Research Policy xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Research Policy

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

Proximity effects on the dynamics and outcomes of scientific collaborations

Felichism W. Kabo[a,1], Natalie Cotton-Nessler[f], Yongha Hwang[c], Margaret C. Levenstein[a,b], Jason Owen-Smith[a,d,e,*]

[a] Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48106-1248, USA
[b] Stephen M. Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109-1234, USA
[c] Taubman College of Architecture and Urban Planning, University of Michigan, 2000 Bonisteel Boulevard, Ann Arbor, MI 48109-2069, USA
[d] Organizational Studies, University of Michigan, 500 S. State St. #3001, Ann Arbor, MI 48109-1382, USA
[e] Sociology, University of Michigan, 500 S. State St. #3001, Ann Arbor, MI 48109-1382, USA
[f] Bentley University, 175 Forest Street, Waltham, MA 02452, USA

ARTICLE INFO

Article history:
Received 24 May 2013
Received in revised form 7 April 2014
Accepted 12 April 2014
Available online xxx

Keywords:
Collaboration formation
Collaboration success
Path overlap
Functional proximity
Social networks
Spatial networks

ABSTRACT

This paper uses path overlap, an innovative measure of functional proximity, to examine how physical space shaped the formation and success of scientific collaborations among the occupants of two academic research buildings. We use research administration data on human subject protection, animal use management, and grant funding applications to construct new measures of collaboration formation and success. The "functional zones" investigators occupy in their buildings are defined by the shortest walking paths among assigned laboratory and office spaces, and the nearest elevators, stairs, and restrooms. When two investigators traverse paths with greater overlap, both their propensity to form new collaborations and to win grant funding for their joint work increase. This effect is robust across two very differently configured buildings. Implications for scientific collaboration and the design and allocation of research space are considered.

1. Introduction

Social scientific studies of scientific work and technological innovation evince a complex relationship with physical space. Some work locates productive innovation efforts in particular locations and organizations (Brown and Duguid, 1991; Hargadon and Douglas, 2001; Lave and Wenger, 1991; Weiser et al., 1999), emphasizing the importance of face-to-face interactions for collaboration formation and the coordination benefits of working jointly across small distances (Allen, 1977). Others make widely dispersed invisible colleges and inter-organizational networks the locus of innovation (Crane, 1972; Powell et al., 1996) and even argue that new information technologies have resulted in the "death of distance" (Cairncross, 2001). Cyber-infrastructure for distant collaboration is increasingly taken to be key to revolutionary work in science and engineering (Atkins et al., 2003) and virtual "collaboratories" are touted as new organizational forms for research and development that need take no account of physical separation (Finholt and Olson, 1997). The report of space's demise may be premature, however, at least with regard to work in environments, such as fundamental science, characterized by a high degree of tacit knowledge (Collins, 1992; Polanyi, 1967). Tacit knowledge is hard to transmit merely through the written word and is best exchanged face-to-face through a range of interactions between individuals (Collins, 1974). Dyadic analysis of scientific collaboration is ripe for rich analysis of knowledge exchange, which can in turn provide new insights into higher order organizational and network phenomena (Mizruchi and Marquis, 2006).

There appear to be few substitutes for dyadic face-to-face interactions in knowledge intensive work. As physical distance increases, the likelihood of collaboration decreases (Olson and Olson, 2000; Olson et al., 2002). Conceptualizing space in terms of the concrete physical layout of workplaces and showing how those layouts shape the dynamics and outcomes of collaborative work takes an essential step toward systematically integrating space into explanations of collaboration, network tie formation, and innovation. We especially emphasize the essential role that individual

* Corresponding author. Tel.: +1 734 936 0700; fax: +1 734 764 6554.
E-mail addresses: fkabo@umich.edu (F.W. Kabo), ncottonnessler@bentley.edu (N. Cotton-Nessler), yonghah@umich.edu (Y. Hwang), maggiel@umich.edu (M.C. Levenstein), jdos@umich.edu (J. Owen-Smith).
[1] Tel.: +1 734 936 0479; fax: +1 734 936 0548.

http://dx.doi.org/10.1016/j.respol.2014.04.007
0048-7333/

movement patterns play in determining the relationship between physical layouts, collaboration dynamics, and the interpersonal collaborations and outcomes of their occupants. In other words, we propose that proximity influences the formation and success of dyadic collaborations only to the extent that the occupants of a space actually use it (by, for instance, walking to and fro) in ways that bring them into face-to-face contact.

Few systematic efforts to theorize the complex relationship between the built environment, which creates patterns of proximity and distance, and the dynamics or outcomes of scientific collaborations have been undertaken. This paper takes steps in that direction by mobilizing a new measure of a classic spatial concept, functional proximity (Festinger et al., 1950), to test hypotheses about how space shapes the formation and early success of scientific collaboration.

We measure the extent to which pairs of investigators share overlapping areas as a means to characterize how individual patterns of movement through research buildings bring investigators into contact with each other (Kabo et al., 2013).[2] We use this measure to explain the formation and subsequent success of new collaborations among researchers engaged in scientific research in two buildings on the campus of a major public research university in the United States. Early stage collaborations are measured using detailed administrative data to identify instances where investigators jointly apply for human subjects protection, animal use management, or an external grant. We measure the initial success of these collaborations by considering whether or not joint grant applications were funded. We find that researchers whose paths overlap more are more likely to collaborate and to collaborate successfully than those who have fewer opportunities for casual contact. Our findings suggest the utility of fine-grained measures of functional proximity for understanding knowledge-intensive collaborative work. Insights into the formation and effects of scientific relationships have broad implications for social networks generally. More specifically, our work makes important contributions to the discourse on physical space as a mechanism for the formation, maintenance, and dissolution of collaborations (Rivera et al., 2010).

2. Space and innovation in organizations

Despite staggering advances in communication technologies, much innovative work is made possible by the face-to-face contacts that take place in physical space. The basic function of buildings, to shelter inhabitants, is relatively well understood. The more social dimensions and functions of buildings have received less systematic attention. Buildings "operate socially in two ways: they constitute the social organisation of everyday life as the spatial configurations of space in which we live and move, and represent social organisation as physical configurations of forms and elements that we see" (Hillier, 2007). In other words, the built environment and particularly the physical layout of workplaces both reflect assumptions about the organization of work and shape possibilities for action and interaction.

Earlier studies of the link between space and work processes focused on the effects that physical distance exerts on processes such as communication. Allen (1977) famously demonstrated that the probability of communication between pairs of engineers dropped precipitously at the 30 m mark. Olson et al. (2002) showed that radical co-location doubled the productivity of software engineers. More recent studies have built on the insight that measures of distance alone miss important aspects of spatial layouts. In particular, this line of research argues that physical relations among multiple spaces structure social interactions and workplace processes. Our work builds on this 'spatial network' approach to characterizing physical space using relational, not just distance, measures.

The most advanced body of theory and methods for studying the relational aspects of space is called 'space syntax.' Space syntax techniques, first developed by Hillier and Hanson (1984), explicitly characterize built spaces as networks by conceptualizing, rooms, passageways, and public spaces as nodes linked by the possibility of direct access or visibility. Viewed in this way, the network topology of a building, campus, or even city makes possible much more nuanced analyses than simple studies of linear distance. This 'functional' approach to proximity (Festinger et al., 1950) is particularly well suited to identifying the conditions under which ease of interaction and the possibility for unexpected, 'passive' contact between co-workers seed, improve, or sustain collaboration.

For example, it may be less important to know that two investigators' offices are separated by a distance of 90 ft than it is note that one has to make three 90° turns to get from one office to another, or that one must pass through shared spaces such as cubicles or private spaces such as the office of an assistant. Each of these facts will render the experience of 90 ft of distance more complex. The first makes it impossible for office occupants to make or maintain visual contact with one another. The second raises the possibility of interruptions and unwanted encounters as one passes through shared space (Backhouse and Drew, 1992). The third requires an intermediate contact with an individual whose physical and social position enable them to control access to an interlocutor (Mechanic, 1962). In other words, the relationships among spaces, not just the distance between them, influence workplace interactions. We seek to measure and understand how the configurational properties of space shape research collaborations and their success.

