Field dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1248	1248	1248	1245	1245	1204	1227	1213	1201	1201
Subpop obs	1098	1098	1098	1095	1095	1054	1077	1063	1051	1051
F test	*F*(25, 1223)	*F*(26, 1222)	*F*(27, 1221)	*F*(26, 1219)	*F*(26, 1219)	*F*(26, 1178)	*F*(26, 1201)	*F*(26, 1187)	*F*(28, 1173)	*F*(29, 1172)
	7.14***	7.57***	7.26***	6.87***	6.86***	6.54***	6.72***	6.54***	6.79***	6.57***

Team size = log (authors); team size squared = [log (authors)]^2.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

significant effect of novelty on impact (Table 3, column 1). Controlling for novelty, column 2 shows that the effect of team size on impact is still strongly positive, net of novelty. Moreover, team size squared has a positive and significant effect on impact (column 3), consistent with the network resources argument (Katz and Martin, 1997) and supporting Hypothesis 4. Based on both results, in column 2 in Table 2 and column 3 in Table 3, we observe different effects of team size on novelty and impact. Fig. 1 illustrates these effects. The figure shows the inverted-U shaped curvilinear effect of team size on novelty across a range of team sizes (in standard deviation units) for an average team having a chemistry journal publication where authors are all in the same private university (based on Table 2, column 2). It also depicts the positive effect of team size on impact and the change in the marginal probability effect at different mean-deviated team sizes, based on the estimation in Table 3, column 3. Accordingly, Fig. 1 supports our Hypotheses 1 and 4, and further shows that there is a moment (about +1 SD from the mean, or about 7 authors) when an increase in team size decreases the expected novelty of the paper, although still increases the chance that the paper has high impact. This figure highlights the dual nature of team size, reflecting both positive and negative forces from knowledge variety (see below) and the additive effects of network resources.

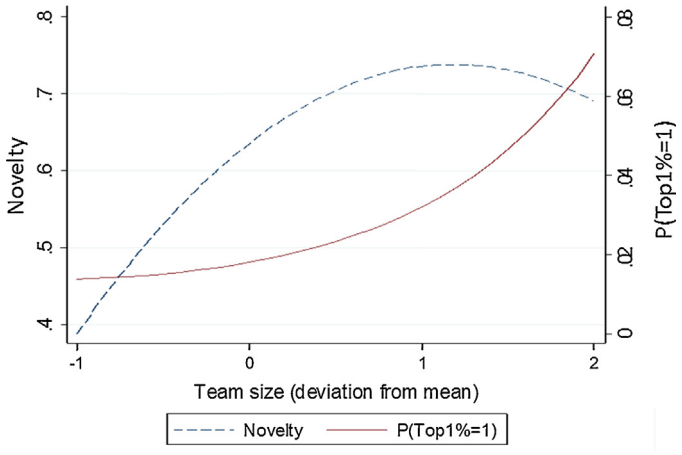

Fig. 1. Effect of team size on novelty and impact.

Knowledge variety directly affects producing a novel output as presented in Table 2. However, it may not directly affect impact of the output after a novel product is produced. The impact of a paper is determined by (1) the novelty of a paper and (2) how effectively the paper is diffused through the academic network. While having effects on the former, knowledge variety does not further affect the later directly. Columns 4–8 show that task and field varieties do not have direct effects on impact, which implies that benefits from task and field varieties are already reflected in novelty and that evaluating usefulness of the output is a distinct process. When adding team size in the model with field and task varieties (column 9), we can see that team size, associated with network size, has a dominant positive effect on impact over team variety, supporting Hypothesis 5. The quadratic term of team size is still positive and significant, net of knowledge variety (column 10).

These results suggest that we should not assume novelty and impact are driven by the same processes. Novelty seems to be primarily driven by knowledge variety of the team while impact is dominantly driven by team size, net of novelty. While impact may reflect quality of ideas, and hence novel output may be more likely to have high impact, impact may also reflect popularity or familiarity, so that team size, as a proxy for larger networks and resources, would have a greater additional effect on impact, net of novelty, than does knowledge variety.

