Received: 2 March 2019 | Revised: 1 August 2019 | Accepted: 7 August 2019 | Published on: 15 October 2019

DOI: 10.1002/smj.3091

RESEARCH ARTICLE

STRATEGIC MANAGEMENT JOURNAL WILEY

Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration

Frank Nagle[1] | **Florenta Teodoridis**[2]

[1]Strategy Department, Harvard Business School, Boston, Massachusetts

[2]Management and Organization Department, University of Southern California, Marshall School of Business, Los Angeles, California

Correspondence
Florenta Teodoridis, PhD, Management and Organization Department, University of Southern California, Marshall School of Business, 701 Exposition Boulevard, HOH 520, Los Angeles, CA 90089.
Email: florenta.teodoridis@marshall.usc.edu

Abstract

Research Summary: We consider the role of individual-level diversification as a mechanism through which skilled researchers engage in successful exploration—recognizing and integrating new knowledge external to one's domains of expertise. To approach an ideal experiment, we (a) employ a matching procedure and (b) exploit the unexpected adoption of Microsoft Kinect as a motion-sensing technology in research. We evaluate the impact of Kinect and its embodiment of new knowledge on a set of ability-matched, diversity-varying researchers without prior experience in motion-sensing and find that diversified researchers explore more successfully than their more specialized peers. We also examine the role of personal preferences and professional incentives as antecedents of diversification and find that culture, age and intellectual freedom are positively associated with the propensity to diversify successfully.

Managerial Summary: Organizations where R&D is core to driving competitive advantage face important tradeoffs when hiring researchers. Specifically, diverse combinations of knowledge generate the most impactful discoveries. Yet, coordinating such combinations increasingly requires larger teams as knowledge accumulation causes researchers to specialize in narrower areas. How should organizations achieve the best balance? We argue and show evidence that diversified researchers, individuals routinely criticized for their lack of knowledge depth, are more likely than specialized researchers of similar ability to integrate new knowledge from beyond their domains of expertise to create impactful innovations. Therefore, organizations aiming to create competitive advantage by pushing the boundaries of knowledge should carefully consider the nuanced tradeoffs between specialized and diversified researchers when strategizing about hiring the optimal types of expertise.

KEYWORDS
diversification, exploration, innovation, recombination, specialization

1 | INTRODUCTION

The role of innovation in creating value and competitive advantage for organizations has long been of interest to the fields of strategy, management, and economics (e.g., Nelson & Winter, 1977; Schumpeter, 1934). In particular, the ability to "explore," namely, to recognize and integrate new external knowledge, that is, outside one's domains of expertise, has continually been shown to help organizations thrive (Chatterji & Fabrizio, 2014; Christensen, 1992; Cohen & Levinthal, 1989; Fleming, 2001; Tushman & Anderson, 1986). Despite this, the precise individual characteristics that allow innovators and, by extension, their organizations, to be among the first to successfully explore by integrating new external knowledge have gone underexamined. In this paper, we explore one such characteristic—namely an individuals' level of knowledge diversification—as a mechanism through which the successful exploration of skilled individuals manifests.

Scholars have revealed that ability is a precursor to successful exploration, both at the individual and the organization level (e.g., Ahuja, Lampert, & Tandon, 2008; Fleming, 2001; Gavetti & Levinthal, 2000; Greve, 2007; Henderson, 1993; King & Tucci, 2002). However, while at the organization level various mechanisms through which ability[1] leads to successful exploration have been analyzed (Cyert & March, 1963; Dothan & Lavie, 2016; Eggers & Kaul, 2018; Fleming & Sorenson, 2004; Greve, 2003), less is known about mechanisms through which the successful exploration of skilled individuals[2] manifests. Understanding this is important as knowledge-based organizations are increasingly relying on scientists, engineers, and researchers to drive value creation and competitive advantage (Agrawal, McHale, & Oettl, 2017; Barth, Davis, Freeman, & Wang, 2017) and because combining broadly across the knowledge frontier has been shown to lead to the most significant discoveries (e.g., Boudreau, Lacetera, & Lakhani, 2011; Chai, 2017; Katila & Ahuja, 2002; Lifshitz-Assaf, 2017; Schilling & Green, 2011; Uzzi, Mukherjee, Stringer, & Jones, 2013; Weitzman, 1998).

[1]There are also numerous studies focused on uncovering factors that contribute to building an ability to successfully explore at the firm level, such as complementary assets (Rothaermel, 2001; Taylor & Helfat, 2009; Tripsas, 1997) and competition pressures (Bayus & Agarwal, 2007; Eggers, 2014; Wu, Wan, & Levinthal, 2014).

[2]We use the phrase "skilled individuals" throughout the paper to denote individuals who have the skillset, and therefore level of ability necessary for exploration.

In this paper, we theorize and provide evidence of the benefits of diversified individuals in integrating new knowledge that is outside their domains of expertise, despite these individuals' perceived disadvantage in knowledge depth.

To evaluate the role of individual-level diversification in integrating new external knowledge, we focus on researchers' propensity to engage with new knowledge embodied in new technology developments outside their current domains of knowledge. This is consistent with Mokyr's (2002) arguments that technology facilitates access to new or costly knowledge.[3]

Identifying a relationship between individual levels of knowledge diversification and the propensity to engage with new technology developments is, however, difficult because, when observing successful exploration, it is unclear if an individual's level of diversification facilitates that successful engagement or if the individual strategically chose to focus on certain domains of knowledge that were promising or if both the level of diversification and the successful engagement are driven by the individual's degree of ability. Ideally, we would like to observe individuals exhibiting similar ability, but varying levels of diversity being exposed to a new and exogenous technology development that is outside their domains of knowledge, and then estimate if individuals with higher levels of diversity exhibit a higher propensity to successfully integrate the new technology.

We attempt to get close to this setting by following a two-step empirical strategy. First, we exploit the unexpected use of Microsoft Kinect in research as a new technology development in motion-sensing research. Kinect, an add-on for Xbox 360, was launched in November 2010 in the video game market but was unexpectedly embraced by the research community as a motion-sensing research technology in fields ranging from artificial intelligence, robotics, and virtual reality to paleontology, education, health care, music, cinematography, market research, and advertising.[4] We follow the interpretation of the events in Teodoridis (2018), which argues that the role of Kinect as a motion-sensing research technology was not anticipated by the research community. Second, we observe the impact of this technology on a sample of researchers where we hold ability constant but allow for varying levels of knowledge diversity by employing coarsened exact matching (CEM) (Iacus, King, & Porro, 2011, 2012).

To capture the role individual-level diversity plays in the propensity to engage with the new Kinect technology, we estimate the propensity of our ability-matched, diversity-varying sample of researchers to publish academic papers referencing the Kinect in the period after the launch. We find that individuals without prior direct experience with motion-sensing—individuals for whom Kinect represents a new technology development outside their domains of expertise—but who are in the top quartile of knowledge diversification, are 3.1 times more likely to engage with Kinect in their research than individuals of similar ability and without motion-sensing experience, but who are in the bottom quartile of knowledge diversification. The effect is even more pronounced when focusing on highly cited output, with diversified researchers being 3.8 times more likely to produce papers in the top 10th percentile of academic papers ranked by

[3]We focus on researchers so we can measure innovative output in a tangible way—academic papers. Although this paper trail of innovation primarily tracks the inventive output of individuals working in research-oriented organizations, existing work shows that within a given field, industrial and academic researchers behave similarly in the context of knowledge creation (Sauermann & Stephan, 2013). Furthermore, the use of publications, rather than patents, is appropriate for addressing the proposed question given our focus on analyzing the role of individual levels of diversification in knowledge creation which occurs at all stages of innovation, not only at the later, patentable stages.

[4]See, for example, kinecthacks.com and blogs.msdn.microsoft.com/kinectforwindows for a compilation of various applications of Kinect outside the gaming industry. A Factiva search on Kinect articles returns close to 20,000 hits for the period 2011–2014.

number of citations. Importantly, the propensity to write highly cited output is not accompanied by an increase in output at the left tail of the impact distribution (less-cited output). Furthermore, the highly cited publications authored by diverse individuals are generated by combining more diverse pieces of knowledge, as evidenced by the spread of these publications' keywords across IEEE domains. Reassuringly, we find that diversity does not play a significant role when the new knowledge is local, namely for individuals with prior expertise in motion-sensing. We interpret these results as providing strong support for our arguments that individual-level diversification plays an important role in the propensity to explore by integrating new knowledge that is outside the individual's domains of knowledge. Finally, we consider several mechanisms behind why a researcher chooses to become diversified in the first place and show evidence suggesting that diversification is both a result of personal preferences and professional incentives. While we are limited by data availability in making strong claims, we offer evidence and discuss the potential influencing role of several factors including culture, age, choice of education path, level of general-purposeness of chosen domains of study, and employer incentives for diversification or specialization.

Our findings contribute to the strategy literature by identifying the role of an important individual-level characteristic in successful exploration: these "jacks of all trades, masters of knowledge" are more effective at integrating new knowledge that is outside their domains of prior knowledge than their more specialized colleagues. Thus, we offer a deeper understanding of individual characteristics that may contribute to firm success, an approach that has been proposed as a central avenue for pushing the boundaries of strategic management research (Felin & Foss, 2005; Foss, 2011; Gavetti, 2005; Teece, 2007). We also help shine a light on the role of the individual in absorptive capacity (Cohen & Levinthal, 1990), which we argue to be particularly important given the increased tendency of firms to rely on knowledge workers to drive value creation and competitive advantage (Fabrizio, 2009; Perkmann et al., 2013).[5] We further contribute to the academic literature on the economics of science and innovation by offering insights into the role of diversified individuals in knowledge creation and into how academics and other researchers pursue their careers. Young researchers, especially in academia, are frequently encouraged to be highly specialized and to focus on a very narrow field (Stephan, 2012), even though distant novel combinations of knowledge are shown to generally lead to the most impactful results (e.g., Boudreau et al., 2011; Chai, 2017; Katila & Ahuja, 2002; Lifshitz-Assaf, 2017; Schilling & Green, 2011; Uzzi et al., 2013; Weitzman, 1998).

Overall, our results suggest that knowledge-based organizations would gain value from considering the potential benefits of diversified researchers. Specifically, organizations seeking to integrate distant knowledge into their innovation efforts should consider hiring such diversified individuals to help increase their ability to explore the knowledge frontier more broadly. Finally, our results suggest that decision-makers in all fields where research integrating distant knowledge is important should reduce the emphasis they place on individual specialization at the expense of diversification.

[5]For example, in our dataset, a scientist at Intel whose research before Kinect had covered diverse topics such as personal activity sensors (similar to the FitBit), privacy concerns related to public WiFi, self-awareness of physical exercise, and human–robot interaction quickly engaged with this new and distant knowledge (relative to his research domains up to that point) to help create a system allowing a robot to play physical board games, an important research contribution in artificial intelligence.

2 | THEORY AND HYPOTHESES DEVELOPMENT

Our goal is to evaluate the role of individual-level diversification in successful exploration. We define exploration as the process of integrating new external[6] knowledge—that is outside of the individual's domains of knowledge—into successful knowledge creation. Integrating new distant knowledge is not necessarily superior to integrating new local knowledge (Kaplan & Vakili, 2015) and both approaches have been found to benefit organizations. We choose to focus on the process of integrating distant pieces of knowledge, which has been shown to be important for leading to novel and impactful innovations in studies from economics (e.g., Nelson & Winter, 1973; Weitzman, 1998), strategic management (Katila & Ahuja, 2002; Schilling & Green, 2011; Uzzi et al., 2013), and open and user innovation (Afuah & Tucci, 2012; Altman, Nagle, & Tushman, 2014; Boudreau et al., 2011; Boudreau, Guinan, Lakhani, & Riedl, 2016; Jeppesen & Lakhani, 2010; Laursen & Salter, 2006; Lifshitz-Assaf, 2017).