Nuanced spatial analysis is particularly important in research-intensive workplaces where serendipitous encounters and easy face-to-face interactions can spark insights and lower the coordination costs associated with collaborative research. The spaces created by buildings designed to support scientific research structure interactions, informal knowledge flows, and with them the likelihood and effectiveness of collaborations. At the same time, these spaces physically instantiate and elaborate social and organizational assumptions about the purposes and processes of research (Gieryn, 2002; Hillier, 1999). Consider one high profile research building on the campus of Stanford University.

The James H. Clark Center for Biomedical Engineering and Sciences opened in 2003 and received rave reviews for its radical, completely open floor plan. Research spaces in the building are both broadly observable (most walls are glass) and completely fluid. All the mechanical necessities of both wet and dry research work are set on wheels, so, in principle, researchers can set up shop anywhere in the building. In a recent article evaluating the building's status as a "truth spot," (Gieryn, 2008) calls the center itself an experiment, asking "can breakthroughs and innovations be nurtured most effectively in a building where the spatial locations of people and equipment can be rearranged at any time, in response to rapidly changing patterns of transdisciplinary interaction and collaboration?" He goes on to note that "Nothing much is secret in the Clark Center, as there are no doors to close off your lab space from other scientists, and even the hoi polloi can watch the action from. . .publicly accessible balconies. Scientists spontaneously bump into each other at lunch in the 256-seat NeXus full service restaurant . . ."

[2] The current paper is one part of a larger program of research. The measure we use here was first detailed in Kabo et al. (2013), which introduces the path overlap measure, describes methods for its construction and conducts validity tests of the new measure. That paper does not address the substantive or theoretical issues involving collaboration formation and early success that are our focus here. Instead, it demonstrates that path overlap is a significant and more robust predictor of overall collaboration rates than more traditional spatial measures such as physical walking distance or topological or turn distance.

With its emphasis on both the planned and the unscripted use of workspace in support of collaborative discovery, Gieryn's description highlights two features of space essential to understanding how building design shapes collaboration and innovation. First, there are 'no secrets' at the Clark Center. The building's very openness and the lack of individual privacy make it relatively easy for investigators to see what colleagues, collaborators, and lab members are doing. Awareness facilitated by openness, in turn, reduces problems of coordination and control associated with uncertain, interdisciplinary research work (Owen-Smith, 2001).

Second, spontaneous face-to-face encounters are 'programmed' into the building's design. This latter feature was also key in the proposed design for the super conducting super collider (SSC) facility that (Galison, 1997) called a 'trading zone' where interactions among scientists and engineers would be expected to facilitate innovation in high-energy physics. We believe such unscripted, face-to-face interactions to be central to the development of new collaborations. Likewise we argue that the coordination benefits that accompany greater ease of access to and awareness of the work of collaborators (Appel-Meulenbroek, 2009) will contribute to success.

Most scientific and technical workplaces fall short of the flexibility in the Clark Center and planned for parts of the SSC. Fully open and flexible floor plans are rare; most scientific buildings are internally differentiated by walls, stairs, doors, passageways, and other physical features that partition space in ways that make some locations more accessible or visible than others. This differentiation means that individuals in different labs or offices have unequal likelihood of bumping into any particular colleague in the course of their daily activity. Similarly, differently positioned scientists will face greater or lesser costs as they seek to direct, monitor, or troubleshoot the work of collaborators or lab members. These two features of the built environment – interactivity and awareness – influence innovation by shaping interactions, informal knowledge transfer, and the coordination of research collaborative work.

2.1. Hypotheses

Intra-organizational networks are forged and maintained in physical space, but social science research is just beginning to examine the intertwined social and spatial dimensions of organizations. The potential for a transformative sociospatial research program that was apparent in the seminal works of Festinger et al. (1950) is still largely untapped. Those scholars noted that space structures social interactions through two main mechanisms. The first mechanism is *physical distance*, which is a good proxy for the costs of dyadic interaction. The second mechanism is the relational attributes of spatial layouts, what they termed *functional distance*.

Functional distance captures the extent to which actors in a specific spatial environment are expected to interact using a relational conception of space. Hillier and Hanson (1984) made important advances in describing space relationally with the introduction of space syntax techniques. However, even that approach can fall short in analyses where the spatial level of analysis is a building or complex of buildings. Space syntax assumes that the relationships between spaces determine interpersonal contact and awareness, without taking into account either varying functional uses of space or varying individual patterns of work. Moreover, the proximity measures generated using space syntax emphasize point-to-point pathways connecting spaces and thus emphasize physical distance rather than functional distance. To develop a truer measure of functional proximity, we must incorporate a sense of how human behavior interacts with spatial layout to produce proximity. Functional proximity must reflect how people occupy and use their workspaces. Patterns of workplace mobility vary from person to person. While everyone enters and exits buildings, uses restrooms and traverses different kinds of workspace in the course of most days, the salience or importance of relevant spaces may vary across individuals or with other factors such as organizational culture and hierarchy.

Our work uses the relational insights of the space syntax tradition to develop a new measure of functional proximity that captures the extent to which individuals inhabit overlapping areas in buildings. People navigate among multiple spaces in the course of their daily work. The extent to which their routine walking paths overlap influences the likelihood of unscripted encounters. Kabo et al. (2013) introduces path overlap as a measure of functional distance and demonstrates the validity of path overlap relative to other measures (linear distance, turn distance). We build on this conception of functional proximity to make two arguments: (1) increasing path overlap between unconnected scientists increases the likelihood of interactions and knowledge sharing and therefore increases the likelihood of new dyadic collaborations and (2) increasing overlap between already collaborating scientists increases effectiveness by easing coordination and monitoring challenges, increasing the likelihood of successful outcomes.

2.1.1. Proximity and new collaboration formation

Beyond well documented claims that social and spatial proximity fosters collaboration and other relationships (Agrawal et al., 2008; McPherson et al., 2001; Sailer and McCulloh, 2012), relatively little theory exists to predict precisely how features of buildings and aspects of the built environment shape face-to-face interactions and knowledge flows at work. Co-location is widely believed to have multiple benefits including enhancing communication flows and frequency (Allen, 1977; Penn et al., 1999), increasing the probability of chance encounters or interactions (Allen and Henn, 2007), and amplifying the quality and impact of collaborative outcomes (Lee et al., 2010). Previous research suggests that communication between individuals drops significantly if they are not in the same building, on the same floor, or if the distance between them is beyond a specific threshold such as 30 m (Allen, 1977; Kraut et al., 1988; Monge and Kirste, 1980).

Moreover, researchers who have paid explicit attention to the ways in which particular features of buildings affect knowledge-intensive work have demonstrated that different parts of a space (e.g. offices, hallways, labs, meeting rooms) are conducive to varying degrees and types of interaction. These studies document the importance of having offices spread through more open laboratory spaces and the key role of organizational distinctions such as research team membership and differences in rank for collaboration formation (Hillier and Penn, 1991; Serrato and Wineman, 1999; Wineman et al., 2009). In order to forge new ties with other people, an individual needs to be exposed to them or have the ability to access them. Exposure to others significantly predicts network tie formation (Currarini et al., 2009, 2010) and the functional proximities created by physical space are important factors in exposure.

Functional proximity is particularly important for explaining relationship formation because this relational notion of space can provide useful insights into how physical layouts foster unplanned, fleeting face-to-face encounters. These kinds of repeated 'nodding' interactions, which Festinger et al. (1950, p. 34) dubbed 'passive contacts,' were the building blocks of friendships between new neighbors in a college dorm for returning World War II veterans. Passive contacts occurred most often when individuals met at the intersection of 'required paths' as they entered or left their apartments. As Festinger et al. (1950, pp. 34–35) note:

> Passive contacts are determined by required paths followed in entering or leaving one's home for any purpose. For example, in going from one's door to the stairway one must pass certain

Please cite this article in press as: Kabo, F.W., et al., Proximity effects on the dynamics and outcomes of scientific collaborations. Res. Policy (2014), http://dx.doi.org/10.1016/j.respol.2014.04.007

apartments; in walking to the butcher shop one must go by certain houses. These specific required paths are determined by the physical structure of the area.