5.4. Robustness tests

In this section, we discuss the robustness of our models in Tables 2 and 3, using alternative measures.[6] First, we did a robustness check using different size measures. Team size can be measured differently by the physical capacity of a team, the number of people available to a team for work, and resource autonomy such as net assets (Kimberly, 1976). As additional measures of team size, we tested (1) the log of the total number of people on the project team: including both co-authors from WoS and non-co-authors (e.g. post-doctoral researchers, technicians and graduate students) from the survey, and (2) the log of project research funds, collected from the survey. The results are qualitatively consistent although some significance levels change. The only results that are not robust are: (a) we do not find an effect for the squared terms of these two alternative size measures on novelty or impact, suggesting that the conventional research team measure as the number of authors captures the inverted-U size-novelty and the additive network effect on impact more clearly, and (b) when we control for size using the log of funds, both funding and knowledge variety variables are still significant predictors of novelty, suggesting that funding measures some aspects of size other than knowledge variety. Second, we tested the effect of field variety using counts of different fields involved in the project (out of 29 fields) as an alternative measure of field variety (Taylor and Greve, 2006) and obtained significantly consistent results. Moreover, since our field variety measures, i.e., Blau index and counts of fields, are based on the information for up to 7 authors, we retested our models limiting to papers with no more than 7 authors and obtained robust results except the effects of squared size on novelty and impact, which is not surprising, since we truncated size at about the bliss point (Fig. 1). Those squared terms still showed consistent direction, but lost significance in this sample of projects with 7 or fewer authors. Third, we tested different variations of our novelty measures. Our novelty measure is −1 times the log of the 10th percentile value of reference-pair-based commonality in year t for all journals. We also constructed this measure excluding 25% of the least cited journals and/or using a three-year window, that is, $t-2$ to t. The novelty measure with year t and excluding 25% of the least cited journals, that with year $t-2$ to t for all journals, and that with year $t-2$ to t and excluding 25% of the least cited journals, all showed consistent results with no change in significance. Fourth, we use (1) a five-year time window citation counts and (2) citation counts up to April, 2013 as alternative measures of impact, and test our models adding controls for publication years and using negative binomial regressions. All results are consistent with those in Table 3 except the squared terms of team size has weaker significance ($p < .15$). Overall, our hypotheses are largely robust to alternative measures of all our dependent and independent variables.

Lastly, there may be some concerns about the causal direction of these relationships. Traditionally group studies treat group characteristics as exogenous variables and investigate their effects on group performance (Cohen and Bailey, 1997; Hulsheger et al., 2009; Stewart, 2006). However, the recent prevalence of self-assembled project teams may challenge this traditional approach (Contractor, 2013; Zhu et al., 2013). Collaborative teams in science are typically self-assembled project teams, and scientists have substantial autonomy to create, maintain, and dissolve collaborative teams. For these self-assembled project teams, "projects are both the goal of, and the reason for, the existence of project teams" (Zhu et al., 2013). In other words, the process of idea generation precedes or co-evolves with the process of team assembly. This view can raise a concern that to create novel output, people assemble teams with larger size and higher field and task variety. However, to begin, this process would suggest that the team members (in forming the team) are invoking the causal process we are modeling. We are assuming that the team members are the carriers of the information used in the project, in particular, that they have a familiarity with a particular published literature that they bring to solving the problems of the team (Tang and Walsh, 2010). Thus, the variety of their backgrounds will affect the likelihood of novel combinations appearing in the final paper. In addition, since our novelty measure is based on the actual knowledge sources used in the final result (which we argue grows from the knowledge bases of the team members), we can argue that, although certain projects might invite certain team structures (size, field variety, or task variety), it is still the case that the assembled team is producing the given result, and, furthermore, it would be difficult, a priori, to predict which knowledge combinations would be used in the final result (and how rare those are compared to the universe of knowledge combination in contemporary publications). To further explore this concern, we test our models in Table 2, limiting our sample to the research projects that resulted in a focal paper that was not the same as originally planned, to minimize the self-assembly effect, and produce consistent results although the squared terms, still having the same direction, become insignificant.[7] Thus, although we cannot be certain about the directionality of the relations between novelty and team size, field or task variety, we argue that the causal direction we are postulating seems most reasonable, and is consistent with the theories (and experimental results) developed from prior work on team size, diversity and performance (Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998).

There can be a related concern that the effect of knowledge variety on novelty is tautological, given that both are measuring combining people or ideas from different sources. However, it is not logically necessary that those trained from different disciplines or involved in different tasks would cite papers from unusual combinations of journals in a given publication (although it is likely, as we argue based on our theory). Moreover, the modest correlations ($r = .25$ between fine field variety and novelty, $r = .26$ between coarse field variety and novelty, and $r = .18$ between task variety and novelty, see Table 1) do not suggest that these are measuring the same concept. Furthermore, we find, like size, that the second-order effects (squares of coarse field and task varieties) are negative (though not quite significantly in field variety), which suggests that while knowledge variety may drive greater novelty in the papers, there is a threshold, and that knowledge variety and novelty are not the same concept.

Finally, in terms of the effects of team characteristics on research impact, the directionality seems more straightforward. The process

[6] All results in this section are available from the contact author on request.

[7] The survey asked "Did the research project that yielded the focal paper proceed as initially planned?", with answers on a five point scale from '1: Largely the same as originally planned' to '5: Quite different than originally planned'. We excluded teams that answered "1" on this question and repeated the regressions in Table 2. All results are qualitative similar (although the negative square term in column 2 is no longer statistically significant).

of impact accumulation is temporally after the process of project development and team assembly, and a research output starts causing impacts only after the project is finished (and the project size is determined).