Novel and impactful innovations are important for organizations because they can be a source of competitive advantage (Cohen & Levinthal, 1990; Murmann, 2003). Since knowledge production is a recombinant process (Fleming, 2001; Schumpeter, 1934), this implies that there is a competitive process in which organizations search for the best combinations of distant knowledge pieces with other (potentially local) knowledge pieces and try to be the first to use them. Scholars studying how organizations can successfully engage in such a recombination process have shown that ability is a precursor to successful exploration, both at the individual and the organization level (e.g., Ahuja et al., 2008; Fleming, 2001; Gavetti & Levinthal, 2000; Greve, 2007; Henderson, 1993; King & Tucci, 2002). However, while there is fairly extensive evidence on several mechanisms through which skilled organizations can achieve successful exploration (Dothan & Lavie, 2016; Eggers & Kaul, 2018; Fleming & Sorenson, 2004; Greve, 2003), less is known about mechanisms of successful exploration of skilled individuals.

We contribute to this literature by drawing attention to individual-level knowledge diversification. At the organization-level, knowledge diversification has been found to lead to an increased propensity of organizations to integrate new distant knowledge (Cohen & Levinthal, 1990; Katila & Ahuja, 2002). We argue that achieving organization-level knowledge diversification is an increasingly complex endeavor that places the role of the individual front and center in exploration attempts due to the continuous increase in knowledge stock which creates both opportunities and challenges for recombining distant pieces of knowledge. In particular, as the burden of knowledge accumulation increases, researchers and scientists are forced to specialize in narrower domains of knowledge (Jones, 2009). This implies that individuals become increasingly likely to focus their knowledge output (producing research) and their knowledge input (consuming research) in the field they are specialized in and are therefore unlikely to be aware of and identify distant knowledge (Toh, 2014). This is important because it suggests that to achieve organization-level diversification, organizations

[6]We use the terms "external knowledge" and "distant knowledge" interchangeably. Search distance is frequently thought of in the realm of product development and new firm-level innovations (Gupta, Smith, & Shalley, 2006; Katila & Ahuja, 2002; Shane, 2000). As Adner and Levinthal (2008) point out, "The distance of search is usually measured as the extent of departure from established routines and behavioral patterns." However, in the realm of knowledge creation and research, we can think of "established routines and behavioral patterns" as the areas in which an individual has performed research before. So, when a researcher with prior experience performing research exclusively in the field of microeconomics publishes a paper that builds on some new piece of knowledge in the field of microeconomics, they were exploiting their existing experience and utilizing local knowledge. However, when the same researcher publishes a paper that builds on some new piece of knowledge in the field of biology, and they have never used knowledge from the field of biology, they are exploring new domains of knowledge by performing a distant search and using distant knowledge.

should ensure access to a wider pool of specialists who can collaborate to combine their narrow-specialized knowledge (Agrawal, Goldfarb, & Teodoridis, 2016; Jones, 2009, 2010). However, this is a costly process. First, it is unclear which specializations should be kept on hand to achieve a combined level of diversification conducive to successful exploration. Second, large teams of specialist collaborators suffer from exponentially increasing coordination costs (Bikard, Murray, & Gans, 2015; Teodoridis, 2018).

An alternative to individual-level specialization is an individual who chooses to respond to the burden of knowledge accumulation by embracing a wider breadth of knowledge albeit at the expense of some knowledge depth (Jones, 2010). Unlike their specialized colleagues, diversified individuals benefit from a higher variety in knowledge breadth and hence are more likely to become aware of new knowledge beyond domains they have produced research in previously. Furthermore, these individuals would be more likely to recognize fruitful combinations of knowledge pieces that include the new external knowledge. The literature on exploration argues that combining distant pieces of knowledge will also lead to more breakthroughs (Fleming, 2001; Schilling & Green, 2011; Uzzi et al., 2013) and that the effect will be more pronounced for combinations that include many rather than few pieces of distant knowledge (Fleming, 2001). However, to achieve breakthroughs, it is important to understand which knowledge pieces to recombine. This can be difficult for specialized individuals (Chai, 2017; Toh, 2014) who rely on knowledge pieces that cover a rather narrow knowledge distance and who do not have experience working with a wider breadth of knowledge domains. Diversified individuals benefit from experience in working across different knowledge domains and hence have a higher propensity to understand what knowledge is necessary for potentially impactful recombinations as well as what knowledge combinations would lead to less impactful recombinations. Furthermore, the diversified individuals gain their experience across different knowledge domains by working with a wider network of collaborators than that of their more specialized colleagues (Tortoriello, McEvily, & Krackhardt, 2015). These networks enable diversified individuals to overcome the limitations of their depth of knowledge and successfully execute on the observed opportunities (Teodoridis, 2018).

More formally, our primary arguments can be summarized in two hypotheses:

Hypothesis (H1). *Individuals who have a higher degree of knowledge diversification have a higher propensity to integrate new knowledge from outside their domains of expertise than those who have a lower degree of knowledge diversification.*

Hypothesis (H2). *When individuals who have a higher degree of knowledge diversification integrate new knowledge from outside their domains of expertise, they do so by combining more categories of knowledge than their less diversified peers, which results in more high-impact output than low-impact output.*

A natural question that arises from these arguments is that of why a skilled individual might become diversified in the first place. In particular, when holding the level of ability constant, why does one individual specialize and another diversify? To better understand the mechanism through which this process occurs, we consider two primary avenues: (a) a personal preference for diversification and (b) professional incentives that encourage one to diversify or to specialize.

First, an individual's decision to diversify or to specialize may be driven by intrinsic preferences. Sociologists and psychologists have long discussed the idea that people are characterized by a variety of intrinsic skills and talents that inform their preferences to pursue a certain career and lifestyle (e.g., Gagne & Briggs, 1974). These tend to manifest in the choices individuals make. For example, it is likely that diversified researchers exhibited a taste for diversity in other walks of life, or earlier in their professional life while, for example, training for their future profession—engaging in attending diverse courses, focusing on more diversified fields of study, or pursuing a variety of different specialties or majors in their degrees. Furthermore, these preferences are likely shaped throughout one's life by the upbringing they receive. For example, social psychology has long shown that different cultures embrace and instill different values on their members (e.g., Hofstede, 1980; Hofstede, Neuijen, Ohayv, & Sanders, 1990). Hofstede (1991) identifies six dimensions of national cultures—power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence—and demonstrates how these attributes influence individuals' behavior in social interactions and in the workplace. Akin to these insights, we argue that these attributes might influence researchers' preferences for diversification. More specifically, the attributes might capture cultural dimensions that influence the level of freedom a researcher perceives they have in following up on their intrinsic preferences and the researcher's assertiveness in pursuing those preferences. Therefore, we expect culture to influence the manifestation of personal preferences for diversification.[7]

Second, the propensity to diversify is likely to be influenced by incentives stemming from a researcher's employer. Researchers employed by an academic institution are generally given more freedom to pursue research topics they find appealing, whereas industry-employed researchers tend to be more focused on supporting their employer's business and have less freedom to study disparate topics (Bush, 1945). This suggests that researchers who chose to work in academia are likely to be more diversified, on average, than their colleagues employed in the private sector. Furthermore, within academia, the incentives for diversification vary over the course of one's career. In particular, early-career researchers have been increasingly encouraged to hyper-focus to try to become the world's expert in a certain domain (Stephan, 2012). This is particularly true early in the career of academic researchers who are on the tenure track or who simply are attempting to establish their reputation in the academic world (Harris, 2018). The incentives to build a reputation necessary for securing employment are likely to also influence private sector researchers' decisions to specialize or diversify, albeit to a milder degree. In aggregate, incentive alignment is thus likely to play a role in the diversification of an individual such that, on the whole, academic researchers are likely to be more diversified than their private sector colleagues. However, the effect is likely driven by academics who are further along in their career.

Overall, our argument is that individuals who chose to respond to the burden of knowledge by embracing a wider breadth of knowledge at the expense of some depth offer a unique opportunity for organizations to facilitate successful exploration—via identification of new distant knowledge and opportunities for recombinations that are more likely to lead to breakthroughs. The benefits these individuals bring are not driven by superior ability, but rather by a choice to invest more in breadth of knowledge rather than in depth of knowledge. This distinction is important because, traditionally, it is assumed that the choice to diversify spreads researchers' capabilities too thin across domains and thus precludes an ability to make substantial contributions to science. For this reason, diversified individuals are often labeled as a “jack of all trades and master of none.” Our arguments highlight a process that demonstrates that, in fact, these individuals play a critical role in the production of knowledge by helping to successfully integrate new distant knowledge into innovation.

[7]Theorizing comprehensively on each cultural dimension is outside the scope of our study. Rather, our goal is to shine a light on culture as a potential influencing factor and set the stage for a deeper dive into these dynamics in future research.

3 | DATA AND EMPIRICAL STRATEGY

To test our hypotheses, we follow a three-step strategy. First, we focus on individual researchers and their innovative output as captured in academic publications. We do so not only to align with our theoretical argument at the individual researcher level but also to gain access to a reliable and measurable paper trail of innovation. Second, we focus on new technology developments as embodiments of new knowledge in line with Mokyr's (2002) arguments that technology facilitates access to knowledge that was previously inaccessible because it either did not exist or was costly to access.

Third, we recognize that, in observational data, a correlation between individual levels of knowledge diversification and a propensity to engage with new external technological developments can be driven by unobserved factors. Specifically, if we are to observe diversified individuals having a higher propensity to engage with such a technology it is unclear if these individuals' level of diversification facilitates that successful engagement, or if the individuals strategically chose to focus on certain domains of knowledge that were promising in facilitating the engagement, or that other unobserved factors, such as ability, drove both the level of diversification and the successful engagement. Ideally, we would like to observe individuals of equal ability and other relevant characteristics but varying levels of diversity being exposed to a new and exogenous technology development that is outside their domains of knowledge, and then estimate if individuals with higher levels of diversity exhibit a higher propensity to successfully integrate the new technology in knowledge creation.

We attempt to get close to this ideal setting through a two-step empirical strategy. First, we exploit the unexpected use of Microsoft Kinect in research as a new technology development in motion-sensing research. Second, we observe the impact of this technology on a sample of researchers where we hold ability constant but allow for varying levels of knowledge diversity. We focus on ability as the main factor that might confound the effect of diversity in line with findings in prior literature that show that ability is a precursor for successful exploration and in line with our theory focused on evaluating the role of individual-level diversity in successful exploration as a mechanism through which individual-level ability to explore manifests. At the same time, we acknowledge and discuss the possibility of other relevant unobserved factors that limit causal interpretations of our findings.

3.1 | Kinect

Microsoft Kinect was launched on November 4, 2010 as an add-on to the Xbox 360 video game system. It allowed users to interact with the games through body gestures rather than using a hand-held controller, similar to the competing devices from Nintendo, the Wii Remote, and from Sony, the PlayStation Move. While both the Wii Remote and the PlayStation Move operated via gesture-recognition strategies, the Kinect was a significant leap forward, as it moved the gesture recognition from a single tracking point to full-body 3D motion capture, along with facial, gesture, and voice recognition. Therefore, Kinect is a significant advance in the knowledge frontier as a physical embodiment of new knowledge in motion-sensing, in line with Mokyr's (2002) arguments on the role of technology in capturing and providing access to the knowledge embedded in its algorithms.