We argue that the formation of new collaborations between neighbors in research buildings will depend on the specific pattern of required paths created by the building's layout and the habitual patterns by which people navigate them. In this sense, we treat collaborations like other types of relationships, that is, as connections that often emerge from unlooked for interactions. In interdisciplinary research settings where such interactions may span fields of expertise and substantive areas of interest the new collaborations may be particularly fruitful when they emerge from such encounters to combine unexpected approaches, questions, or techniques (Schumpeter, 1942).

In the terms we develop above, greater functional proximity between a given pair of investigators increases their passive contact and thus the likelihood they will share information, recognize common interests and eventually collaborate. Likewise linear distance between primary workspaces will decrease passive contacts regardless of investigators' functional proximity. Thus, we hypothesize:

H1. As the functional proximity between two individual researchers increases, the likelihood that they will form a new collaboration increases.

H2. As the physical distance between two individual researchers increases, the likelihood that they will form a new collaboration decreases.

2.1.2. Early success of collaborations

We expect physical distance to decrease and functional proximity to increase the likelihood of collaboration formation by altering rates of passive contact. Likewise, we expect the arrangement of physical space and investigators' relative positions in it to influence the success of new collaborations. That is, we argue that proximity and distance are also linked to collaborative outcomes.

Increased awareness stemming from higher rates of passive contact throughout the workday makes it possible for new collaborators to jointly identify and react to technical or conceptual issues that might arise in the course of early stage research. Physical distance, in contrast, makes it more difficult for collaborators to seek one another out when such issues arise. This suggests that increases in physical distance will decrease the likelihood of success while increases in functional proximity will increase that likelihood, but the logic behind each is slightly different.

While the benefits of physical proximity for knowledge intensive work are well known (c.f. Olson et al., 2002), the logic behind more functional conceptions of distance requires further elaboration. Our prediction that functional proximity increases the likelihood of early success for new collaborations rests on several related ideas. First, we note that collaborations in laboratory-based sciences represent significant coordination challenges because knowledge bases can differ across disciplines and approaches (Knorr-Cetina, 1999). The collaborations between academic scientists that we observe are, in reality, the joint efforts of investigators' labs. Coordinating the work of two research groups requires significant organizational savvy as both resources and attention must be apportioned to meet uncertain long and short term goals effectively (Owen-Smith, 2001). In such fast moving environments the ability to 'touch base' informally throughout the workday and investigators' awareness of the state of their partners' efforts will increase the likelihood of success on a technically challenging research frontier. Attracting and keeping the attention and effort of colleagues in a timely fashion may be as important to successful discovery as the original conception of an idea in these sorts of environments (Ocasio, 1997).

Part of the uncertainty associated with early stage collaborations comes from the highly tacit character of scientific research work. Whether people are troubleshooting and interpreting microscopic images (Lynch, 1985), attempting to replicate cutting edge experiments (Collins, 1992) or jointly debugging software (Olson et al., 2002), being able to intervene in, demonstrate, or reference hard to articulate aspects of a technical task often requires face-to-face interaction (Bechky, 2003). Moreover, problems are likely to occur at unpredictable times in the course of a workday or week. The passive contacts that were a wellspring for collaboration formation become a mechanism to troubleshoot, coordinate, and share information between researchers on an informal but routine basis.

In other words, once a new collaboration is formed, its early success depends on coordinating the work of many researchers as they work (often under deadlines) to construct new techniques, concepts and protocols. The success, for example, of a National Institutes of Health proposal depends in part on the quality of the initial data used to prove the proposed concept and approach. If greater functional proximity makes it easier for research investigators to shift from the conceptual stage of collaboration to the actual work of generating such data, then collaborations between proximate investigators will be more successful in their early stages. If the technical challenges faced by new, organically emerging, and often interdisciplinary collaborations involve the need to share highly tacit knowledge through difficult to predict face-to-face interactions at work, then greater functional proximity will also increase the speed and quality of new research efforts contributing to early stage success. For all these reasons we argue:

H3. As the functional proximity between two investigators increases, the likelihood of collaborative success increases.

H4. As the physical distance between two investigators increases, the likelihood of collaborative success decreases.

3. Data, methods, and variables

We build on prior work (Kabo et al., 2013) to test new hypotheses using extensive administrative data on collaborations and grants for similar sets of investigators in two relatively new buildings dedicated to interdisciplinary translational research on the campus of a large, public university in the United States. Our analyses for investigators in Building 1 (BLD1) and Building 2 (BLD2) span the period from 2006 to 2010. BLD1 was first occupied in 2006 while BLD2 was opened in 1997. We thus compare results for collaboration formation and success in a building that had long been occupied and one that was newly constructed and staffed. Because the occupants of each building change from year to year, we follow the population of researchers who were present in each building in 2006, at the time of BLD1's opening.[3] This includes 180 people in BLD1, and 128 people in BLD2.[4] We use human resource, spatial

[3] There are two potential sources of new collaborations especially in BLD1 which was first occupied in the middle of our study time period. First, people new to the Medical School could move into the building; any collaboration they form with others will be a new one. Second, people who were previously at the Medical School but were not co-located could choose to be proximate to one another. Such a choice might be elected in order to enhance the likelihood that they collaborate. The latter represents a key source of endogeneity that we take steps to mitigate.

[4] Office locations for the investigators we studied remained stable in 84% of investigator-years we observed. Just 4% of investigators gave up one office in order to move to another, while another 7% either added or lost office space without appearing to make an unambiguous move. Finally, 5% of investigator years were unable to be categorized due to missing data and other issues. In both these buildings, investigator locations are very stable with fewer than 11% of observations involving even the possibility of a change in office location.

Please cite this article in press as: Kabo, F.W., et al., Proximity effects on the dynamics and outcomes of scientific collaborations. Res. Policy (2014), http://dx.doi.org/10.1016/j.respol.2014.04.007

allocation and utilization, publication, and research administration data to capture early stage collaborations and develop nuanced measures of functional proximities among researchers.

3.1. Generation of spatial networks

Our ability to accurately locate individuals in space as well as map relations between their spaces required the generation of a spatial network for each of the buildings. Using electronic floor plans in the form of ArcGIS shapefiles, we generated spatial networks for BLD1 and BLD2. First, we broke down the building layout into smaller spatial elements. The connector spaces or hallways were decomposed into smaller path-contingent units which included demarcations of the connector spaces adjoining labs and offices as thresholds (Kabo et al., 2013). Primary spaces such as labs and offices, and public or circulation spaces such as break rooms, restrooms, elevators, and stairways were treated as discrete elements. Second, we constructed a spatial network by connecting the centroids of all spaces contingent on physical accessibility between adjacent or contiguous spaces. Finally, the inter-floor spatial networks in the building are connected through stairs and elevators. For each calendar year, we mapped each individual's position on the spatial network in each building using administrative data on workspace allocation (i.e. office and lab assignments).

We identified an individual's functional zone as the area bounded by his or her lab, office, nearest relevant restroom and closest elevator and stairs.[5,6] We take the shortest walking paths among these points to represent the 'required paths' an individual investigator will commonly traverse in the course of daily work. For every given pair of investigators in a building we used these functional zones to calculate a new measure of functional proximity that we call *path overlap*. Fig. 1 offers a graphic illustration of this concept. Taking two hypothetical researchers who are co-located on a floor in BLD1, the heavy black and double gray lines trace the shortest paths that connect the five types of spaces in Person 1's and Person 2's functional zones, respectively. Note that an individual's paths lie entirely within the area of his or her functional zone. In this conception, overlap between a dyad's functional zones operationalizes functional proximity as path overlap. For the two investigators pictured here, path overlap is the aggregate instances of their two paths running alongside each other. In this example, the overlap results from a shared elevator. By inference, our two hypothetical investigators are more likely to bump into each other when entering or exiting the building, provided at least that they keep more or less the same working hours. As a result, chance encounters between the hypothetical investigators in Fig. 1 are more likely earlier in the morning and later in the afternoon or evening (Table 1).