6. Conclusions

There is increasing interest in team science as a research domain and as a topic for policy discussion (Falk-Krzesinski et al., 2010; Fiore, 2008; Stokols et al., 2008; Vogel et al., 2013; Wuchty et al., 2007). Using new data on a large sample of research projects, we examine the drivers of creativity in team science. We begin by unpacking the concept of creativity into the distinct (though related) concepts of novelty and impact. We then show the effects of team size, field variety and task variety on both novelty and impact.

Our results show that team size has an inverted-U relation with novelty. Furthermore, this size effect on novelty is driven largely by the relation between size and knowledge variety, such that once we control for field or task variety, the size effect becomes insignificant. Thus, to the extent that increasing team size adds people from distinct knowledge domains, the chance of more novel outputs increases (up to a point, after which the negative effects of large teams dominate). Thus, the key novelty benefit from size is that it tends to increase knowledge variety. This result suggests that teams benefit from aggregating members with distinct knowledge bases. However, creating novel outcomes requires not only the divergent process of idea generation but also the convergent process of idea evaluation. While diversity may contribute to generating novel ideas, integration is important for identifying, evaluating and selecting the best novel ideas (Harvey and Kou, 2013; Singh and Fleming, 2010; Skilton and Dooley, 2010). Therefore, integration issues may make it difficult to manage such large and diverse groups in the process of producing novel outcomes (Jordan, 2006).

Team size also has a strong independent effect on impact, even after controlling knowledge variety and novelty of output. This is consistent with the interpretation that impact is an ex post judgment of the output and may depend on accretion of members' networks or resources for encouraging wider use of the output by others as well as by themselves. This suggests the need for further analyses of the social processes that affect the scientific impact of a paper net of its novelty and highlights the multidimensional nature of creativity, consisting of both novelty and impact, with each having distinct relations to team characteristics.

Our findings have several policy implications. First, our finding that half of teams in our sample involve a single field suggests significant opportunity for encouraging more interdisciplinarity in scientific teams. Thus, initiatives such as NSF's interdisciplinary research initiatives and programs such as Collaborative Interdisciplinary Team Science at NIH may be important for encouraging and facilitating greater knowledge variety in scientific teams. However, we also show that there are declining effects as knowledge variety increases. Thus, while encouraging knowledge variety is an important policy goal, too much variety can produce little marginal benefit and can even reduce novelty. Learning more about this bliss point, and how management of teams can help shift this curve upward is a key research agenda for S&T policy and management of innovation (Nooteboom, 2008).

Although our data represent scientific research teams in universities, many of these processes may also apply to research teams in firms (Pelz and Andrews, 1976; Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). These results suggest significant benefits from information-processing related diversity for science and for tasks such as research, development, design or other work areas that depend on drawing on diverse knowledge to generate original ideas and solve complex problems (Harvey, 2014; Page, 2007; Taylor and Greve, 2006; Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). This may not generalize to other dimensions of diversity such as race/ethnicity, gender, nationality or age/tenure (Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). Further work is needed to investigate how novelty and impact are affected by these other dimensions of diversity (which are often the target of S&T policies designed to increase the diversity of the STEM workforce). Furthermore, prior work in organizational behavior suggests that the relation between knowledge variety and performance may be mediated and moderated by a variety of process and contingency factors, including social categorization, communication and conflict, and the task relevance of the diversity dimensions (Van Knippenberg et al., 2004; Williams and O'Reilly III, 1998). Future work should elaborate these mediating and moderating effects for a fuller understanding of the dynamics of scientific collaborations. Van Knippenberg et al. (2004) argued that almost any dimension of diversity can be related to information elaboration, and to process loss from inter-group biases, suggesting the need to check for the contingencies under which elaboration is greater or lesser. Our work suggests that field variety (interdisciplinarity) and task variety (division of labor) are two key dimensions of task-related diversity. In addition, other aspects of diversity, such as gender, race/ethnicity and tenure may also affect creativity of the group, independent of or interacting with field or task variety. Further work should examine these demographic factors in team science. These arguments from organizational behavior research suggest that policies to develop tools for managing team science may help limit process loss associated with diverse scientific teams (Vogel et al., 2013). In addition, the management challenges of team science are contingent on the strategy and structure of the project, in particular how radical and how broad are the research goals of the project, suggesting the need to develop coordination and control mechanisms that are matched to the projects' strategic goals (Jordan, 2006). As science increasingly deals with boundary spanning problems that require drawing on the expertise of a large and diverse team, there is increasing need to understand the management challenges in such a team, in order to successfully produce novel solutions to pressing scientific and technical problems.

Our results also show the utility of a new bibliometric measure of novelty, which allows an objective and a priori measure of this aspect of creativity, independent of an ex post measure of impact (see also Uzzi et al., 2013). We encourage use of such metrics as a new tool for science policy indicators. Furthermore, our results suggest that the widely used metric of citations (impact) is only a rough proxy for the novelty of a finding, and that social factors such as team size affect impact independently of novelty. This suggests two cautions when using citations as a proxy for creativity: (1) citations may not map closely to the novelty of a finding, although it may be novelty that we are most interested in encouraging, especially in basic science, and (2) citations may be driven by social factors that are distinct from the quality of the research, suggesting the need for a more nuanced interpretation of this widely used metric. Thus, one key policy implication of these findings is to put novelty as a goal on the agenda of S&T policy. Our results suggest that a focus on impact (citations) may distract from the more fundamental goal of generating novel results. Evaluation systems should consider the novelty of the research, not just publication or citation counts. Fortunately, our work (along with that of Uzzi and colleagues) offers a new metric that can be incorporated into university evaluation systems interested in understanding the novelty of scientific outputs.