Furthermore, the role of Kinect as motion-sensing research technology was not anticipated by the research community. Although Kinect was launched with great anticipation, at no time before the launch did Microsoft or any other party promote, link, or suggest using the Kinect technology outside its intended purpose as a gaming device. The starting point of the unexpected adoption of Kinect in research can be traced back to the bounty offered by AdaFruit Industries on the very day of Kinect's launch. AdaFruit, an electronics hobbyist company influential in the open hardware community, offered a bounty in search of someone who could develop and distribute an open source driver for Kinect. The driver would make it possible for researchers and enthusiasts to access the Kinect motion-sensing algorithms and use them to integrate with any project of their choice.

Hours after AdaFruit made the search for an open source driver public, Microsoft voiced its disapproval on CNET, saying that it "*does not condone the modification of its products …. With Kinect, Microsoft built in numerous hardware and software safeguards designed to reduce the chances of product tampering. Microsoft will continue to make advances in these types of safeguards and work closely with law enforcement and product safety groups to keep Kinect tamper-resistant*" (Terdiman, 2010). AdaFruit did not withdraw the contest but rather tripled its bounty. Six days later, on November 10, 2010, a Spanish technology enthusiast, Hector Martin Cantero, released an open source driver and won the bounty (BBC News, 2010). As the unexpected Kinect effect in research began to rapidly unfold, Microsoft recognized the benefit of Kinect for research and, essentially, approved of its use for such purposes although this was not the original intention.[8]

3.2 | Data collection

We collect data on academic publications of researchers in computer science, electrical engineering and electronics, as available through IEEE Xplore, the bibliographical database maintained by the Institute of Electrical and Electronics Engineers (IEEE). We collect data on every academic publication, early-access publication, and conference proceeding during a 14-year period from 2001 to 2014 (inclusive), resulting in 2,492,451 publications.

We estimate the propensity of diversified researchers to engage with Kinect based on an 8-year subset of this data, from 2007 to 2014 (1,776,125 publications). This represents 4 years of data before and 4 years after the launch of Kinect. The estimation period is substantial considering that the publication cycle is fairly short in computer science, electrical engineering and electronics, and conference proceedings are often the primary outlet for disseminating knowledge in these fields. We use the remainder of the data (2001–2006) to better estimate researchers' experience in academic research measured as number of years of active publication since 2001 (researchers' age).[9]

Next, we construct our dataset at the individual level, while taking advantage of the IEEE-curated unique author identifiers. IEEE identifies 1,391,313 unique names authoring over the period 2001–2014. We restrict our analysis to researchers who publish at least one paper in the 4-year period before Kinect's launch (2007–2010), which reduces the sample to 342,872 researchers. We focus on this subset for two main reasons. First, we want to ensure that our estimations account for researchers' pre-Kinect productivity, which is important for our strategy of controlling for ability. Second, we need to observe researchers for a period before Kinect's launch to determine their degree of diversification across research areas, which is our primary variable of interest.

[8]Researchers engaged with Kinect in a broad set of projects with applications ranging from security to market research, improving the ability of robots to navigate complex landscapes and sudden changes in scenery, helping individuals with impaired abilities, such as allowing the blind to hear an accurate and timely description of their surrounding environment as they attempt to walk within a room, and improving medical procedures, such as the ability to track cameras traveling within a patient during surgery, or simulating custom joint prosthetics.

[9]In an ideal scenario, we would know how many years it has been since a researcher finished their degree and became research active. However, data limitations prevent this. Therefore, we capture age via how long a researcher has been active during the complete period of time we observe in our data (i.e., starting in 2001). Although this is not perfect, it does help control for whether or not someone just started doing research at the beginning of our sample window (research age = 1), or if they have been research-active for 10+ years (research age = 10), or anything in between.

Within this group, we further reduce the number of authors in our dataset by eliminating outlier author IDs that have more than 50 or fewer than three publications in the 4-year period before Kinect's launch. We eliminate researchers with fewer than three publications to ensure that our results on diversification are not driven by comparisons with unproductive individuals who would, mechanically, appear as researchers with a low degree of diversification. This is an important early step towards obtaining our final sample that aims to control for researchers' ability. Note that the group of researchers with fewer than three publications includes occasional authors, such as industry partners and researchers from other domains outside computer science, electrical engineering and electronics. There are 156,688 researchers with fewer than three publications in the 4-year period before Kinect's launch. We also eliminate researchers with more than 50 publications in the 4-year period to ensure that our results are not driven by outliers on the higher end of the productivity spectrum.[10] We limit the maximum number of publications to 50 to align with the anecdotal view of realistic productivity in the fields of computer science, electrical engineering and electronics. There are 3,200 researchers with over 50 publications in the 4-year period before Kinect's launch, less than 1% of the sample. The resulting sample includes 182,984 researchers. Our results remain robust to considering lower or higher cut-off values, including using the full sample.[11]

3.3 | Sample construction and empirical strategy

We are interested in identifying how individual levels of knowledge diversification influence the propensity to engage with the new Kinect technology in a sample of researchers with comparable degrees of ability at the time of the Kinect's arrival. We infer individual-level diversification from researchers' breadth of academic publications across knowledge areas. We measure engagement with Kinect by tracking researchers' publications that reference this technology after its arrival. We restrict to comparable degrees of individual-level pre-Kinect ability through a combination of estimation controls and a matching procedure.

We start by using two features of the IEEE database: (a) the ability to search the full text of all publications included in the IEEE bibliographical database, and (b) the fact that IEEE assigns a limited set of keywords to publications out of a controlled hierarchical vocabulary of nearly 9,000 words. The first feature of the IEEE Xplore database helps us identify publications that refer to Kinect. We search the full text and metadata of all publications included in the IEEE using the keyword "Kinect."[12] Next, we label authors of at least one such identified Kinect publication as a Kinect author, that is, a researcher who successfully engaged with the new technology. All other researchers in our dataset are labeled non-Kinect authors. We also use the search feature of the IEEE database to identify researchers with knowledge in motion-sensing, the knowledge domain to which Kinect belongs. This is needed to distinguish between researchers for whom the Kinect represents local knowledge and those for whom it represents knowledge outside their domains, a distinction that is key to our theoretical arguments. To do so, we repeat our search in the full text and metadata of all publications in our dataset using a set of keywords that describe motion-sensing research topics. We

[10]In addition, this helps address concerns related to potentially inaccurate name disambiguation in the IEEE database that might incorrectly assign individuals with the same name to the same author identifier. Such an error is not uncommon in bibliographical databases and generally occurs when the names are very common. We carefully review the set of authors with more than 50 publications and observe that the list indeed is composed of common names.

[11]All additional robustness results not included in the manuscript, mentioned here and thereafter, are not shown due to space constraints but are available from the authors upon request.

[12]In our robustness tests, we also use a more restrictive definition of Kinect publications whereby we search only in the metadata for the keyword "Kinect." The results remain consistent.

follow the same set of keywords in Teodoridis (2018). The keywords were carefully selected through conversations with experts and cross-referenced with IEEE's taxonomy. We focus on the 4-year period before Kinect's launch (2007–2010) since we seek to identify researchers who had or did not have local domain knowledge at the time the new knowledge embodied in the Kinect device became available. We label authors with at least one such identified motion-sensing publication as a motion-sensing author, and all other as nonmotion-sensing authors, that is, researchers for whom Kinect is new knowledge that is outside their prior domains of knowledge.

The second feature of the IEEE Xplore database helps us calculate an index of diversification at the individual researcher level. The IEEE taxonomy groups publications under 51 main research areas (Appendix Table A1). We focus exclusively on the IEEE set of research areas because the taxonomy provides a stable and thus tractable classification of scholars' research portfolio areas. Furthermore, our estimates are conservative using this approach since the research areas defined under the IEEE taxonomy are at the highest level in the taxonomy. To calculate the individual-level diversity index, we begin by collecting all IEEE-assigned keywords per author for the 4-year period before Kinect's launch (2007–2010).[13] We only use the period before Kinect's launch since the focus is on estimating the role of individual-level diversification in the propensity to engage with new knowledge brought by the launch of Kinect. As such, the relevant individual-level characteristics are the ones observed before the arrival of Kinect. Next, we refer to the IEEE's taxonomy to identify the main research area (out of the 51 IEEE-identified areas) for each keyword. We proceed by constructing a list of main research areas per author and the corresponding keywords used in his/her publications. With these data, we construct a measure of diversification of research areas at the individual level that adjusts for the fact that the probability of diverse keywords increases with the number of publications per author. First, we measure the frequency of occurrence of each research area at the author level for publications between 2007 and 2010. Specifically, we count the total number of keywords assigned to each of the 51 IEEE top categories across all authors' papers published between 2007 and 2010. Next, we convert the count to percentages and calculate the Euclidian distance in the multidimensional space of the 51 research areas.[14] We focus on percentages rather than counts of publications because our goal is to capture variation in knowledge breadth while controlling for within variation in knowledge depth. Note that, by construction, the measure is less than or equal to 1 and is never 0. The measure is lowest when the percentages per research area are equally spread or when the level of diversification of research portfolio areas is highest. Thus, for mathematical convenience, we construct the diversification measure to be equal to 1 minus the calculated Euclidian length. The higher the value, the higher the diversity of research areas at the individual level:

$$DiverseIndex_i = 1 - \sqrt{\sum_{k=1}^{51} CategoryPercentage_{ik}^2} \tag{1}$$

where i is the individual researcher and $CategoryPercentage_{ik}^2$ represents the squared percentage of keywords of researcher i in each category k of the 51 high-level categories of the IEEE taxonomy.

Last, we restrict our estimations to a set of researchers who exhibit similar levels of ability before the launch of Kinect. To construct our sample, we employ the CEM method (Iacus et al., 2011, 2012) which pairs individuals based on specified characteristics. In our case, the goal is to pair

[13]The IEEE taxonomy remains unchanged over this period.

[14]By definition, Euclidian distance is equal to the square root of the Herfindahl index. The results remain robust when considering a diversification measure based on the Herfindahl index alone.

individuals who did and did not successfully engaged with Kinect after its launch based on their observed ability in the period before the launch. This approach allows us to compare individuals of equal ability and to observe if the individuals who successfully engaged with Kinect are the ones characterized by higher levels of diversity before Kinect. In our CEM procedure, we capture the ability level of individuals before Kinect through a total of nine attributes: four covariates representing the total number of publications weighted by citations[15] for each researcher, per year, for the 4 years before Kinect's launch (e.g., one covariate for each year from 2007–2010); four covariates representing the total number of co-authors for each researcher, per year, for the 4 years before Kinect's launch (2007–2010); and the research age of each individual calculated as the number of years since the first observed publication in our large dataset going back to 2001. We also include the total number of publications weighted by citations over the entire 4-year period before Kinect's launch (2007–2010), the total number of co-authors over the same period, and age squared as controls in all of our estimations to capture any remaining variation from pre-Kinect time trends and nonlinear effects of age that are not captured by our CEM procedure. Furthermore, we consider a CEM procedure with weights to make use of as much of our data as possible; CEM with weights considers a richer set of matched individuals based on both exact matches of paired individuals as well as pairs where the match is constructed with weights when an exact match does not exist.[16] Following this approach, we obtain a CEM sample comprised of 104,587 authors. Our results remain robust to considering only the subset of exact matches, albeit with some loss of statistical power and hence ability to more robustly interpret coefficients due to the smaller number of observations.