3.2. Measures and models

3.2.1. Dependent variables

3.2.1.1. New collaboration. For any given year from 2006 to 2010, we consider a collaboration to be new if it is the first time we observe two individuals jointly submitting either a new human subjects protection (Institutional Review Board, or IRB) application, a new animal research protocol (University Committee on Use and Care of Animals, or UCUCA), or a new grant application. Our dependent variable *new collaboration* is an indicator equal to one if a pair of researchers we have not previously observed collaborating initiate one of these three applications. This measure represents the earliest indication of a new collaboration we identify using administrative data. To the best of our knowledge this is a measure that has never been used before in the published literature. We use our joint measure of collaboration to track baseline rates of new project formation across our two samples plus a control sample of other medical school investigators. We recognize that left censoring may mean that some of the collaborations we observe may not, in fact, be new.

The BLD1 and BLD2 populations are similar in terms of their status within the Medical School, demographic factors, and level of research output relative to other units at the Medical School (Fig. 2). Fig. 2 compares annual new collaborations at BLD1, BLD2, and among a matched sample of researchers drawn from the Medical School with the same distribution of rank, department, and degree. This similarity in other characteristics simplifies our examination of how differences in spatial arrangements shape collaboration patterns and outcomes in these interdisciplinary research facilities.

3.2.1.2. Grants awarded. We analyze grants funded by external sponsors as a proxy for the success of a collaborative dyad. We analyze both funded and unfunded grant applications for the years 2006–2010. We create a dummy variable, *grants awarded*, equal to one if a pair of investigators we observe filing an IRB, UCUCA, or grant application in the $t-1\ldots t-3$ window preceding year t is jointly awarded a grant in the year t.

3.2.1.3. Prior collaboration. We create a variable to capture any instances of dyadic collaboration for the $t-1\ldots t-3$ window preceding year t, regardless of whether or not the dyad had any previous IRBs, UCUCAs, or grant applications. Our time period of interest is from 2001 to 2010, which accounting for the $t-1\ldots t-3$ window means that the effective coverage period is from 2004 to 2010. The resultant dependent variable, *prior collaboration*, is an indicator equal to one if, for any of the years in the $t-1\ldots t-3$ window, a pair of researchers jointly initiates any of the following: a new IRB application, a new UCUCA protocol, or a new grant application.

3.2.2. Independent variables

3.2.2.1. Path overlap. We apply *path overlap* – calculated as the total length in feet of overlapping paths in the intersection of two individuals' functional zones – as our measure of functional proximity. Recall Fig. 1, which illustrates path overlap between two individuals in BLD1.[7]

[5] Including both stairs and elevators in the functional zone allows us to control for behaviors such as people entering a floor using the elevator but choosing to exit the floor using the stairs in lieu of waiting for the elevator.

[6] While we focus our analysis on the overlaps accruing to the functional zone formed by the aggregate of offices, labs, restrooms, elevators, and stairs, we tested the reliability and validity of our approach by running analyses of overlaps created using smaller subsets of these spaces. Even though our findings were qualitatively robust to these disparate specifications of zone overlap, it is noteworthy that the biggest difference was in the subsets that included the office spaces relative to those that did not. Removing office spaces from the overlap set made no difference in BLD1. However, in BLD2 the removal of offices made path overlap non-significant in one model – that is, while controlling for physical linear distance. This finding is readily explained by the fact that the office spaces on a typical BLD2 floor plan were clustered around the core of the building, while BLD1 offices were arranged linearly on one side of the building (alongside the main façade). Removal of the BLD2 offices (which are very central) was therefore more likely to result in a lowered likelihood of encounters between investigators who had labs on either end of the floor plan. Conversely, BLD1 offices did not have the same centrality and connections were still possible between other spaces even in the absence of office spaces. A more detailed discussion of this phenomenon is in the sensitivity analysis section.

[7] In unreported sensitivity analyses we generate multiple measures of overlap using different combinations of our five basic types of spaces – offices (O), labs (L), elevators (E), stairs (S), and restrooms (R). We tested the robustness of our analyses using four additional combinations of spaces: OL, OLE, OLER, and OLRS. Our results were robust suggesting that the differences between the overlaps were

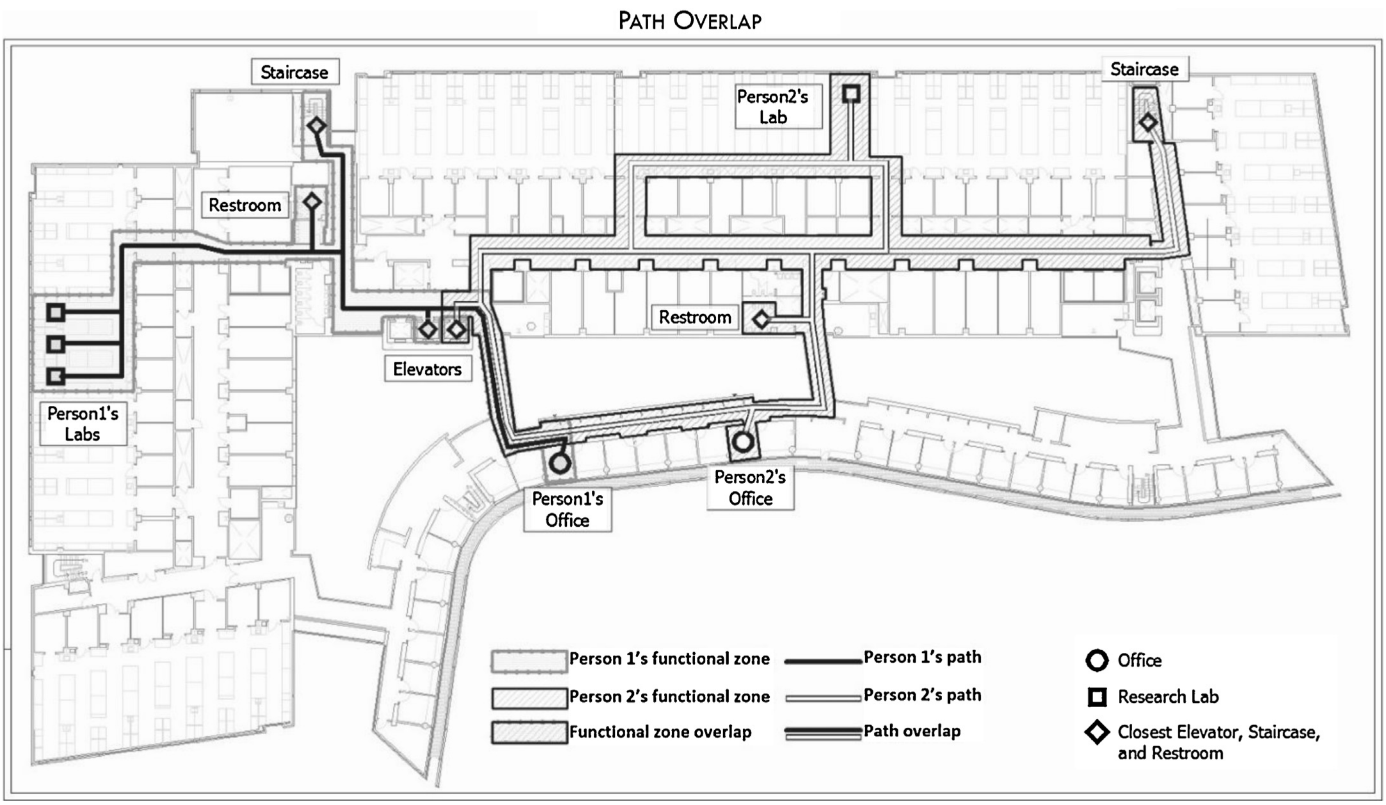

Fig. 1. An illustration of the path overlap measure. Note that the shared or public spaces that bound each person's functional zone in the example above are the elevators, stairs, and the restrooms.

Table 1
Descriptions of the variables in the study.