A related concern is that the emphasis on citation impact (as well as productivity) may lead to funding strategies that encourage overly large scientific teams, with potentially adverse effects on novelty. For example Hicks and Katz (2011) argued that science funders should concentrate funding in the most productive scientists, leading to ever larger labs (and greater inequality in funding). However, while the resulting large labs may be especially productive, and even, based on our findings, produce very high citation rates, they may not be especially novel in their research outputs. Our results thus suggest caution in encouraging larger teams (especially if the larger teams do not increase variety or do not integrate well). A related concern is that too much detailed division of labor (task variety) in large teams may result in overly specialized researchers that may not be able to contribute to the elaboration of knowledge needed to generate novel results (Hackett, 1990; Walsh and Lee, 2013). The negative second-order effect of task variety highlights this concern.

A limitation of our study is that we are only observing projects that resulted in at least one published WoS paper. Thus, papers with very low expected impact and/or very low novelty may be excluded from our population. Put differently, we are defining our population as those outputs that made it into the published literature, and then comparing the relative novelty and impact among that population. One issue of using WoS papers for this study is that the data are truncated favoring successful ideas while novel ideas that failed to be published are not observed. This issue may cause tension between our theoritcal arguments and empirical operationalization. Novelty and usefulness are described as distinct properties. For novelty, we are interested in combinatorial novelty, independent of usefulness (Fleming, 2001; Uzzi et al., 2013), and this property of novelty is an inherent quality of the paper, which can be measured in an ex ante and objective fashion. On the other hand, impact is a result of both novelty and usefulness (Whitley, 1984), and may even lean more towards the dimension of usefulness (Fleming, 2001) or popularity (Bentley, 2007). Furthermore, impact is realized through a social process interacting with the community and is therefore utimately an ex post and subjective judgment (Amabile, 1983; Singh and Fleming, 2010; Whitley, 1984). In short, we aim to unpack two components of creativity: novelty is combinatorial originality, and impact is the result of both novelty and usefulness. However, the failures are truncated in our data, that is, the ideas seen as useless are excluded. In other words, all our observations have passed a threshold of usefulness, and novel but useless ideas are not observed. It is worth considering how these results might be affected if we could observe idea generation before the filtering of the usefulness criteria that lead to publication. Future work on group processes and idea generation may help develop our understanding of the relations between team structure (size, field variety and task variety) and the novelty of ideas generated independently of any filter for viability or usefulness. For example, Taylor and Greve (2006) argued that knowledge diverse teams should produce wider varience in outputs.

The results in this study suggest that scientific creativity may be related to the composition and organization of the work teams in which the science is conducted (Andrews, 1976). Furthermore, creativity has two distinct components: novelty and impact, and these components have a nuanced set of relations to each other and to team size and knowledge variety. Thus, even given these relations, it is not obvious that increasing team size or knowledge variety will necessarily lead to greater novelty from scientific teams. Given the offsetting effects of size or knowledge variety (especially task variety), additional members may cause a net decline in novelty. And, hence, it may be the case that the observed sizes are optimized given the problems being addressed. It is important to keep these relationships in mind when developing policies that promote team science. Understanding team project composition and management may be key to implementing team science policies in a way that will increase the novelty and impact of scientific projects.

Appendix A. Field lists

Tables A.1 and A.2.

Table A.1
Journal fields for WoS.

1.	Agricultural science
2.	Biology & Biochemistry
3.	Chemistry
4.	Clinical medicine
5.	Computer science
6.	Environment/Ecology
7.	Economics & Business
8.	Engineering
9.	Geosciences
10.	Immunology
11.	Materials science
12.	Mathematics
13.	Microbiology
14.	Molecular biology & Genetics
15.	Multidisciplinary
16.	Neuroscience & Behavior
17.	Pharmacology
18.	Physics
19.	Plant & Animal science
20.	Psychiatry/Psychology
21.	Social sciences, general
22.	Space sciences

For post-stratification for sampling weights, we merge Economics & Business into Social sciences, and assign Multidisciplinary into one of disciplinary fields. Therefore, we use 20 fields for field controls with Chemistry as a reference group.

Table A.2
Researcher fields of expertise.