Taken together, we believe these sample construction steps help us generate a dataset that comes close to the ideal setting where individuals of equal ability, but varying levels of diversity, are exposed to a new technology development. At the same time, we recognize that ability is a complex attribute to accurately capture. Although we proxy for ability using individuals' observed research output in line with prior research (e.g., Azoulay, Stuart, & Wang, 2013; Conti, Gambardella, & Mariani, 2013; Waldinger, 2012) and well-known norms in research evaluation procedures such as tenure decisions, we recognize that additional attributes, such as place of graduation, history of employment, grants and other awards would have been enriching. Unfortunately, we do not have data on such additional attributes, but we believe that the citation-weighted publication portfolio is a telling proxy of individuals' ability to conduct research, one that also implicitly captures the benefits of training, intellectual capacity, and other factors that might be correlated with ability, and one that we exploit in multiple ways in our sample construction to incorporate as much of its richness as possible.

Specifically, and first, our measure of diversity has a built-in mechanism to avoid the trap of mechanically confounding an increase in volume of publication with an increase in diversity. For example, with our measure, a researcher with a publication portfolio of 10 papers in 10 different research areas will have the same calculated index of diversification as another researcher with

[15]Specifically, we sum up citations and counts of publications, such that each publication is counted as one plus its total number of citations. Robustness checks confirm that the results hold when matching on citations and publications separately, rather than in a combined measure.

[16]Considering that, in research, the norms (and other factors) are such that they incentivize specialization, it is reasonable to assume that the diversified researchers who survive are, in average, more productive than their specialized colleagues. In other words, the fact that not all diversified individuals can be exact matched with specialized individuals characterized by the same level of ability, as per our definition, can be a result of the current incentive structure in science which favors specialized individuals. Thus, by employing weighted matching (CEM), we aim to make the most of the otherwise truncated observational data the nature offers i.e. less able diversified individuals, unlike their less able specialized colleagues, have a higher probability of being eliminated from science given the current set of norms and incentives.

20 publications, two in each of the 10 research areas. Furthermore, this approach ensures that we remain true to our theoretical focus on knowledge breadth versus depth; our focus is on researchers who have the same amount of experience and that experience is either (a) spread across multiple domains (wider breadth, shallower depth) or (b) concentrated in one domain or a narrow set of domains (narrower breadth, deeper depth). Second, and because this approach is not fault-proof for controlling for individuals' volume of publication (i.e., in our example, to reach this particular level of diversification, a researcher needs a minimum of 10 papers) we turn to the CEM procedure to ensure that the effect we measure for our diversified authors is not driven by their potentially higher average volume of publications. In addition, in our CEM procedure, we account for the impact of these publications, as measured through the number of citations received. Third, we extend our matching procedure to account for the research age of the individuals in our sample and for their number of coauthors. We match on age because the ability to produce good research has been shown to increase with experience in research (e.g., Azoulay et al., 2013). We match on the number of coauthors because individuals can increase their number of publications by engaging in collaboration with more individuals. Fourth, we include additional covariates in our regression estimates that control for the total number of citation-weighted publications in the period before Kinect, thus capturing any remaining variation due to, for example, time trends and for a nonlinear effect of age.

Furthermore, our CEM approach is more conservative than regression estimates that include the covariates used in the matching procedure as controls alone. The reason for this stems from the CEM process that excludes from the matched sample those individuals for who a suitable exact or weighted match could not be located. This is important because it ensures that our sample includes a counterfactual for each exploring researcher included in the sample. Absent this approach, an estimation using regressions with controls alone could provide results driven by outliers in the group of exploring researchers for whom a comparable nonexploring researcher does not exist. Indeed, when employing this non-CEM approach as a robustness check the results are consistent in sign, however, as expected, the magnitude increases driven by outliers otherwise excluded from our matched sample due to a lack of a matched ability counterfactual.

Our main estimating equation is a cross-sectional probability model with CEM weights:

$$KinectAuthor_i = I(\alpha DiverseIndex_i + \beta DiverseIndex_i * MotionSensingAuthor_i + \theta X_i + \varepsilon_i) > 0 \qquad (2)$$

where i is the individual researcher and X_i is a vector of control variables that includes the indicator variable $MotionSensingAuthor_i$ equal 1 for researchers who published at least one motion-sensing paper before Kinect's launch, between 2007 and 2010, and 0 otherwise. In addition, X_i includes a variable capturing the affiliation of researchers, either in the public or private sector. We collect this information based on the affiliation listed in the IEEE profile of researchers in the period before Kinect's launch. We locate and confirm the affiliation of 83,983 individuals in our sample and we include a dummy variable to account for the remaining cases where we could not verify researchers' affiliation. We further distinguish between researchers with an industry affiliation, a total of 18,669 individuals, and academic researchers, a total of 65,314 individuals. We do so to confirm that our results are not a phenomenon that occurs only in an academic environment but is representative of research behavior in both industrial and academic settings (Sauermann & Stephan, 2013).

The dependent variable is an indicator variable equal to 1 for author i who publishes at least one paper referencing Kinect during the 4-year period after Kinect's launch (2011–2014), and 0 otherwise. The coefficient of interest α captures the propensity of ability-matched diversified researchers, identified as such based on their publication behavior before the launch of Kinect, to refer to Kinect in their academic publications after the launch. We interpret a positive estimated value of this coefficient as indicating that a higher level of diversification of research portfolio areas in the period before Kinect predicts a higher propensity to engage with the new knowledge brought about by the arrival of Kinect. The coefficient β captures how the effect of diversification manifests for researchers with prior local knowledge, namely for researchers with prior experience in motion-sensing, the knowledge domain of Kinect.

Our approach is not without limitations and causal interpretations should be made with care. Specifically, and in addition to the already mentioned limitations, it is possible that other attributes that are relevant for researchers' propensity to explore remain unobserved in our empirical strategy. We believe our empirical strategy captures the most central attributes, namely those that proxy for individuals' ability, thus allowing us to get closer to drawing causal implications but not without limitations. Furthermore, it is important to note that the boundaries of our attempts to get closer to causal estimates end with our evaluation of the role of diversified individuals in exploration. We do not deny the role of certain unobservable attributes, such as curiosity or a taste for diversification, or that of other factors, such as employer incentives that might explain why certain individuals become diversified in the first place. In fact, we explicitly theorize about and empirically explore the role of personal preferences and that of employer incentives as potential antecedents to diversification.

3.4 | Descriptive statistics

We conduct all our main estimations on the matched sample but present descriptive statistics for the full sample as well. From the full sample of 182,984 researchers, we identify 4,705 who published at least one Kinect paper during the period 2011–2014. The remaining 178,279 researchers represent the full sample of non-Kinect authors. Table 1, Panel 1 shows that Kinect authors are generally more productive than non-Kinect authors during the 4-year period before Kinect's launch (2011–2014). Specifically, Kinect authors publish more papers, receive more citations, and have more co-authors than non-Kinect authors do. Furthermore, Kinect authors also exhibit a higher level of diversification and are on average 1 year older than non-Kinect authors.

These differences, some of which are most likely attributable to the larger variance in the non-Kinect author sample, motivate the use of the CEM methodology to ensure that our group of Kinect authors and our group of non-Kinect authors are comparable in ability as measured by productivity, number of co-authors, and age in the period before Kinect's launch. Furthermore, the differences foreshadow some of our theoretical arguments about the process through which the benefits of diversification manifest, such as larger networks of collaborators,[17] and about the antecedents of diversification, such as publication age, and thus further motivate the use of the CEM methodology. In other words, to get closer to causally identifying the impact of diversity on the propensity to engage with the new technology, it is important to isolate the role of diversification from other factors that vary with diversity, including both diversity-inducing and other diversity-consequent factors.

Table 2 shows the CEM balance on all covariates included in the matching. Table 1, Panel 2 shows the same descriptive statistics as Panel 1 but for the matched sample. The matching reduces the number of Kinect authors to 2,994 and the number of non-Kinect authors to 101,593. In this sample, both groups of authors have similar levels of productivity in the period before Kinect's launch. However, as preliminary support for our arguments, the difference in diversification persists.

In our estimations, we present results using three measures of diversification. First, we show results using a continuous measure of diversification equal to our diversification index calculated

[17]Although not the main focus of our paper, we empirically explore the role of collaboration in more detail in Appendix B.

TABLE 1 Descriptive statistics

Variable	Panel 1: Unmatched sample					Panel 2: Matched sample (CEM)				
	Observations	Min	Max	Mean	*SD*	Observations	Min	Max	Mean	*SD*
Diversification measure										
All authors	182,984	0	0.804	0.645	0.074	104,587	0	0.793	0.646	0.072
Kinect authors	4,705	0	0.798	0.688	0.058	2,994	0	0.790	0.674	0.060
Non-Kinect authors	178,279	0	0.804	0.644	0.074	101,593	0	0.793	0.645	0.073
Number of publications 2007–2010										
All authors	182,984	3	50	8.493	7.587	104,587	3	50	7.318	4.926
Kinect authors	4,705	3	50	13.091	10.934	2,994	3	50	7.398	4.953
Non-Kinect authors	178,279	3	50	8.371	7.441	101,593	3	44	7.316	4.925
Number of co-authors 2007–2010										
All authors	182,984	3	1,073	41.356	48.709	104,587	3	511	29.855	22.253
Kinect authors	4,705	4	662	58.956	60.652	2,994	4	509	29.742	22.324
Non-Kinect authors	178,279	3	1,073	40.892	48.268	101,593	3	511	29.859	22.251
Number of citations 2007–2010										
All authors	182,984	0	1,612	17.952	39.462	104,587	0	198	11.992	16.876
Kinect authors	4,705	0	900	33.735	57.518	2,994	0	179	12.082	16.691
Non-Kinect authors	178,279	0	1,612	17.536	38.785	101,593	0	198	11.989	16.881
Number of citations-weighted pubs										
All authors	182,984	3	1,662	26.445	43.813	104,587	3	217	19.310	19.186
Kinect authors	4,705	3	932	46.826	63.566	2,994	3	217	19.480	19.127
Non-Kinect authors	178,279	3	1,662	25.907	43.039	101,593	3	206	19.187	19.187
Author age 2001–2010										
All authors	182,984	1	10	6.572	2.808	104,587	1	10	6.973	2.699
Kinect authors	4,705	1	10	7.621	2.556	2,994	1	10	7.005	2.709
Non-Kinect authors	178,279	1	10	6.545	2.809	101,593	1	10	6.972	2.699

Abbreviation: CEM, coarsened exact matching.