Variable	Type	Description
New collaboration dyad	Dummy	Equal to 1 if the collaboration is the first ever instance of an IRB application, UCUCA protocol, or grant application for the dyad members and zero otherwise
Grants awarded	Dummy	Equal to 1 if a grant application by the dyad members was successful that year, that is, obtained funding
Prior collaboration	Dummy	Equal to 1 if the dyad filed a new IRB application, UCUCA protocol, or grant application – even if they have worked together in the past – within the previous three years and zero otherwise ($t-1\ldots t-3$ cumulative measure)
Collaborativeness	Continuous	Sum of the number of collaborations (IRBs, animal protocols, grant proposals, and co-publications) each person in the dyad has with all other people in each of their respective samples (BLD1 or BLD2) including the dyad itself
Path overlap	Continuous	Length of the overlap in feet of the paths in the functional zones of the two people in the dyad
Physical distance	Continuous	Distance in feet between the labs of the two people in the dyad
Same building	Dummy	Equal to 1 if dyad members were in the same building for any part of the three prior years ($t-1\ldots t-3$ cumulative measure)
Same floor	Dummy	Equal to 1 if dyad members were on the same floor of the same building for any part of the three prior years ($t-1\ldots t-3$ cumulative measure)
Same department	Dummy	Equal to 1 if the two people in a dyad were in the same department that year (yearly measure) or any part of the three prior years ($t-1\ldots t-3$ cumulative measure)
Jobcode	Categorical	Equal to 0 when both people in dyad had academic or tenured/tenure-track positions Equal to 1 when exactly one person in the dyad had an academic position Equal to 2 when neither person in dyad had an academic position
Year	Dummy	Excluded year is 2006

3.2.2.2. *Physical distance.* This is the straight-line (or "as the crow flies") shortest distance in feet between the centroids of the individuals' labs. For individuals without any lab assignments, we calculate distances to others using their offices. Because of its simplicity, the straight-line conceptualization of physical distance has been used extensively by geographers and social scientists (Abramovsky and Simpson, 2011; Berry et al., 2010; Daraganova et al., 2012; Mok and Wellman, 2007; Moodysson and Jonsson, 2007; Pitts, 1978; Sailer and McCulloh, 2012). We compare it with our more nuanced path overlap measure in order to determine the advantages of functional proximity measures.

3.2.3. *Control variables*

3.2.3.1. *Collaborativeness.* The extent to which a potential dyad cements a collaboration relationship might merely be an artifact of self-selection; the potential dyad members might already be favorably disposed to initiating new ties because their research interests or personality types lend themselves to collaboration. To control for this effect, we create a *collaborativeness* count variable which is, for a given year, the sum of the IRBs, UCUCAs, grant applications, and co-publications in which each of the two individuals was involved

inconsequential. Therefore we chose to focus our analysis on the most inclusive definition of overlap, the one encompassing all of the five types of spaces (or OLERS).

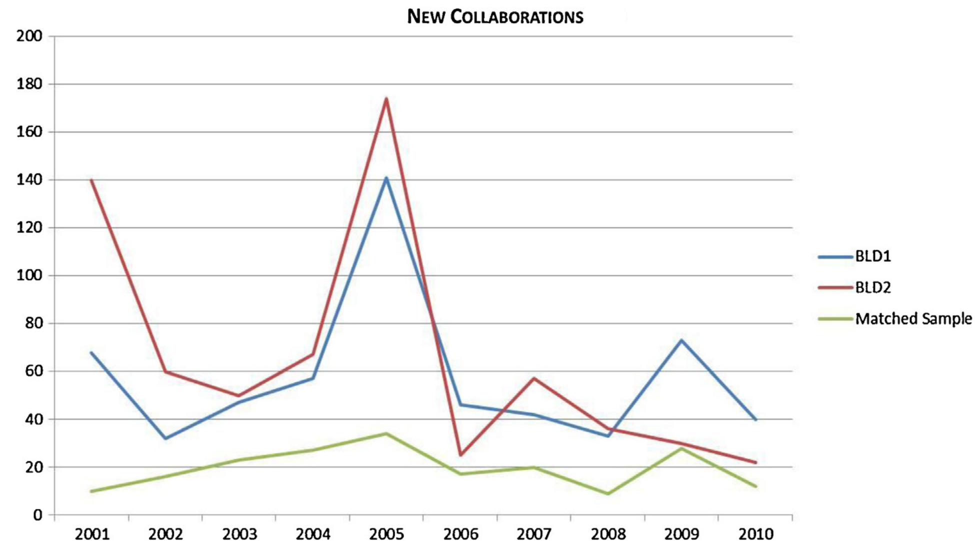

Fig. 2. The yearly count of new collaborations for researchers in the three populations: BLD1, BLD2, and a matched sample from the Medical School.

with *all* other people in their respective buildings, including the dyad itself.

3.2.3.2. Same building (t−1...t−3). Our coarsest measure of spatial co-location captures whether the dyad members are in the same building for any part of the $t-1\ldots t-3$ window preceding the year t in the period from 2004 to 2010. This is a dummy variable named *same building* equal to one if the pair is in the same building at least once and zero otherwise. We include this measure because over our period of study, occupants moved into and out of BLD1 and BLD2, from and to other locations on campus. Prior to 2006, all occupants of BLD1 resided in other buildings.

3.2.3.3. Same floor (t−1...t−3). This is a more refined measure of co-location that indicates whether the dyad members are on the same floor of the same building (equal to one if they were and zero if they were not) for any part or the entire $t-1\ldots t-3$ window preceding year t in the range 2004–2010. The variable is named *same floor*.

3.2.3.4. Same department (t & t−1...t−3). Research has established that individuals are more likely to collaborate with others in their group or department (Agneessens and Wittek, 2012; Kossinets and Watts, 2006; Wineman et al., 2009). We control for this similarity of research interests by creating a *same department* variable equal to one if the dyad members are employed by the same department in the year t and coded zero otherwise (yearly measure). For the cumulative 3-year measure, the variable is equal to one if the dyad members are employed by the same department at any time in the $t-1\ldots t-3$ window preceding year t and coded zero otherwise for the time period from 2004 to 2010.

3.2.3.5. Same jobcode (t−1...t−3). Based on the data from the human resources dataset, the researchers in BLD1 and BLD2 have positions that can be categorized as tenure/tenure-track (henceforth referred to as "academic"), clinical, and research. We create a dummy named *same jobcode* to capture whether a pair of researchers had the same job code using this three-category system for the $t-1\ldots t-3$ window preceding year t for the time period 2004–2010.

3.2.3.6. Jobcode. We are also interested in capturing whether or not differences in dyad formation and collaboration success are merely the artifacts of a social hierarchy conditioned by scientists having academic, clinical, or research positions. For our purposes, those three job codes effectively collapse into academic and non-academic (other) given that the BLD1 population has a negligible number of clinical-clinical dyads while the BLD2 sample has very low counts of research-research dyads. On this basis we create a *jobcode* categorical variable with three possible values: academic-academic, academic-other (where "other" is either research or clinical), and other-other for each year.

3.2.3.7. Year. To capture variations in external funding or other incentives to collaborate, we create a categorical *year* variable representing each year in the period 2004–2010.

3.3. Statistical analysis and model specification

We estimate dyadic rare events logit regression models to test our hypotheses about the effects of physical and functional distance on collaboration formation. To test predictions about the relationship between those proximities and early collaborative success (a joint grant being funded) we estimate Heckman probit selection models.

3.3.1. Proximity and new collaboration dyad formation

From Table 2 we see that the percentage of new collaboration dyads relative to all dyads for the years 2001–2010 ranges from 0.44 to 2.73%. Therefore, we treat the formation of new dyads as rare events and apply a statistical model suitable to analysis of outcomes that are very infrequent. With ordinary logistic regression models, the maximum likelihood estimator underestimates the probability that $Y=1$ and overestimates the probability that $Y=0$ (King and Zeng, 2001). One way to correct this is to weight the data to compensate for differences in the incidences of ones in the sample ($\bar{y}$) versus the population (τ). The weighted log-likelihood is then expressed as:

$$\ln L_w(\beta|y) = -\sum_{i=1}^{n} w_i \ln(1 + e^{(1-2Y_i)X_i\beta}) \qquad (1)$$

Table 2
Yearly summary of the number of newly formed collaboration dyads.