	Disaggregate 29 fields (i.e., university department level)	Aggregate 8 fields (i.e., NSF directorate level)
1	Agricultural science	Biological sciences
2	Biology & Biochemistry	
3	Botany, Zoology	
4	Microbiology	
5	Molecular biology & Genetics	
6	Environmental studies/Ecological science	
7	Earth science	Geosciences
8	Computer science	Computer science
9	Information Engineering	
10	Engineering, Electrical & Electronic	Engineering
11	Engineering, Environmental	
12	Engineering, Material	
13	Engineering, Mechanical	
14	Engineering, Medical	
15	Engineering, Urban	
16	Engineering, Chemical	
17	Clinical medicine	Medical sciences
18	Immunology	
19	Pharmaceutical science, Toxicology	
20	Space science	Mathematics and Physical sciences
21	Chemistry	
22	Materials science	
23	Mathematics/statistics	
24	Physics	
25	Social science	Social and behavioral sciences
26	Economics & Management	
27	Neuroscience & Behavioral science	
28	Psychiatric medicine/Psychology	
29	Other	Other

References

Adams, J., Black, G., Clemmons, J.R., Stephan, P.E., 2005. Scientific teams and institutional collaborations: evidence from U.S. Universities, 1981–1999. Research Policy 34, 259–285.

Aldrich, H.E., Pfeffer, J., 1976. Environments of organizations. Annual Review of Sociology 2, 79–105.

Amabile, T.M., 1982. Social psychology of creativity: a consensual assessment technique. Journal of Personality and Social Psychology 43, 997–1013.

Amabile, T.M., 1983. The social psychnology of creativity: a componential conceptualization. Journal of Personality and Social Psychology 45, 357–376.

Amason, A.C., Sapienza, H.J., 1997. The effects of top management team size and interaction norms on cognitive and affective conflict. Journal of Management 23, 495–516.

Ancona, D.G., Caldwell, D.F., 1992. Bridging the boundary: external activity and performance in organizational teams. Administrative Science Quarterly 37, 634–665.

Andrews, F.M., 1976. Creative process. In: Pelz, D.C., Andrews, F.M. (Eds.), Scientists in Organizations. , 2nd ed. University of Michigan Institute for Social Research, Ann Arbor, MI, pp. 337–366.

Antonelli, C., Franzoni, C., Geuna, A., 2011. The organisation, economics and policy of scientific research. Industrial and Corporate Change 20, 201–213.

Association of University Technology Managers, 2007. AUTM U.S. Licensing Activity Survey: FY 2007. AUTM, Deerfield, IL.

Aube, C., Rousseau, V., Tremblay, S., 2011. Team size and quality of group experience: the more the merrier? Group Dynamics: Theory Research and Practice 15, 357–375.

Baer, M., 2012. Putting creativity to work: the implementation of creative ideas in organizations. Academy of Management Journal 55, 1102–1119.

Barley, S.R., Bechky, B., 1994. In the backrooms of science: notes on the work of science technicians. Work and Occupations 21, 85–126.

Baron, R.M., Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51, 1173.

Bentley, R.A., 2007. Why do team-authored papers get cited more? Science (New York, NY) 317, 1496–1498, author reply 1496–1498.

Blau, P.M., 1970. A formal theory of differentiation in organizations. American Sociological Review 35, 201–218.

Blau, P.M., Heydebrand, W.V., Stauffer, R.E., 1966. The structure of small bureaucracies. American Sociological Review 31, 179–191.

Brown, C., Ross, M.L., 2013. Fermilab: Science at Work., pp. 42.18.

Carayol, N., Matt, M., 2004. Does research organization influence academic production? Laboratory level evidence from a large European university. Research Policy 33, 1081–1102.

Child, J., 1973. Predicting and understanding organization structure. Administrative Science Quarterly 18, 168–185.

Clark, K.B., Chew, W.B., Fujimoto, T., 1987. Product development in the world auto industry. Brookings Papers on Economic Activity 3, 729–781.

Cohen, S.G., Bailey, D.E., 1997. What makes teams work: group effectiveness research from the shop floor to the executive suite. Journal of Management 23, 239–290.

Cohen, W.C., Levinthal, D.A., 1990. Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35, 128–152.

Cole, S., Cole, J.R., 1967. Scientific output and recognition. American Sociological Review 32, 377–390.

Contractor, N., 2013. Some assembly required: leveraging web science to understand and enable team assembly. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 371, 1–14.

Cullen, J.B., Anderson, K.S., Baker, D.D., 1986. Blau's theory of structural differentiation revisited: a theory of structural change and scale. Academy of Management Journal 29, 203–299.

Cummings, J.N., Kiesler, S., 2005. Collaborative research across disciplinary and organizational boundaries. Social Studies of Science 35, 703–722.

Cummings, J.N., Kiesler, S., Zadeh, R., Balakrishnan, A., 2013. Group heterogeneity increases the risks of large group size: a longitudinal study of productivity in research groups. Psychological Science 24, 880–890.

Dahlin, K.B., Behrens, D.M., 2005. When is an invention really radical?: Defining and measuring technological radicalness. Research Policy 34, 717–737.