TABLE 2 Coarsened exact matching (CEM) balance

CEM balance	Full sample			Matched sample (CEM)		
	Kinect authors	Non-Kinect authors	*t*-Stat.	Kinect authors	Non-Kinect authors	*t*-Stat.
Citation-weighted publication count 2007	15.324	8.015	23.48	6.137	6.131	0.03
Citation-weighted publication count 2008	15.035	8.305	23.13	5.802	5.773	0.17
Citation-weighted publication count 2009	8.694	5.038	26.02	3.788	3.732	0.64
Citation-weighted publication count 2010	7.773	4.549	28.91	3.753	3.669	1.09
Co-author count 2007	13.726	9.537	14.73	6.369	6.490	0.72
Co-author count 2008	14.336	10.319	16.50	7.193	7.328	0.78
Co-author count 2009	14.638	9.999	22.12	7.478	7.458	0.14
Co-author count 2010	16.255	11.036	23.62	8.702	8.583	0.73
Total citation-weighted publication count 2007–2010	46.826	25.907	32.42	19.480	19.305	0.49
Total co-author count 2007–2010	58.956	40.892	25.15	29.742	29.859	0.28
Author age	7.621	6.545	25.99	7.005	6.972	0.64
Observations	4,705	178,279		2,994	101,593	

using Equation (1). Second, we create a dummy measure of diversification equal to 1 if the focal researcher has an index of diversification in the top half of the distribution of the diversification index of all authors, namely above 0.646, and 0 otherwise. Third, we create a set of quartile dummies of diversification where the omitted category is the bottom 25th percentile of the diversification distribution (bottom quartile). Specifically, the omitted category is composed of researchers with a diversification index below 0.596. The quartile of diversification in the bottom 25th to 50th percentiles of the distribution (second quartile) is composed of researchers with a diversification index above 0.596 but below 0.646. The quartile of diversification in the top 50th to 75th percentiles of the distribution (third quartile) is composed of researchers with a diversification index above 0.646 but below 0.687. Finally, the quartile of diversification above the top 75th percentile (top quartile) is composed of researchers with a diversification index above 0.687.

4 | RESULTS

4.1 | Diverse researchers and the propensity to successfully engage with new external knowledge

The results shown in Table 3 are consistent with H1, that among individuals for whom Kinect represents new knowledge that is outside of their prior domains of knowledge, that is, nonmotion-sensing researchers, those with a higher degree of diversification have a higher propensity to engage with Kinect in research. Specifically, estimates of a logit model using our three measures of diversification show that the propensity to write Kinect papers increases with increased diversification ($p < .01$ in all models), and that the effect holds only for researchers who were not involved in motion-sensing prior to the launch of Kinect (the coefficient of the diversification measure interacted with an

TABLE 3 Diversification and the propensity to write Kinect papers

DV = 1 if author published at least one Kinect paper and 0 otherwise; matched sample						
	Continuous diversification		**Above median dummy diversification**		**Quartiles of diversification**	
Diversification before Kinect (2007–2010)	0.059 (0.004)	0.068 (0.004)	0.653 (0.047)	0.738 (0.050)		
Diversification before Kinect in bottom 25th to 50th percentiles					0.413 (0.078)	0.415 (0.084)
Diversification before Kinect in 50th to 75th percentiles					0.744 (0.074)	0.758 (0.080)
Diversification before Kinect in 75th to 100th percentiles					1.021 (0.072)	1.146 (0.076)
Motion-sensing author	1.786 (0.065)	5.639 (0.763)	1.826 (0.065)	2.260 (0.118)	1.793 (0.065)	2.339 (0.204)
Diversification before Kinect (2007–2010) × Motion-sensing author		−0.056 (0.011)		−0.538 (0.140)		
Diversification before Kinect in bottom 25th to 50th percentiles × Motion-sensing author						−0.174 (0.251)
Diversification before Kinect in 50th to 75th percentiles × Motion-sensing author						−0.313 (0.239)
Diversification before Kinect in 75th to 100th percentiles × Motion-sensing author						−0.797 (0.224)
Total citation-weighted publications before (2007–2010)	−0.001 (0.002)	−0.001 (0.002)	−0.001 (0.002)	−0.001 (0.002)	−0.001 (0.002)	−0.001 (0.002)
Total co-authors before (2007–2010)	−0.007 (0.002)	−0.007 (0.002)	−0.005 (0.002)	−0.005 (0.002)	−0.006 (0.002)	−0.006 (0.002)
Author age	−0.045 (0.066)	−0.062 (0.065)	−0.040 (0.066)	−0.044 (0.065)	−0.045 (0.066)	−0.056 (0.066)
Author age sq	0.002 (0.004)	0.003 (0.003)	0.002 (0.004)	0.002 (0.004)	0.002 (0.004)	0.003 (0.004)
Unable to obtain affiliation (flag)	−1.293 (0.090)	−1.299 (0.090)	−1.312 (0.089)	−1.314 (0.089)	−1.303 (0.090)	−1.309 (0.089)
University affiliation (flag)	0.089 (0.055)	0.089 (0.055)	0.097 (0.055)	0.096 (0.055)	0.089 (0.055)	0.090 (0.055)
LL	−12,488.84	−12,467.13	−12,527.51	−12,516.69	−12,494.90	−12,474.94
Observations	104,587	104,587	104,587	104,587	104,587	104,587

Note: The data is a cross-section at the author level. All models are logit with robust standard errors. Estimations presented as coefficient (*SE*).

indicator variable capturing involvement in motion-sensing in the period before Kinect has a large 95% confidence interval that includes zero, meaning the effect is indistinguishable from zero). We include baseline effects without interaction terms in columns 1, 3 and 5, and estimations including interaction terms in columns 2, 4 and 6. Columns 1 and 2 show estimates using our continuous measure of diversification based on researchers' publication portfolios in the 4 years before Kinect's launch (2007–2010), as described in Equation (1). In columns 3 and 4, we replace the continuous measure of diversification with a dummy variable equal to 1 if the researcher ranks above the median level of the diversification distribution. In columns 5 and 6, we further break down this covariate to capture the effect on quartiles across the diversification distribution. The results indicate that researchers above the median level of the diversification distribution are 2.1 times more likely to write a paper using Kinect than researchers with a diversification level below the median. Furthermore, the effect increases linearly with the magnitude of the diversification index. Researchers in the second quartile are 1.5 times more likely to engage with Kinect than researchers in the bottom quartile, while researchers in the third quartile are 2.1 times more likely to do so. Researchers in the top quartile are 3.1 times more likely to include Kinect in their research than researchers in the bottom quartile. In all cases, the effect of diversity is positive only when integrating new distant knowledge. When the knowledge is local, namely for researchers with prior knowledge in motion-sensing (the knowledge domain of Kinect) diversification does not offer an advantage over specialization. This is reassuring as our theory is about the benefits of diversification in integrating new knowledge that is outside the individual's prior knowledge domains; when the knowledge is local, it is expected that generalists will not have an advantage over specialized individuals who are also familiar with that local knowledge space, which allows them to also successfully integrate the new knowledge (Shane, 2000; Teodoridis, Bikard, & Vakili, 2018).

Although we do not directly theorize about the direction and magnitude of these interaction terms, it is important to acknowledge the Ai and Norton (2003) critique, that the sign and magnitude of marginal effect of these terms are not necessarily the same as the sign and the magnitude of the interaction coefficients. We take several steps to demonstrate that the critique does not influence the conclusion we can draw from these estimations. First, we do not rely on interpreting marginal effects but rather interpret odds ratios, a regression output that is free from the Ai and Norton (2003) critique (e.g., Buis, 2010). Second, the Ai and Norton (2003) concern does not hold in the case of nonlinear models with binary interaction terms (Green, 2010; Kolasinski & Siegel, 2010; Puhani, 2012), like the ones in the models presented in columns 4 and 6. Third, we repeat the estimation separately, using split samples, for the groups of researchers with and without prior experience in motion-sensing (Appendix Table A2). We observe that in all models the coefficient of the diversification measures is positive ($p < 0.01$) for the subsample of nonmotion-sensing researchers, while the effect diminishes in magnitude with larger standard errors in the subsample of motion-sensing researchers. Fourth, our results remain robust to using a linear probability model (Appendix Table A3). Angrist and Pischke (2009) show that there is little qualitative difference between a linear probability model and a logit specification, with the advantage that the Ai and Norton (2003) critique does not apply to linear estimation models. Taken together, we argue that these steps are reassuring in our interpretation of the odds ratios of the interaction terms in our main logit specification as supporting evidence for our main effect estimations.

Our Table 3 results persist when employing our alternative method of identifying Kinect papers, based exclusively on metadata searches (Appendix Table A4). This definition of Kinect papers is more conservative than our main specification since it excludes those academic publications that mention Kinect only in the body of the text, but not in keywords, abstract, or title. Not only do our results hold, but the magnitude persists, further strengthening the argument that, within the group of nonmotion-sensing researchers, the more diversified researchers are the ones more likely to engage with Kinect in their research.

Furthermore, since we estimate our models on a cross-sectional dataset where the four publication years after Kinect are aggregated, we also want to ensure our results are not driven by any particular year, especially a year far after the Kinect launch event. As such, we repeat our main estimation on subsets of the data, one for each of the 4 years after Kinect's launch. Here, the dependent variable is 1 if, in a given year, the author published a Kinect paper for the first time (i.e., they had not published a Kinect paper in a previous year). We include these results in Appendix Table A5 and observe that the effect of diversification on the propensity to engage with Kinect begins immediately in 2011, the first year after Kinect's launch, and persists in following years. We present these results using the continuous measure of diversification since this approach is most conservative. Our results remain robust when considering the dummy and quartile covariates.

Next, we turn to testing our second hypothesis by restricting the dependent variable to equal 1 only for those Kinect authors who published at least one Kinect paper ranking in the top 10% of papers in IEEE Xplore by citation count. We identify such highly cited papers relative to the entire population of publications, not only relative to work referencing Kinect.[18] In other words, to ensure that we capture the propensity to produce high-impact research, we identify those Kinect publications that enter the ranks of the top 10% most-cited papers of all papers published in computer science, electrical engineering and electronics between 2011 and 2014, our 4 years of interest after Kinect's launch.[19] To confirm robustness to the definition of "highly cited," we also consider a top 5% threshold, and the results remain substantively similar.

We present results of this one-tailed test of H2 in Table 4 to allow for a more accurate interpretation of the magnitude of the coefficient and then utilize a two-tailed estimation to complete the H2 testing. As before, in columns 1 and 2 we show estimates of a logit model using the continuous measure of diversification based on researchers' publication portfolios in the 4 years before Kinect's launch (2007–2010). In columns 3 and 4, we replace the continuous measure of diversification with a dummy variable equal to 1 if the researcher ranks above the median of the diversification distribution, and in columns 5 and 6 we further break down this covariate to capture the effect on quartiles across the diversification distribution. All models show results that support H2, that researchers with a higher degree of knowledge diversification are more likely to produce high-impact research using new distant knowledge than those with a lower degree of diversification. As before, the sign and magnitude of the interaction terms are reassuring in that they suggest that the effect of diversification is diminished when the knowledge is local. Specifically, nonmotion-sensing researchers above the median of the diversification distribution are 2.5 times more likely to produce highly impactful papers using Kinect than nonmotion-sensing researchers with a diversification level below the median. Furthermore, the effect increases linearly with the magnitude of the diversification index. Nonmotion-sensing researchers in the second quartile are 1.3 times more likely to produce impactful papers using Kinect than nonmotion-sensing researchers in the bottom quartile, while nonmotion-sensing researchers in the third quartile are 2 times more likely to do so and nonmotion-sensing researchers in the top quartile of the diversification distribution are 3.8 times more likely to include Kinect in their research than nonmotion-sensing researchers in the bottom quartile of the diversification distribution. As before, the results persist when turning to our split sample estimation (Appendix Table A6) and to our linear probability estimation (Appendix Table A7). Furthermore, the results are robust to considering our more restrictive definition of Kinect publications, and per-year estimation models.

[18]We confirm that Kinect papers are no more or less likely to be highly cited than papers on other topics.