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
New dyads	162	62	73	90	306	76	105	72	108	66
Total dyads	5943	7609	10,083	13,528	16,015	17,332	13,792	11,581	10,576	9768
% new dyads	2.73%	0.81%	0.72%	0.67%	1.91%	0.44%	0.76%	0.62%	1.02%	0.68%

The weights are: $w_1 = \tau/\bar{y}$ and $w_0 = (1-\tau)/(1-\bar{y})$ and $w_i = w_1 Y_i + w_0(1-Y_i)$, and the weight w_i can then be inputted into the logit model using the ReLogit software program (King and Zeng, 2001; Tomz et al., 1999).

We estimate rare event logit regressions for scientists in each building in order to test hypotheses H1 and H2 for the years 2006–2010. These models cover this time period because the fine-grained spatial location data needed to generate the path overlap measure are only available starting in 2006. The general logit regression model is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 PATH + \beta_2 PHYS + \beta_3 COLL + \beta_4 DEP + \beta_5 JOB + \beta_6 YEAR + \hat{e} \quad (2)$$

where p is the probability of forming a new collaboration dyad or tie; *PATH* is the path overlap between dyad members; *PHYS* is the straight-line physical distance between dyad members; *COLL* is the total collaborativeness of the individuals in the dyad; *DEP* is the whether dyad members are in the same department; *JOB* is the job code of the dyad members; *YEAR* is the yearly fixed effects; and $\hat{e}$ is the error term.

3.3.2. Collaboration success

We model the likelihood of grants being awarded (indicating collaboration success) to investigators, contingent on observation of a new collaboration in order to limit potential selection bias. However, given the complexities of grant competitions, a grant may be awarded at any time up to a few years after a collaboration is initiated. Therefore, we use a t–1. . .t–3 window for new collaborations: We model the likelihood of a grant being awarded in year t, contingent on a new collaboration having been formed in years t–1, t–2, *or* t–3. Using the same sample of researchers as in the analysis above, we estimate two sets of Heckman probit selection models corresponding to BLD1 and BLD2 to test hypotheses H3 and H4. While the move into BLD1 occurred in 2006, our models include the three years immediately preceding t in order to evaluate both pre- and post-move spatial effects for the entire time range of our study (2001–2010). Therefore, our Heckman probit selection models effectively cover the years from 2004 to 2010.

Our Heckman selection model consists of two probit equations: a selection equation to model the likelihood that a given dyad forms a new collaboration in the t–1. . .t–3 window preceding year t, and an outcome equation to model the likelihood that a new grant is funded contingent on this new collaboration. The general probit equation models the probability that $Y = 1$ using the cumulative standard normal distribution function and is expressed as:

$$Pr\left(Y = \frac{1}{(X_1, X_2, X_3)}\right) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \hat{e}) \quad (3)$$

where Φ is the cumulative normal distribution function and X_1–X_3 is the predictor variables.

For the two binary variables in our model, formation of a new collaboration in the t–1. . .t–3 window (Y_S) and whether or not a grant is successfully funded (Y_O), the Heckman probit model (Borgoni and Billari, 2002; Heckman, 1976, 1979; Van de Ven and Van Praag, 1981) assumes that there is an underlying relationship between the two such that the outcome dependent variable, Y_O, is only observed for the conditions where $Y_S = 1$. Assuming a latent variable that captures the unobservable propensity of individuals to form new collaborations (Y_{SL}), then $Y_S = 1$ if $Y_{SL} \geq 0$ and $Y_S = 0$ if $Y_{SL} < 0$. Should a new collaboration be formed, then the dyad faces another binary outcome, Y_O, indicating whether their grant application is funded. To the outcome variable we can also add a latent variable, Y_{OL}, to capture the propensity for a dyad to have successful grant applications. Then, $Y_O = 1$ if $Y_{OL} \geq 0$ and $Y_O = 0$ if $Y_{OL} < 0$. The selection equation, which describes the probability of the formation of a new collaboration, is simplified as:

$$Pr\left(Y_S = \frac{1}{X_S}\right) = \eta X_S \quad (4)$$

where X_S is the set of predictors explaining the latent propensity for individuals to form new collaborations (for the t–1. . .t–3 window preceding the year t: co-location in the same building, co-location on the same floor, departmental affiliation, job code or type, and year).

The outcome equation is defined or observed only if $Y_S = 1$ and describes whether or not a grant application is successfully funded. It is simplified as:

$$Pr\left(Y_O = \frac{1}{X_O}\right) = \beta X_O \quad (5)$$

where X_O is the set of predictor and control variables explaining the latent propensity for dyads to submit successful grant applications (path overlap, physical distance, and collaborativeness).

4. Results

The summary statistics and pairwise correlations of the variables are given in Tables 3a–4b with breakdowns by building (BLD1, BLD2) and regression model (rare event logit, Heckman probit).[8] For the models predicting the formation of a new collaboration (Tables 3a and 3b), the correlations are either small or moderate. The low mean value of the binary dependent variable further confirms its low rates of incidence (c.f. Table 2) meaning that rare events models are superior to ordinary logit regressions. In the case of the collaboration success models (Tables 4a and 4b) the correlations are small or moderate with one of the exceptions being the correlation between being on the same floor and in the same department. The implication is that researchers are more likely to initiate collaborations with those are on the same floor or in the same department. As expected, individuals from the same departments are more likely to be on the same floor compared to those from different departments.

4.1. Findings: new collaboration formation

Tables 5a and 5b report results from rare event logit models of collaboration formation in BLD1 and BLD2, respectively.[9] For both

[8] Note that the OLSET overlap is used in constructing the correlation tables. Correlations using the other overlap definitions are identical with the exception of the "path overlap" variable.

[9] Recall that we generated multiple overlaps using the five spaces that define an individual's functional zone. Across all types of overlap the hypothesized relationships between path overlap, physical distance, and the formation of collaborations hold. Therefore, we present only the models for the least restrictive definition of investigator's functional zones using five different spaces.

Table 3a
Overall summary statistics and correlations of variables, rare events logit models BLD1 [OLSET].

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7
1. New collaboration dyad (Y)	0.006	0.078	0	1	1.000						
2. Path overlap	45.232	139.277	0	1044.5	0.094	1.000					
3. Physical distance	235.392	93.186	0	490.4	−0.059	−0.487	1.000				
4. Collaborativeness	35.378	45.433	0	466	0.027	0.029	0.007	1.000			
5. Same department	0.077	0.267	0	1	0.118	0.447	−0.312	0.020	1.000		
6. Jobcode	NA	NA	0	2	−0.033	−0.016	0.029	−0.241	−0.020	1.000	
7. Year	NA	NA	2006	2010	0.021	0.013	−0.015	0.337	0.023	−0.131	1.000

Table 3b
Overall summary statistics and correlations of variables, rare events logit models BLD2 [OLSET].

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7
1. New collaboration dyad (Y)	0.020	0.138	0	1	1.000						
2. Path overlap	71.412	149.588	0	793.3	0.061	1.000					
3. Physical distance	158.807	53.301	0	264.1	−0.047	−0.637	1.000				
4. Collaborativeness	76.363	85.686	0	651	0.084	0.041	0.001	1.000			
5. Same department	0.176	0.381	0	1	0.126	0.152	−0.165	0.243	1.000		
6. Jobcode	NA	NA	0	2	−0.103	−0.030	−0.051	−0.365	−0.051	1.000	
7. Year	NA	NA	2006	2010	0.026	0.006	−0.041	0.415	0.066	−0.128	1.000

buildings, the models control for dyad members' collaborations with others, departmental affiliation, differences in job type, and yearly differences in the base rates of dyadic collaborativeness.[10] The independent variables, path overlap and physical distance, capture functional proximity and physical proximity, respectively.[11]

For each building, goodness-of-fit testing using the Bayesian Information Criterion (BIC) revealed very strong support for the full – that is, with path overlap, physical distance, and controls – models (M4 and M8) over the models with controls only (M1 and M5), path overlap and control variables (M2 and M6), and physical distance and controls (M3 and M7). Therefore, the full models, M4 and M8, will henceforth be used to evaluate our hypotheses.