De Bellis, N., 2009. Bibliometrics and Citation Analysis: from the Science Citation Index to Cybermetrics. Scarecrow Press, Lanham, MD.

de Solla Price, D.J., 1986. Little Science, Big Science – and Beyond. Columbia University Press, New York.

Etzkowitz, H., 1983. Entrepreneurial scientists and entrepreneurial universities in American academic science. Minerva 21, 198–233.

Evers, F.T., Bohlen, J.M., Warren, R.D., 1976. The relationships of selected size and structure indicators in economic organizations. Administrative Science Quarterly 21, 326–342.

Falk-Krzesinski, H.J., Börner, K., Contractor, N., Fiore, S.M., Hall, K.L., Keyton, J., Spring, B., Stokols, D., Trochim, W., Uzzi, B., 2010. Advancing the science of team science. Clinical and Translational Science 3, 263–266.

Fiore, S.M., 2008. Interdisciplinarity as teamwork: how the science of teams can inform team science. Small Group Research 39, 251–277.

Fleishman, J.A., 1980. Collective action as helping-behavior: effects of responsibility diffusion on contributions to a public good. Journal of Personality and Social Psychology 38, 629–637.

Fleming, L., 2001. Recombinant uncertainty in technological search. Management Science 47, 117–132.

Fleming, L., Mingo, S., Chen, D., 2007. Collaborative brokerage, generative creativity, and creative success. Administrative Science Quarterly 52, 443–475.

Ford, C.M., 1996. Theory of individual creative action in multiple social domains. Academy of Management Review 21, 1112–1142.

Franzoni, C., 2010. Exploratory Attitude and Creativity in Science, Science for Innovation: Exploiting and Strengthening the Linkage. Research Institute of Economy, Trade and Industry, Tokyo, Japan.

Garfield, E., 1973. Citation and distinction. Nature 242, 485.

Gooding, R.Z., Wagner III, J.A., 1985. A meta-analytic review of the relationship between size and performance: the productivity and efficiency of organizations and their subunits. Administrative Science Quarterly 30, 462–481.

Guimera, R., Uzzi, B., Spiro, J., Amaral, L.A.N., 2005. Team assembly mechanisms determine collaboration network structure and team performance. Science 308, 697–702.

Hackett, E.J., 1990. Science as a vocation in the 1990 – the changing organizational culture of academic science. Journal of Higher Education 61, 241–279.

Hare, A.P., 1952. A study of interaction and consensus in different sized groups. American Sociological Review 17, 261–267.

Hargadon, A.B., Bechky, B.A., 2006. When collections of creatives become creative collectives: a field study of problem solving at work. Organization Science 17, 484–500.

Harrison, D.A., Klein, K.J., 2007. What's the difference? Diversity constructs as separation, variety, or disparity in organizations. Academy of Management Review 32, 1199–1228.

Harvey, S., 2014. Creative synthesis: exploring the process of extraordinary group creativity. Academy of Management Review 39, 324–343.

Harvey, S., Kou, C.-Y., 2013. Collective engagement in creative tasks: the role of evaluation in the creative process in groups. Administrative Science Quarterly 58, 346–386.

Heinze, T., Shapira, P., Rogers, J.D., Senker, J.M., 2009. Organizational and institutional influences on creativity in scientific research. Research Policy 38, 610–623.

Hemlin, S., Allwood, C.M., Martin, B.R., Mumford, M.D., 2013. Creativity and Leadership in Science, Technology, and Innovation. Routledge, New York, NY.

Hicks, D., Katz, J.S., 2011. Equity and excellence in research funding. Minerva 49, 137–151.

Hicks, D.M., Katz, J.S., 1996. Where is science going? Science, Technology, & Human Values 21, 379–406.

Hollingsworth, J.R., 2006. A path-dependent perspective on institutional and organizational factors shaping major scientific discoveries. In: Hage, J., Meeus, M. (Eds.), Innovation, Science, and Institutional Change. Oxford University Press, Oxford, pp. 423–442.

Hollingsworth, R., 2004. Institutionalizing excellence in biomedical research: the case of Rockefeller University. In: Stapleton, D.H. (Ed.), Creating a Tradition of Biomedical Research. Rockefeller University Press, New York.

Hong, L., Page, S.E., 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proceedings of the National Academy of Sciences 101, 16385–16389.

Hong, W., 2008. Domination of a scientific field: capital struggle in a Chinese isotope lab. Social Studies of Science 38, 543–570.

Hudson, J., 1996. Trends in multi-authored papers in economics. The Journal of Economic Perspectives 10, 153–158.

Hulsheger, 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, 1128–1145.

Jehn, K.A., Northcraft, G.B., Neale, M.A., 1999. Why differences make a difference: a field study of diversity, conflict, and performance in workgroups. Administrative Science Quarterly 44, 741–763.

Jones, B.F., 2009. The burden of knowledge and the death of the renaissance man: is innovation getting harder? Review of Economic Studies 76, 283–317.

Jordan, G.B., 2006. Factors influencing advances in basica and applied research: variation due to diversity in research profiles. In: Hage, J., Meeus, M. (Eds.), Innovation, Science, and Institutional Change. Oxford University Press, Oxford.