[19]As discussed above, the publication cycle in these fields is fairly short. Therefore, citations accrue more quickly than in other fields such as management and economics. Hence, a 4-year post-period captures a significant portion of citations.

TABLE 4 Diversification and the propensity to write top-cited Kinect papers

DV = 1 if author published at least one Kinect paper in the top 10th percentile of the citation distribution of all papers published between 2011 and 2014, and 0 otherwise; matched sample

	Continuous diversification		Above median dummy diversification		Quartiles of diversification	
Diversification before Kinect (2007–2010)	0.064 (0.012)	0.074 (0.014)	0.841 (0.130)	0.909 (0.142)		
Diversification before Kinect in bottom 25th to 50th percentiles					0.229 (0.227)	0.285 (0.249)
Diversification before Kinect in 50th to 75th percentiles					0.738 (0.204)	0.706 (0.230)
Diversification before Kinect in 75th to 100th percentiles					1.152 (0.196)	1.326 (0.213)
Motion-sensing author	1.805 (0.131)	5.116 (1.502)	1.839 (0.128)	2.143 (0.299)	1.789 (0.131)	2.418 (0.468)
Diversification before Kinect (2007–2010) × Motion-sensing author		−0.048 (0.022)		−0.362 (0.329)		
Diversification before Kinect in bottom 25th to 50th percentiles × Motion-sensing author						−0.464 (0.605)
Diversification before Kinect in 50th to 75th percentiles × Motion-sensing author						−0.164 (0.525)
Diversification before Kinect in 75th to 100th percentiles × Motion-sensing author						−0.907 (0.497)
Total citation-weighted publications before (2007–2010)	0.013 (0.003)	0.013 (0.003)	0.012 (0.003)	0.012 (0.003)	0.013 (0.003)	0.013 (0.002)
Total co-authors before (2007–2010)	−0.013 (0.004)	−0.013 (0.004)	−0.011 (0.003)	−0.011 (0.003)	−0.013 (0.003)	−0.013 (0.003)
Author age	−0.257 (0.164)	−0.272 (0.164)	−0.265 (0.164)	−0.266 (0.164)	−0.269 (0.165)	−0.286 (0.165)
Author age sq	0.012 (0.009)	0.013 (0.009)	0.012 (0.009)	0.013 (0.009)	0.013 (0.009)	0.013 (0.009)
Unable to obtain affiliation (flag)	−1.484 (0.249)	−1.492 (0.249)	−1.492 (0.249)	−1.496 (0.249)	−1.489 (0.250)	−1.503 (0.250)
University affiliation (flag)	0.005 (0.133)	0.004 (0.133)	0.012 (0.133)	0.010 (0.133)	0.001 (0.133)	−0.001 (0.132)
LL	−2,348.54	−2,346.21	−2,347.46	−2,346.81	−2,341.52	−2,336.93
Observations	104,587	104,587	104,587	104,587	104,587	104,587

Note: The data is a cross-section at the author level. All models are logit with robust standard errors. Estimations presented as coefficient (*SE*).

In testing H2, a remaining concern is that the increased propensity to produce highly cited papers might be an artifact of simply producing more output. To address this issue, we extend our analysis to consider the change in publication propensity at the right tail of the citation distribution relative to changes in the left tail. More specifically, in Appendix Table A8 we replace our dependent variable with an indicator variable equal to 1 if the focal researcher published more Kinect papers ranked in the top rather than in the bottom 10th percentile of the citation distribution and in Appendix Table A9, we consider an alternative dependent indicator variable, equal to 1 if the focal researcher published more cited papers than noncited papers.[20] In all cases, we continue to find support for our hypothesized effects of diversification; the results persist when turning to our split sample estimation (Appendix Tables A10 and A11) and our linear probability estimation (Appendix Tables A12 and A13).

Finally, to complete the test for hypothesis H2, which states that the higher propensity for breakthroughs is achieved because diversified authors combine more categories of knowledge than their less diversified peers, we examine the number of IEEE domains the Kinect papers belong to as evidenced by the keywords used in each paper. Table 5 shows the results of this analysis. We follow our main estimating equation where the dependent variable is equal to 1 if the author published Kinect papers that list a combination of keywords that indicates an above the sample median number of IEEE domains and 0 otherwise. Table 5's columns break down diversification in the same manner as before such that columns 1 and 2 use a continuous measure of diversification, columns 3 and 4 use a binary above/below the median dummy of diversification, and columns 5 and 6 use quartiles of diversification. All columns show that the more diversified an individual is, the higher the propensity to produce a Kinect paper that combines an above the median number of knowledge categories as evidenced by the above the median number of IEEE domains that characterize the paper.

4.2 | Antecedents of diversification

Having established the role of diversification in integrating new distant knowledge, we now turn to shedding some light on the mechanisms behind why a researcher may become diversified in the first place. We theorized about two primary manners that are likely to drive diversification: personal preferences and professional incentives. Testing these aspects comprehensively is difficult and requires far more detailed data than that in our study. Nevertheless, we attempt to explore these mechanisms to the best of our ability. On the personal preferences side, we take a step in this direction and find that (a) diversity in educational background does not impact diversification, (b) diversification is not driven by a choice to specialize in a general-purpose domain,[21] and (c) cultural values appear to influence preferences for diversity. On the professional incentives side, we consider two possible mechanisms—publication age (the number of years since a researcher's first publication) and employer-type (academic vs. industry). We find that (a) more experienced researchers tend to be more diversified, and (b) this is especially the case for academic researchers.

4.2.1 | Personal preferences—Diversity in educational background

To examine the role of variety in topics of study, we attempted to collect detailed resume information for a random sub-sample of 200 individuals from our primary CEM sample (100 random individuals with a diversification level above median, and 100 random individuals with a diversification level below median).We found limited information for 130 of these individuals that contains sufficient details about their programs of study to provide a glimpse into potential difference in taste for

[20]Given the skewed nature of citations, especially in the fields of computer science, electrical engineering and electronics most papers in our sample have zero citations.

[21]The fact that diversification is not driven by individuals who specialize in general-purpose domains also acts as a robustness test to our main findings, in the sense that it demonstrates that our results are driven by truly diversified individuals and not by researchers who happen to work in domains that are more general-purpose, and hence naturally suited to be combined with other domains.

TABLE 5 Diversification and the propensity to write Kinect papers that belong to a broader set of IEEE domains

DV = 1 if author published Kinect papers that belong to an above median (within the set of Kinect papers) number of IEEE categories and 0 otherwise; matched sample

	Continuous diversification		Above median dummy diversification		Quartiles of diversification	
Diversification before Kinect (2007–2010)	0.053 (0.005)	0.062 (0.006)	0.528 (0.060)	0.597 (0.064)		
Diversification before Kinect in bottom 25th to 50th percentiles					0.429 (0.101)	0.424 (0.108)
Diversification before Kinect in 50th to 75th percentiles					0.572 (0.098)	0.593 (0.104)
Diversification before Kinect in 75th to 100th percentiles					0.942 (0.093)	1.030 (0.097)
Motion-sensing author	1.566 (0.077)	5.289 (0.974)	1.611 (0.076)	1.993 (0.146)	1.569 (0.077)	2.035 (0.266)
Diversification before Kinect (2007–2010) × Motion-sensing author		−0.054 (0.014)		−0.477 (0.170)		
Diversification before Kinect in bottom 25th to 50th percentiles × Motion-sensing author						−0.121 (0.319)
Diversification before Kinect in 50th to 75th percentiles × Motion-sensing author						−0.341 (0.306)
Diversification before Kinect in 75th to 100th percentiles × Motion-sensing author						−0.645 (0.286)
Total citation-weighted publications before (2007–2010)	−0.004 (0.002)	−0.004 (0.002)	−0.004 (0.002)	−0.004 (0.002)	−0.004 (0.002)	−0.004 (0.002)
Total co-authors before (2007–2010)	−0.006 (0.002)	−0.006 (0.002)	−0.004 (0.002)	−0.004 (0.002)	−0.005 (0.002)	−0.005 (0.002)
Author age	0.058 (0.087)	0.043 (0.087)	0.067 (0.087)	0.063 (0.087)	0.061 (0.087)	0.054 (0.087)
Author age sq	−0.003 (0.005)	−0.002 (0.005)	−0.003 (0.005)	−0.003 (0.005)	−0.003 (0.005)	−0.003 (0.005)
Unable to obtain affiliation (flag)	−1.337 (0.117)	−1.341 (0.117)	−1.357 (0.117)	−1.359 (0.117)	−1.346 (0.117)	−1.350 (0.117)
University affiliation (flag)	0.002 (0.068)	0.002 (0.067)	0.012 (0.068)	0.012 (0.067)	0.002 (0.068)	0.003 (0.068)
LL	−7,724.56	−7,713.90	−7,754.11	−7,749.58	−7,728.17	−7,721.86
Observations	104,587	104,587	104,587	104,587	104,587	104,587

Note: The data is a cross-section at the author level. All models are logit with robust standard errors. Estimations presented as coefficient (*SE*).

diversification between diverse and specialized researchers. In other words, we sought to gather information on these researchers' undergraduate, masters (if any), and doctoral topics of study to better understand if there is a relationship between diversity of educational training (e.g., different degrees in different topics) and diversity of research. In aggregate, we found no significant difference between the educational paths taken by diversified researchers and those of specialists.

4.2.2 | Personal preferences—Generalness of research field

To understand if diversification is related to a propensity to enter a research field that is more general-purpose, we considered (a) the ratio of diversified individuals to all individuals publishing in an IEEE category, (b) the percentage of keywords in an IEEE domain that appear in another IEEE domain as a measure of how general or specialized a domain is (e.g., if all of the keywords in a domain also appear in other domains, then that domain is highly general whereas a domain that has a very limited number of keywords appearing in other domains is, by comparison, highly specialized), and (c) the correlation between these two measures. Although we do find that there is some variance by domain in the percentage of authors that are diversified, the range is reasonably tight across all IEEE domains (from 52 to 75% with a 5th percentile of 55% and a 95th percentile of 69%) indicating that although some domains have more diversified individuals than others, there are no outliers in either direction. Next, we consider the percentage of keyword overlap between domains and find a wide spread of keyword overlap (from 27 to 94%) which indicates that some categories are more general-purpose than others. However, the distribution of researchers who are more diverse or more specialized has a very low correlation index (of only 0.075) with this measure capturing the degree of specialization versus general-purpose of a domain. In other words, the existence of diverse researchers is not a consequence of the fact that some knowledge domains are more general-purpose than others.