4.1.1. Path overlap

Overall, the results provide strong support for H1 and partial support for H2 (models M4 and M8). Path overlap has a significant effect on the formation of new collaborations in both BLD1 and BLD2. To better understand the impact of increasing or decreasing path overlap, we calculate the marginal effects of path overlap on the likelihood that new collaborations will form. In BLD1, a 100-foot increase in the extent to which a pair of researchers' walking paths overlap results in a 19.7% increase in the likelihood that they will form a new collaboration. The numbers for BLD2 are strikingly similar. A 100-foot increase in path overlap leads to a 19.8% increase in the likelihood of a new collaboration forming. These marginal effects of path overlap at BLD1 and BLD2 are statistically indistinguishable, suggesting robustness to building layout differences.

4.1.2. Physical distance

Increasing the physical distance between a pair of researchers by 100 ft reduces the likelihood of their forming a new collaboration dyad by 36.1% and 25.5% at BLD1 and BLD2, respectively. However, the effect of physical distance is significant at BLD1 but not at BLD2, providing only partial corroboration for H2. As there is higher correlation between path overlap and physical distance in BLD2 it is likely that the path overlap measure picks up other effects that are also associated with distance. There are undoubtedly building layout and topology factors that influence the effects of our two proximity measures; functional proximity as operationalized by path overlap appears more robust to these influences.

4.1.3. Control variables

The controls perform largely as expected with respect to their effects on the formation of new collaboration dyads. That is, dyads composed of individuals with high overall levels of collaborativeness are more likely to form new collaborations. Moreover, departmental affiliation is a strong indicator of the potential for two individuals to form a new collaboration. The results also suggest that job types capture differences in the level and role of individuals in research activity and therefore in research collaborations: in both BLD1 and BLD2, dyads composed of individuals who both had "academic" job types formed new collaborations at significantly higher rates than the other two types of dyads.

4.2. Findings: collaboration success

Tables 6a and 6b report Heckman probit sample selection models of early collaboration success contingent on the existence of a prior collaboration. The predictors for the selection equation are: same building, same floor, same department, same jobcode, jobcode, and year dummies.[12] The first four predictors cover the $t-1\ldots t-3$ window preceding year t while the last two are categorical variables. Note that the binary variable "same jobcode" simply indexes whether both members of a potential collaboration

[10] From Tables 3a and 3b we see that *collaborativeness* appears to be skewed. Consequently, we also ran relogit models where we logged *collaborativeness*. In these logged models, there was no change in the significance of the path overlap and physical distance predictors. Therefore, we present only the models using the unlogged *collaborativeness* variable.

[11] Many studies that use distance as a variable log it to account for the fact that its distribution tends to be skewed. In our case, the physical distances were either not skewed (BLD1) or were not skewed to the point that it made sense to use the lognormal distribution (BLD2). However, as an additional robustness check, we also considered models where the path overlaps and physical distances were logged alongside the *collaborativeness* control. In both BLD1 and BLD2, the significance of path overlap held. While physical distance became significant at BLD2, we decided to use the models with the unlogged variables as they were more stable in terms of the numbers of observations across the different overlap types.

[12] The selection equation in the Heckman probit model does not correct for the sparseness of the dependent variable as was done using the rare events logit models. To account for this, we also ran unconditioned rare events logits on collaboration success to test the possibility that bias introduced into the selection model by rare events or by the exclusion of relevant controls from our outcome model might alter our findings. Key results were robust to this sensitivity analysis.

Table 4a
Overall summary statistics and correlations of variables, Heckman probit selection models BLD1 [OLSET].

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11
1. Grants awarded (Y_O)	0.005	0.072	0	1	1.000										
2. Path overlap	45.232	139.277	0	1044.5	0.167	1.000									
3. Physical distance	235.392	93.186	0	490.4	−0.107	−0.487	1.000								
4. Collaborativeness	35.378	45.433	0	466	0.092	0.029	0.007	1.000							
5. Prior collaborations, t−1. . .t−3 (Y_S)	0.968	0.177	0	1	0.005	0.020	−0.022	−0.154	1.000						
6. Same building, t−1. . .t−3	0.862	0.345	0	1	0.011	0.046	−0.018	0.006	0.046	1.000					
7. Same floor, t−1. . .t−3	0.584	0.493	0	1	0.018	0.312	−0.131	−0.334	0.159	0.475	1.000				
8. Same department, t−1. . .t−3	0.535	0.499	0	1	0.015	0.182	−0.101	−0.370	0.179	0.394	0.842	1.000			
9. Same jobcode, t−1. . .t−3	0.851	0.356	0	1	−0.051	−0.043	0.040	−0.301	0.125	0.103	0.328	0.365	1.000		
10. Jobcode	NA	NA	0	2	−0.047	−0.016	0.029	−0.241	0.204	0.088	0.182	0.194	0.540	1.000	
11. Year	NA	NA	2001	2010	0.015	0.013	−0.015	0.450	−0.577	0.127	−0.461	−0.571	−0.287	−0.090	1.000

Table 4b
Overall summary statistics and correlations of variables, Heckman probit selection models BLD2 [OLSET].

Variable	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11
1. Grants awarded (Y_O)	0.013	0.113	0	1	1.000										
2. Path overlap	71.412	149.588	0	793.3	0.107	1.000									
3. Physical distance	158.807	53.301	0	264.1	−0.091	−0.637	1.000								
4. Collaborativeness	76.363	85.686	0	651	0.118	0.041	0.001	1.000							
5. Prior collaborations, t−1. . .t−3 (Y_S)	0.597	0.490	0	1	−0.002	0.017	−0.026	−0.199	1.000						
6. Same building, t−1. . .t−3	0.957	0.203	0	1	0.013	0.012	−0.056	−0.024	0.006	1.000					
7. Same floor, t−1. . .t−3	0.651	0.477	0	1	−0.003	0.428	−0.278	−0.337	0.184	0.289	1.000				
8. Same department, t−1. . .t−3	0.639	0.480	0	1	0.020	0.086	−0.093	−0.253	0.204	0.227	0.707	1.000			
9. Same jobcode, t−1. . .t−3	0.851	0.356	0	1	−0.091	−0.047	−0.023	−0.411	0.178	−0.036	0.318	0.377	1.000		
10. Jobcode	NA	NA	0	2	−0.107	−0.030	−0.051	−0.365	0.128	−0.067	0.164	0.203	0.585	1.000	
11. Year	NA	NA	2001	2010	−0.007	0.006	−0.041	0.518	−0.164	−0.053	−0.476	−0.478	−0.281	−0.060	1.000

Please cite this article in press as: Kabo, F.W., et al., Proximity effects on the dynamics and outcomes of scientific collaborations. Res. Policy (2014), http://dx.doi.org/10.1016/j.respol.2014.04.007

Table 5a
The effects of path overlap and physical distance on new dyad formation at BLD1 [OLSET].

Variables	M1 New dyad	M2 New dyad	M3 New dyad	M4 New dyad
Path overlap		0.00281***		0.00180***
		(0.000290)		(0.000363)
Physical distance, lab			−0.00713***	−0.00448***
			(0.000978)	(0.00103)
Collaborativeness	0.00474***	0.00565***	0.00561***	0.00589***
	(0.000882)	(0.000909)	(0.000898)	(0.000914)
Same department	2.093***	1.511***	1.540***	1.347***
	(0.146)	(0.170)	(0.150)	(0.165)
Jobcode Academic–Academic (reference)				
Jobcode Academic–Other	−0.675***	−0.600***	−0.662***	−0.615***
	(0.149)	(0.152)	(0.148)	(0.150)
Jobcode Other–Other	−1.002***	−1.053***	−1.058***	−1.070***
	(0.231)	(0.235)	(0.229)	(0.233)
Constant	−5.435***	−5.650***	−3.928***	−4.613***
	(0.183)	(0.187)	(0.232)	(0.259)
Observations	43,429	43,429	43,429	43,429

Note: Dummies for each year from 2006–2010 were included in the regression models. However, the coefficients are not reported as they were not very informative.
Robust standard errors in parentheses.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, +$p < 0.1$.

dyad share the same general type of job. In contrast, the categorical variable "jobcode" captures the different combinations of job types that appear in potential collaboration dyads. For the outcome equation, the independent variables are path overlap and physical distance. The outcome models control for the overall collaborativeness of the individuals in the dyad. For each building, BIC measures of fit for the probit outcome equations only suggest that the full models (M12-OUT and M16-OUT) are stronger than either the models with only the collaborativeness control (M9-OUT and M13-OUT), path overlap and the control (M10-OUT and M14-OUT), and physical distance and the control (M11-OUT and M15-OUT). Our discussions of the path overlap and physical distance variables will henceforth be on the basis of the full models.