Kalton, G., 1983. Introduction to Survey Sampling. Sage, Newbury Park, CA.

Katz, J.S., 1999. The self-similar science system. Research Policy 28, 501–517.

Katz, J.S., 2000. Scale-independent indicators and research evaluation. Science and Public Policy 27, 23–36.

Katz, J.S., Martin, B.R., 1997. What is research collaboration? Research Policy 26, 1–18.

Kiesler, C.A., 1969. Group Pressure and Conformity. Macmillan, New York.

Kimberly, J.R., 1976. Organizational size and structuralist perspective: a review, critique, and proposal. Administrative Science Quarterly 21, 571–597.

Larivière, V., Gingras, Y., 2010. On the relationship between interdisciplinarity and scientific impact. Journal of the American Society for Information Science and Technology 61, 126–131.

Latour, B., Woolgar, S., 1979. Laboratory Life. Sage, Beverly Hills.

Lavie, D., Drori, I., 2012. Collaborating for knowledge creation and application: the case of nanotechnology research programs. Organization Science 23, 704–724.

Lee, E.S., Forthofer, R.N., Lorimor, R.J., 1989. Analyzing Complex Survey Data. Sage, Newbury Park, CA.

Manners, G.E., 1975. Another look at group size, group problem-solving, and member consensus. Academy of Management Journal 18, 715–724.

Martin, B.R., Irvine, J., 1983. Assessing basic research: some partial indicators of scientific progress in radio astronomy. Research Policy 12, 61–90.

Melin, G., 2000. Pragmatism and self-organization: research collaboration on the individual level. Research Policy 29, 31–40.