4.2.3 | Personal preferences—Culture

To explore the impact of cultural values on preference for diversity, we follow a two-step approach. First, we identify the likely country of origin for each researcher in our sample. We do so by employing one of the most popular algorithms that infers the country of origin form a given name: NamSor. The algorithm employs a Naïve Bayes classification algorithm and outputs a score that captures the probability that a name reflects a certain country of origin. We are able to obtain such a score for 98,277 out of 104,587 researchers comprising our main CEM sample. Second, we map the resulting likely country of origin to the numerical values of the six Hofstede culture dimensions for that country.[22]

Next, we regress these six culture measures on our continuous measure of diversification, in an ordinary least squares regression using our main CEM sample with the same set of controls utilized in all our other estimations. We include the results of this estimation in Table 6. The first column utilizes the full CEM sample, whereas column 2 restricts the analysis to those researchers for whom the NamSor inferred country of origin is the same as their affiliation country (a more restrictive analysis). Column 3 shows results for the group of researchers for whom the country of affiliation does not match the NamSor inferred country of origin. Researchers for whom we do not have affiliation data are excluded from columns 2 and 3, but included in the column 1 estimations. The split sample results in columns 2 and 3 aim to account for some of the uncertainty in country assignment that might not be captured by the NamSor score and also to provide a glimpse into potential differences in the role of culture for migrants relative to researchers who operate in their country of origin. Furthermore, because NamSor does not provide a clear interpretation of their country identification accuracy score, we repeat our estimation with different cut-offs for this

[22]Data on Hofstead country cultural dimension is available here: https://geerthofstede.com/culture-geert-hofstede-gert-jan-hofstede/6d-model-of-national-culture/

TABLE 6 Diversification propensity is likely influenced by culture

DV = continuous diversification measure; matched sample	Full CEM sample	Only authors for who their affiliation country is the same as the inferred culture country	Only authors for who their affiliation country is not the same as the inferred culture country
Power distance	0.092 (0.018)	0.154 (0.045)	−0.016 (0.026)
Individualism	−0.110 (0.028)	−0.074 (0.038)	−0.096 (0.049)
Masculinity	0.040 (0.023)	−0.056 (0.065)	0.075 (0.036)
Uncertainty avoidance	−0.058 (0.012)	−0.098 (0.025)	−0.035 (0.020)
Long-term orientation	−0.023 (0.019)	0.011 (0.028)	−0.005 (0.030)
Indulgence	0.087 (0.032)	0.078 (0.040)	0.115 (0.058)
Author age	0.251 (0.019)	0.230 (0.028)	0.303 (0.028)
Motion-sensing author	3.100 (0.226)	3.319 (0.336)	2.785 (0.329)
Total citation-weighted publications before (2007–2010)	0.010 (0.006)	−0.011 (0.009)	0.016 (0.007)
Total co-authors before (2007–2010)	0.062 (0.006)	0.090 (0.008)	0.068 (0.008)
Unable to obtain affiliation (flag)	−1.128 (0.188)	N/A	N/A
University affiliation (flag)	0.418 (0.140)	0.263 (0.202)	0.552 (0.191)
R^2	0.089	0.116	0.102
Observations	98,277	41,261	36,184

Note: The data is a cross-section at the author level. All models are ordinary least squares with robust standard errors. Estimations presented as coefficient (*SE*).
Abbreviation: CEM, coarsened exact matching.

score. All results remain consistent. This estimation approach is certainly not without limitations, especially given the fact that (a) algorithms such as NamSor are more accurate for certain countries than for others, (b) the population of some countries, such as that of the United States, is characterized by a high heterogeneity in names of different origins, and (c) the population of scientists exhibits relatively high levels of mobility. For these reasons, we limit our interpretation of the results to noting that culture does seem to play a role in influencing the propensity of individuals to diversify, although the results are only correlational. We hope that our results motivate other scholars to investigate these issues in a more detailed and robust manner than our data permits.

4.2.4 | Professional incentives—Age and university affiliation

We theorized about two primary characteristics to consider on the potential role of professional incentives for diversification: publication age (the number of years since a researcher's first publication) and employer-type (academic vs. industry). Table 7 shows the results of this analysis. Column 1 shows that, on average, each one additional year spent in the research profession is correlated with a 0.3% ($p < .01$) increase in the level of diversification as captured by our Euclidian measure. Column 2 shows that the role of age in leading to higher levels of diversification is slightly concave (and not linear), such that the increase in diversity with age slows over time. However, the inflection point of the parabola is outside of the range of the age variable indicating that the relationship is one of diminishing returns rather than an inverted-U. Additionally, both columns 1 and 2 show that the propensity to diversify is higher for researchers employed by universities than for researchers working in the private sector ($p < .01$). In column 3 we include an interaction term between age and university affiliation to explicitly show that the positive relationship between age and diversification is more than twice as strong for academic affiliated researchers as it is for industry affiliated researchers. To understand the role of academic tenure in this process, in column 4, we look at a binary split of author age at 7 years. Normal tenure cycles take between 5 and 7 years from first publication in the IEEE domain, so

TABLE 7 Diversification propensity increases with publication age, albeit with diminishing returns; the effect is strongest for academics

DV = continuous diversification measure; matched sample				
Author age	0.164	0.899	0.514	
	(0.018)	(0.129)	(0.247)	
Author age sq		−0.040	−0.022	
		(0.007)	(0.014)	
Author age × University affiliation (flag)			0.592	
			(0.285)	
Author age sq × University affiliation (flag)			−0.027	
			(0.016)	
Author age 7 years or more				0.568
				(0.176)
Author age 7 years or more × University affiliation (flag)				0.548
				(0.210)
Motion-sensing author	2.970	2.988	2.975	2.990
	(0.217)	(0.218)	(0.218)	(0.218)
Total citation-weighted publications before (2007–2010)	0.007	0.007	0.007	0.008
	(0.006)	(0.006)	(0.006)	(0.006)
Total co-authors before (2007–2010)	0.066	0.066	0.066	0.067
	(0.006)	(0.006)	(0.006)	(0.006)
Unable to obtain affiliation (flag)	−0.932	−0.934	−0.940	−0.947
	(0.180)	(0.180)	(0.180)	(0.179)
University affiliation (flag)	0.733	0.738	−2.285	0.289
	(0.136)	(0.136)	(1.146)	(0.183)
R^2	0.075	0.076	0.076	0.074
Observations	104,587	104,587	104,587	104,587

Note: The data is a cross-section at the author level. All models are ordinary least squares with robust standard errors. Estimations presented as coefficient (*SE*).

this cutoff can represent an important shift in incentives from a situation where a young researcher is encouraged to stay highly focused on one domain before they obtain tenure, to when they are free to pursue what they like after tenure. The estimation indicates that the level of diversification of university researchers increases by 1.76% ($p < .01$) after the 7-year mark relative to before, whereas that of researchers employed in the private sector increases only by 0.89% ($p < .01$). These results are robust to other cutoffs between 5 and 9 years after first publication.

Because our measure of age is from the time of publication but our sample only starts in 2001, we are likely underestimating the true research age of researchers that had already published prior to 2001. Therefore, as a robustness check, we repeat our analysis on the subset of researchers with no publications in 2001 or 2002. Since most researchers in the IEEE domains publish at least one paper a year, by only looking at individuals that do not publish in 2001 or 2002, we are likely capturing only researchers that publish in 2003 or later for the first time. Hence, our calculated publication age is likely a more accurate measure in this subsample. We include these results in Appendix Table A14. The results are consistent with those in Table 7 showing that although this concern may be valid, it does not substantially alter our primary findings. Furthermore, the effect of age on diversification for researchers in the private sector is further reduced in line with our theoretical arguments. In aggregate, these results show that researchers' decision to diversify their research interests are indeed shaped by their professional incentives.

5 | DISCUSSION AND CONCLUSION

We examine the role of individual knowledge diversification in integrating new distant knowledge—that is outside of the individual's domains of knowledge. Our study is motivated by the central, yet understudied, role of the individual in influencing the innovation performance of organizations. We focus on exploration, an endeavor that was shown to contribute to the competitive advantage of organizations (Barth et al., 2017; Cohen & Levinthal, 1990) and find evidence consistent with our hypotheses that diversification is a mechanism through which skilled individuals achieve a higher propensity to successfully engage with new distant knowledge and do so in a manner that produces highly impactful output. Furthermore, we show that the highly impactful publications authored by these diverse individuals are generated by combining more diverse pieces of knowledge. In addition, we consider several factors that might explain why a skilled researcher might choose to become diversified in the first place and show evidence suggesting that diversification is both a result of personal preferences and professional incentives.

To shed light on the role of individual-level diversification as a mechanism through which skilled individuals engage in successful exploration, our empirical strategy attempts to get close to an ideal setting where equally skilled individuals with varying levels of diversity are exposed to the arrival of a new knowledge that is outside their domains of expertise. To do so, we construct a sample where we evaluate the propensity of ability-matched diversity-varying researchers in computer science, electrical engineering and electronics to publish academic papers referencing Kinect, a technology that arrived unexpectedly in motion-sensing research. We focus on the Kinect technology developments as embodiments of new knowledge in line with Mokyr's (2002) arguments that technology facilitates access to knowledge that is otherwise inaccessible because it did not exist or because it was costly to access.

While our empirical strategy offers certain benefits that allow us to get closer to causal interpretations of the findings, the approach is not without limitations. First, there might be theoretically-relevant, empirically unobserved attributes that we cannot capture in our analysis. By following prior research and norms in academic evaluations to proxy for researchers' ability using their observed research output, number of collaborators and research age we believe we have captured the most relevant factors. Second, our analysis is conditional on observed selection into different levels of diversification, and thus is not causally informing on the antecedents of diversification such as personal preferences or professional incentives. Rather, our analysis provides evidence that emphasizes the point that these factors matter and that they should be studied in more detail to robustly inform managers and policy makers on strategic actions regarding antecedents of diversification. Third, we study researchers engaging with a particular type of new knowledge in a particular area of science—computer science, electrical engineering and electronics. It is possible that at least some of the observed magnitudes reflect idiosyncratic aspects of this setting. Our hope is that our study provides enough compelling evidence to shine a light on the role of individual-level diversification and to start a conversation that encourages scholars to gain access to more comprehensive or granular data to refine and extend our findings.

While we focus on the individual level, it is important to note that, ultimately, our goal is to inform organization-level decisions that lead to successful exploration. Knowledge diversification at the organization level has long been recognized as a requirement for successful exploration (Cohen & Levinthal, 1990; Katila & Ahuja, 2002). However, given the increase in knowledge accumulation, achieving knowledge diversification across individuals in an organization is becoming increasingly costly, and is therefore a strategic decision that needs to be carefully evaluated. In particular, increased knowledge accumulation has been found to lead to a trade-off between depth and breadth of knowledge at the individual level (Jones, 2009, 2010; Schilling & Green, 2011). Thus, capitalizing on the benefits of external knowledge combinations is increasingly a team effort rather than an individual endeavor (Agrawal et al., 2016; Jones, 2009; Wuchty, Jones, & Uzzi, 2007), a fact that suggests increased complexity in organization-level hiring decisions.

Our study contributes to the literature examining how organizations can best use their limited resources to integrate new external knowledge in impactful and productive ways, which is often referred to as absorptive capacity (Cohen & Levinthal, 1989, 1990). While there are many studies on the topic (e.g., Eggers & Kaplan, 2009; Escribano, Fosfuri, & Tribó, 2009; Lane, Salk, & Lyles, 2001; Lenox & King, 2004; Pacheco-de-Almeida & Zemsky, 2007), most focus on firm-level characteristics, despite Cohen and Levinthal's (1990) argument that the concept of absorptive capacity occurs not only at the organization level but also at the individual level. After all, individuals are at the root of the knowledge creation process, and it has been posited that diverse knowledge in an individual allows for learning and problem solving that leads to innovation (Simon, 1985). Furthermore, our results provide a deeper understanding of what individual characteristics allow for more successful exploration through distant search, the benefits of which are long-term and lasting (March, 1991) and allow organizations to succeed during rapid technological shifts (Christensen, 1992; Christensen, Suárez, & Utterback, 1998; Cohen & Levinthal, 1994; Tushman & Anderson, 1986).