Merton, R.K., 1940. Bureaucratic structure and personality. Social Forces 18, 560–568.
Merton, R.K., 1968. The Matthew effect in science. Science 159, 56–63.
Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. University of Chicago Press, Chicago, IL.
Meyer, M.W., 1972. Size and the structure of organizations. American Sociological Review 37, 434–440.
Milliken, F.J., Martins, L.L., 1996. Searching for common threads: understanding the multiple effects of diversity in organizational groups. Academy of Management Review 21, 402–433.
Moed, H.F., 2005. Citation Analysis in Research Evaluation. Springer, Dordrecht, The Netherlands.
Mueller, J.S., Melwani, S., Goncalo, J.A., 2011. The bias against creativity: why people desire but reject creative ideas. Psychological Science.
National Academy of Sciences, 2004. Facilitating Interdisciplinary Research. National Academies Press, Washington, DC.
National Research Council, 1995. Research Doctorate Programs in the United States: Continuity and Change. The National Academies Press, Washington, DC.
Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Belknap Press of Harvard University Press, Cambridge, MA.
Nemet, G.F., Johnson, E., 2012. Do important inventions benefit from knowledge originating in other technological domains? Research Policy 41, 190–200.
Newman, J.M., Cooper, E., 1993. Determinants of academic recognition: the case of the journal of applied psychology. Journal of Applied Psychology 78, 518–526.
Nooteboom, B., 2008. Learning and innovation in inter-organizational relationships. In: Ebers, M., Smith Ring, P. (Eds.), Handbook of Inter-Organizational Relationships. Oxford University Press, Oxford, pp. 607–634.
NRC, 2014, 2014. Capturing Change in Science, Technology, and Innovation: Improving Indicators to Inform Policy. The National Academies Press, Washington, DC.
Obstfeld, D., 2005. Social networks, the tertius iungens orientation, and involvement in innovation. Administrative Science Quarterly 50, 100–130.
Page, S.E., 2007. The Difference. Princeton University Press, Princeton, NJ.
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, 1–28.
Pelz, D.C., Andrews, F.M., 1976. Scientists in Organizations, 2nd ed. University of Michigan Institute for Social Research, Ann Arbor, MI.
Pfeffer, J., Salancik, G.R., 1978. The External Control of Organizations: a Resource Dependence Perspective. Harper and Row, New York, NY.
Pieterse, A.N., Knippenberg, D.v., Dierendonck, D.v., 2013. Cultural diversity and team performance: the role of team member goal orientation. Academy of Management Journal 56, 782–804.
Rafols, I., Meyer, M., 2010. Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. Scientometrics 82, 263–287.
Rogers, J.D., 2010. Citation analysis of nanotechnology at the field level: implications of R&D evaluation. Research Evaluation 19, 281–290.
Schoenmakers, W., Duysters, G., 2010. The technological origins of radical inventions. Research Policy 39, 1051–1059.
Schumpeter, J.A., 1939. Business Cycles: a Theoretical, Historical, and Statistical Analysis of the Capitalist Process, 1st ed. McGraw-Hill Book Company Inc., New York, London.
Shibayama, S., Walsh, J.P., Baba, Y., 2013. Social Organisation of Life Science Laboratories. The Organisation, Economics and Policy of Scientific Research, Torino, Italy.
Shrum, W., Genuth, J., Chompalov, I., 2007. Structures of Scientific Collaboration. MIT Press, Cambridge, MA.
Simonton, D.K., 1999. Origins of Genius: Darwinian Perspectives on Creativity. Oxford University Press, New York.
Simonton, D.K., 2003. Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives. Psychological Bulletin 129, 475–494.
Simonton, D.K., 2004. Creativity in Science: Chance, Logic, Genius, and Zeitgeist. Cambridge University Press, Cambridge, UK; New York.
Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns. Management Science 51, 756–770.
Singh, J., Fleming, L., 2010. Lone inventors as sources of breakthroughs: myth or reality? Management Science 56, 41–56.
Skilton, P.F., Dooley, K.J., 2010. The effects of repeated collaboration on creative abrasion. Academy of Management Review 35, 118–134.
Smith, A., 1776. Wealth of Nations.
Steele, T.W., Stier, J.C., 2000. The impact of interdisciplinary research in the environmental sciences: a forestry case study. Journal of the American Society for Information Science 51, 476–484.
Stephan, P., 2012. How Economics Shapes Science. Harvard University Press.
Stewart, G.L., 2006. A meta-analytic review of relationships between team design features and team performance. Journal of Management 32, 29–55.
Stokols, D., Hall, K.L., Taylor, B.K., Moser, R.P., 2008. The science of team science: overview of the field and introduction to the supplement. American Journal of Preventive Medicine 35, S77–S89.
Swatez, G.M., Unpublished dissertation 1966. Social Organization of a University Laboratory. University of California, Berkeley, CA.
Tang, L., Walsh, J.P., 2010. Bibliometric fingerprints: name disambiguation based on approximate structure equivalence of cognitive maps. Scientometrics 84, 763–784.
Taylor, A., Greve, H.R., 2006. Superman or the fantastic four? Knowledge combination and experience in innovative teams. Academy of Management Journal 49, 723–740.
Traweek, S., 1988. Beamtimes and Lifetimes. Harvard University Press, Cambridge, MA.
Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical combinations and scientific impact. Science 342, 468–472.
Uzzi, B., Spiro, J., 2005. Collaboration and creativity: the small world problem. American Journal of Sociology 111, 447–504.
Valderas, J.M., 2007. Why do team-authored papers get cited more? Science (New York, NY) 317, 1496.
Van Knippenberg, D., De Dreu, C.K.W., Homan, A.C., 2004. Work group diversity and group performance: an integrative model and research agenda. Journal of Applied Psychology 89, 1008–1022.
van Raan, A.F.J., 2006a. Performance-related differences of bibliometric statistical properties of research groups: cumulative advantages and hierarchically layered networks. Journal of the American Society for Information Science and Technology 57, 1919–1935.
van Raan, A.F.J., 2006b. Statistical properties of bibliometric indicators: research group indicator distributions and correlations. Journal of the American Society for Information Science and Technology 57, 408–430.
Ven De Ven, A.H., Delbecq, A.L., Koenig Jr., R., 1976. Determinants of coordination modes within organizations. American Sociological Review 41, 322–338.
Vogel, A.L., Hall, K.L., Fiore, S.M., Klein, J.T., Bennett, L.M., Gadlin, H., Stokols, D., Nebeling, L.C., Wuchty, S., Patrick, K., Spotts, E.L., Pohl, C., Riley, W.T., Falk-Krzesinski, H.J., 2013. The Team Science Toolkit: enhancing research collaboration through online knowledge sharing. American Journal of Preventative Medicine 45, 787–789.
Walsh, J.P., Lee, Y.-N., 2013. The bureaucratization of science. In: American Sociological Association Annual Meeting, New York, NY.
Wang, J., 2014. Unpacking the Matthew effect in citations. Journal of Informetrics 8, 329–339.
West, B.T., Berglund, P., Heering, S.G., 2008. A closer examination of subpopulation analysis of complex-sample survey data. Stat Journal 8, 520–531.
Whitley, R., 1984. The Intellectual and Social Organization of the Sciences. Oxford University Press, Oxford.
Williams, K.Y., O'Reilly III, C.A., 1998. Demography and diversity in organizations. Research in Organizational Behavior 20, 77–140.
Woodman, R.W., Sawyer, J.E., Griffin, R.W., 1993. Toward a theory of organizational creativity. Academy of Management Review 18, 293–321.
Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA.
Wuchty, S., Jones, B., Uzzi, B., 2007. The increasing dominance of teams in the production of knowledge. Science 316, 1036–1039.
Yong, K., Sauer, S.J., Mannix, E.A., 2014. Conflict and creativity in interdisciplinary teams. Small Group Research 45, 266–289.
Zhu, M., Huang, Y., Contractor, N.S., 2013. Motivations for self-assembling into project teams. Social Networks 35, 251–264.
Zuckerman, H., 1967. Nobel laureates in science: patterns of productivity, collaboration, and authorship. American Sociological Review 32, 391–403.