More broadly, the study offers insights into the career paths of researchers and scientists. Although institutional norms in both firms and research organizations frequently demonstrate preferences for specialization, our results show that individuals with high levels of knowledge diversity play an important role in pushing the knowledge frontier forward in critical ways. Furthermore, this role might grow in importance with increased knowledge accumulation and divisions into even narrower knowledge areas. In aggregate, our study contributes to calls for more individual-level perspectives to better understand the microfoundations of strategy (Felin & Foss, 2005; Foss, 2011; Gavetti, 2005; Teece, 2007) by drawing attention to the possibility that rather than being a "jack of all trades and master of none," individuals with high levels of knowledge diversity might play an important role as a "jack of all trades and master of knowledge."

ACKNOWLEDGEMENTS

Authors listed alphabetically; both authors contributed equally. We thank Ajay Agrawal, Michael Bikard, Jeff Furman, Avi Goldfarb, Keyvan Vakili, and the journal's editorial team for substantive discussions of the analysis, and seminar participants at various conferences and seminars for constructive feedback. All errors are our own.

REFERENCES

Adner, R., & Levinthal, D. (2008). Doing versus seeing: Acts of exploitation and perceptions of exploration. *Strategic Entrepreneurship Journal*, *2*(1), 43–52.

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review*, *37*(3), 355–375.

Agrawal, A., Goldfarb, A., & Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics*, *8*(1), 100–128.

Agrawal, A., McHale, J., & Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, *46*, 853–867.

Ahuja, G., Lampert, C. M., & Tandon, V. (2008). Moving beyond Schumpeter: Management research on the determinants of technological innovation. *The Academy of Management Annals*, *2*, 1–98.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129.

Altman, E., Nagle, F., & Tushman, M. (2014). Innovating without information constraints: Organizations, communities, and innovation when information costs approach zero. In M. A. Hitt, C. Shalley, & J. Zhou (Eds.), *Oxford handbook of creativity, innovation, and entrepreneurship*. Oxford, England: Oxford University Press.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Azoulay, P., Stuart, T., & Wang, Y. (2013). Matthew: Effect or fable? *Management Science*, *60*(1), 92–109.

Barth, E., Davis, J. C., Freeman, R. B., & Wang, A. J. (2017). *The effects of scientists and engineers on productivity and earnings at the establishment where they work* (NBER working paper #23484). National Bureau of Economic Research, Cambridge, MA.

Bayus, B. L., & Agarwal, R. (2007). The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. *Management Science*, *53*, 1887–1902.

BBC News. (2010). *Kinect hacked days after release*. Retrieved from http://www.bbc.co.uk/news/technology-11742236

Bikard, M., Murray, F., & Gans, J. (2015). Exploring trade-offs in the organization of scientific work: Collaboration and scientific award. *Management Science*, *61*(7), 1473–1495.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, *62*(10), 2765–2783.

Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*, *57*(5), 843–863.

Buis, M. L. (2010). Stata tip 87: Interpretation of interactions in nonlinear models. *The Stata Journal*, *10*(2), 305–308.

Bush, V. (1945). *Science the endless frontier: A report to the President. United States*. Washington, DC: Office of Scientific Research and Development.

Chai, S. (2017). Near misses in the breakthrough discovery process. *Organization Science*, *28*(3), 379–595.

Chatterji, A. K., & Fabrizio, K. R. (2014). Using users: When does external knowledge enhance corporate product innovation? *Strategic Management Journal*, *35*(1), 1427–1445.

Christensen, C. M. (1992). Exploring the limits of the technology S-curve. Part I: Component technologies. *Production and Operations Management*, *1*(4), 334–357.

Christensen, C. M., Suárez, F. F., & Utterback, J. M. (1998). Strategies for survival in fast-changing industries. *Management Science*, *44*(12-part-2), S207–S220.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: The two faces of R&D. *The Economic Journal*, *99*(397), 569–596.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Cohen, W. M., & Levinthal, D. A. (1994). Fortune favors the prepared firm. *Management Science*, *40*(2), 227–251.

Conti, R., Gambardella, A., & Mariani, M. (2013). Learning to be Edison: Inventors, organizations, and breakthrough inventions. *Organization Science*, *25*(3), 833–849.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.

Dothan, A., & Lavie, D. (2016). Resource reconfiguration: Learning from performance feedback. In T. B. Folta, C. E. Helfat, & S. Karim (Eds.), *Resource redeployment and corporate strategy. Advances in Strategic Management* (Vol. 35, pp. 319–369). Bingley, England: Emerald Group Publishing Ltd..

Eggers, J. P. (2014). Competing technologies and industry evolution: The benefits of making mistakes in the flat panel display industry. *Strategic Management Journal*, *35*, 159–178.

Eggers, J. P., & Kaplan, S. (2009). Cognition and renewal: Comparing CEO and organizational effects on incumbent adaptation to technical change. *Organization Science*, *20*(2), 461–477.

Eggers, J. P., & Kaul, A. (2018). Motivation and ability? A behavioral perspective on the pursuit of radical invention in multi-technology incumbents. *Academy of Management Journal*, *61*(1), 67–93.

Escribano, A., Fosfuri, A., & Tribó, J. A. (2009). Managing external knowledge flows: The moderating role of absorptive capacity. *Research Policy*, *38*(1), 96–105.

Fabrizio, K. R. (2009). Absorptive capacity and the search for innovation. *Research Policy*, *38*(2), 255–267.

Felin, T., & Foss, N. J. (2005). Strategic organization: A field in search of micro-foundations. *Strategic Organization*, *3*, 441–455.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, *47*(1), 117–132.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, *25*, 909–928.

Foss, N. J. (2011). Invited editorial: Why micro-foundations for resource-based theory are needed and what they may look like. *Journal of Management*, *37*(5), 1413–1428.

Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. Oxford, England: Holt, Rinehart & Winston.

Gavetti, G. (2005). Cognition and hierarchy: Rethinking the microfoundations of capabilities' development. *Organization Science*, *16*(6), 599–617.

Gavetti, G., & Levinthal, D. A. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly*, *45*, 113–137.

Green, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Econometric Letters*, *107*(2), 291–296.

Greve, H. R. (2003). A behavioral theory of R&D expenditures and innovations: Evidence from ship-building. *Academy of Management Journal*, *46*, 685–702.

Greve, H. R. (2007). Exploration and exploitation in product innovation. *Industrial and Corporate Change*, *16*, 945–975.

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The interplay between exploration and exploitation. *Academy of Management Journal*, *49*, 693–706.

Harris, M. S. (2018). *How to get tenure: Strategies for successfully navigating the process*. New York, NY: Routledge.

Henderson, R. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic equipment alignment industry. *The Rand Journal of Economics*, *24*, 248–270.

Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage Publications.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London, England: McGraw-Hill.

Hofstede, G., Neuijen, B., Ohayv, D. D., & Sanders, G. (1990). Measuring organizational cultures: A qualitative and quantitative study across twenty cases. *Administrative Science Quarterly*, *35*, 286–316.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, mpr013, *20*(1), 1–24.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, *21*(5), 1016–1033.

Jones, B. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *Review of Economic Studies*, *76*(1), 253–281.

Jones, B. (2010). As science evolves, how can science policy? *NBER Innovation Policy and the Economy*, *11*, 103–131.

Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435–1457.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *The Academy of Management Journal*, *45*(6), 1183–1194.

King, A. A., & Tucci, C. L. (2002). Incumbent entry into new market niches: The role of experience and managerial choice in the creation of dynamic capabilities. *Management Science*, *48*, 171–186.

Kolasinski, A. C., & Siegel, A. F. (2010). *On the economic meaning of interaction term coefficients in non-linear binary response regression models* (Working paper SSRN #1668750).

Lane, P. J., Salk, J. E., & Lyles, M. A. (2001). Absorptive capacity, learning, and performance in international joint ventures. *Strategic Management Journal*, *22*(12), 1139–1161.

Laursen, K., & Salter, A. (2006). Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms. *Strategic Management Journal*, *27*(2), 131–150.

Lenox, M., & King, A. (2004). Prospects for developing absorptive capacity through internal information provision. *Strategic Management Journal*, *25*(4), 331–345.

Lifshitz-Assaf, H. (2017). Dismantling knowledge boundaries at NASA: From problem solvers to solution seekers. *Administrative Science Quarterly* in press.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87.

Mokyr, J. (2002). *The gifts of Athena*. Princeton, NJ: Princeton University Press.

Murmann, J. P. (2003). *Knowledge and competitive advantage: The coevolution of firms, technology, and national institutions*. Cambridge, London: Cambridge University Press.

Nelson, R. R., & Winter, S. G. (1973). Toward an evolutionary theory of economic capabilities. *The American Economic Review*, *63*(2), 440–449.

Nelson, R. R., & Winter, S. G. (1977). In search of useful theory of innovation. *Research Policy*, *6*(1), 36–76.

Pacheco-de-Almeida, G., & Zemsky, P. (2007). The timing of resource development and sustainable competitive advantage. *Management Science*, *53*(4), 651–666.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., … Sobrero, M. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy*, *42*(2), 423–442.

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models. *Economics Letters*, *115*(1), 85–87.

Rothaermel, F. T. (2001). Incumbent's advantage through exploiting complementary assets via interfirm cooperation. *Strategic Management Journal*, *22*, 687–699.

Sauermann, H., & Stephan, P. (2013). Conflicting logics? A multidimensional view of industrial and academic science. *Organization Science*, *24*(3), 889–909.

Schilling, M. A., & Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, *40*(10), 1321–1331.

Schumpeter, J. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Shane, S. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science*, *11*(4), 448–469.

Simon, H. A. (1985). What we know about the creative process. *Frontiers in Creative and Innovative Management*, *4*, 3–22.

Stephan, P. (2012). *How economics shapes science*. Cambridge, MA: Harvard University Press.

Taylor, A., & Helfat, C. E. (2009). Organizational linkages for surviving technological change: Complementary assets, middle management, and ambidexterity. *Organization Science*, *20*, 718–739.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350.

Teodoridis, F. (2018). Understanding team knowledge production: The interrelated roles of technology and expertise. *Management Science*, *64*(8), 3469–3970.

Teodoridis, F., Bikard, M., & Vakili, K. (2018). Creativity at the knowledge frontier: The impact of specialization in fast- and slow-paced domains. *Administrative Science Quarterly* in press.

Terdiman, D. (2010). Bounty offered for open-source Kinect driver. *CNET News*. Retrieved from https://www.cnet.com/news/bounty-offered-for-open-source-kinect-driver/

Toh, P. K. (2014). Chicken, or egg, or both? The interrelationship between a firm's inventor specialization and scope of technologies. *Strategic Management Journal*, *35*(5), 723–738.

Tortoriello, M., McEvily, B., & Krackhardt, D. (2015). Being a catalyst of innovation: The role of knowledge diversity and network closure. *Organization Science*, *26*(2), 423–438.

Tripsas, M. (1997). Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*, *18*, 119–142.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, *31*, 439–465.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *Review of Economic Studies*, *79*, 838–861.

Weitzman, M. (1998). Recombinant growth. *Quarterly Journal of Economics*, *113*(2), 331–360.

Wu, B., Wan, Z., & Levinthal, D. A. (2014). Complementaryassets as pipes and prisms: Innovation incentives and trajectory choices. *Strategic Management Journal*, *35*, 1257–1278.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Nagle F, Teodoridis F. Jack of all trades and master of knowledge: The role of diversification in new distant knowledge integration. *Strat. Mgmt. J*. 2020;41:55–85. https://doi.org/10.1002/smj.3091