4.2.1. Selection equation: prior collaborations, t−1...t−3 window

The selection equations for M12 and M16 model the likelihood that a given dyad formed a collaboration during the preceding three years. For both BLD1 and BLD2 scientists, the propensity to form collaborations is significantly and positively influenced by co-presence in the same building, co-location on the same floor, being in the same department, and having the same type of job. The categorical job type predictors behave similarly across buildings. Dyads whose individual members are both regular faculty ("academic" dyads) are no more likely to have prior collaborations than are pairs that mix academic faculty with research and/or clinical faculty. However, dyads whose members are both research or clinical faculty are more likely to have prior collaborations than "academic" dyads.

4.2.2. Outcome equation: grants awarded

4.2.2.1. Path overlap. We find positive effects of path overlap on the likelihood that new collaborations will result in successful grant proposals in both buildings. The columns labeled M12-OUT and M16-OUT in Tables 6a and 6b report our estimates of the likelihood that a given dyad receives an externally funded grant conditional on having established a collaboration (by filing a joint IRB application or animal research protocol or grant application) in the prior three

Table 5b
The effects of path overlap and physical distance on new dyad formation at BLD2 [OLSET].

Variables	M5 New dyad	M6 New dyad	M7 New dyad	M8 New dyad
Path overlap		0.00247***		0.00181**
		(0.000437)		(0.000679)
Physical distance, lab			−0.00684***	−0.00294
			(0.00166)	(0.00234)
Collaborativeness	0.00241**	0.00259**	0.00288***	0.00276***
	(0.000815)	(0.000805)	(0.000822)	(0.000825)
Same department	1.801***	1.748***	1.719***	1.723***
	(0.190)	(0.190)	(0.189)	(0.190)
Jobcode Academic–Academic (reference)				
Jobcode Academic–Other	−0.615***	−0.578**	−0.673***	−0.611**
	(0.185)	(0.190)	(0.188)	(0.191)
Jobcode Other–Other	−1.329***	−1.299***	−1.405***	−1.339***
	(0.351)	(0.357)	(0.354)	(0.357)
Constant	−4.762***	−5.040***	−3.676***	−4.494***
	(0.247)	(0.258)	(0.355)	(0.507)
Observations	9876	9876	9876	9876

Note: Dummies for each year from 2006 to 2010 were included in the regression models. However, the coefficients are not reported as they were not very informative.
Robust standard errors in parentheses.
****$p < 0.001$, **$p < 0.01$, *$p < 0.05$, +$p < 0.1$.

Table 6a
Effects of path overlap and physical distance on grants awarded at BLD1, conditional on new collaborations in previous three years [OLSET].

Variables	M9-OUT Grants awarded	M9-SEL Prior collab	M10-OUT Grants awarded	M10-SEL Prior collab	M11-OUT Grants awarded	M11-SEL Prior collab	M12-OUT Grants awarded	M12-SEL Prior collab
Path overlap			0.00192***				0.000989***	
			(9.12e−05)				(0.000123)	
Physical distance, lab					−0.00614***		−0.00401***	
					(0.000338)		(0.000404)	
Collaborativeness	0.00552***		0.00501***		0.00542***		0.00524***	
	(0.000209)		(0.000301)		(0.000304)		(0.000314)	
Same building, $t-1\ldots t-3$		0.297***		0.236**		0.236**		0.239**
		(0.0528)		(0.0802)		(0.0802)		(0.0801)
Same floor, $t-1\ldots t-3$		0.184***		0.130***		0.132***		0.130***
		(0.0280)		(0.0293)		(0.0294)		(0.0293)
Same department, $t-1\ldots t-3$		0.749***		0.712***		0.711***		0.708***
		(0.0411)		(0.0454)		(0.0452)		(0.0453)
Same jobcode, $t-1\ldots t-3$		0.181***		0.260***		0.261***		0.259***
		(0.0337)		(0.0426)		(0.0427)		(0.0426)
Jobcode Academic–Academic (reference)								
Jobcode Academic–Other		0.00963		−0.0739+		−0.0751+		−0.0719+
		(0.0327)		(0.0419)		(0.0419)		(0.0419)
Jobcode Other–Other		0.387***		0.337***		0.336***		0.340***
		(0.0429)		(0.0532)		(0.0533)		(0.0532)
Constant	−2.717***	1.511***	−3.094***	1.534***	−1.787***	1.534***	−2.288***	1.534***
	(0.0279)	(0.0320)	(0.0515)	(0.0346)	(0.0568)	(0.0346)	(0.0861)	(0.0346)
Observations	88,312	88,312	45,862	45,862	45,862	45,862	45,862	45,862

Note: Dummies for each year from 2004 to 2010 were included in the regression models. However, the coefficients are not reported as they were not very informative. Standard errors in parentheses.
****$p<0.001$, **$p<0.01$, *$p<0.05$, +$p<0.1$.

years. We find that path overlap is a significant and positive predictor of successful grants in both BLD1 and BLD2, thus supporting H3.

Interpretation of the coefficients in the probit models M12-OUT and M16-OUT is not as straightforward as it is for linear or logit regressions as the marginal effect varies with the level of the independent variables. We report the marginal effects of path overlap on the likelihood of a successful grant evaluated at the mean level of path overlap. Conditional on the existence of a prior collaboration, a 100-foot increase in the path overlap between two researchers yields a 33.4% and 22.8% increase in the likelihood that the pair will receive an external grant in BLD1 and BLD2, respectively.

4.2.2.2. Physical distance. The models also provide support for H4; physical distance is a significant predictor of successful grants in BLD1 and BLD2. However, the effect is much stronger at BLD1 ($p<.001$) than it is at BLD2 ($p=.021$). This finding reinforces the rare events logit models results that suggest that the physical distance measure is more sensitive to the building layout or topology than is path overlap.

Table 6b
Effects of path overlap and physical distance on grants awarded at BLD2, conditional on new collaborations in previous three years [OLSET].

Variables	M13-OUT Grants awarded	M13-SEL Prior collab	M14-OUT Grants awarded	M14-SEL Prior collab	M15-OUT Grants awarded	M15-SEL Prior collab	M16-OUT Grants awarded	M16-SEL Prior collab
Path overlap			0.00127***				0.000841**	
			(0.000187)				(0.000261)	
Physical distance, lab					−0.00393***		−0.00196*	
					(0.000627)		(0.000849)	
Collaborativeness	0.00355***		0.00336***		0.00344***		0.00340***	
	(0.000163)		(0.000267)		(0.000269)		(0.000271)	
Same building, $t-1\ldots t-3$		−0.218***		0.102		0.0956		0.100
		(0.0652)		(0.109)		(0.109)		(0.109)
Same floor, $t-1\ldots t-3$		0.296***		0.278***		0.287***		0.279***
		(0.0360)		(0.0426)		(0.0425)		(0.0427)
Same department, $t-1\ldots t-3$		0.538***		0.595***		0.590***		0.592***
		(0.0385)		(0.0445)		(0.0445)		(0.0445)
Same jobcode, $t-1\ldots t-3$		0.191**		0.312***		0.317***		0.314***
		(0.0706)		(0.0910)		(0.0913)		(0.0913)
Jobcode Academic–Academic (reference)								
Jobcode Academic–Other		0.0912		−0.0573		−0.0625		−0.0594
		(0.0702)		(0.0897)		(0.0899)		(0.0899)
Jobcode Other–Other		0.653***		0.457***		0.450***		0.456***
		(0.0867)		(0.107)		(0.107)		(0.107)
Constant	−2.492***	1.337***	−2.788***	0.984***	−2.085***	0.990***	−2.476***	0.987***
	(0.0354)	(0.0748)	(0.108)	(0.108)	(0.100)	(0.108)	(0.165)	(0.108)
Observations	33,302	33,302	12,251	12,251	12,251	12,251	12,251	12,251

Note: Dummies for each year from 2004 to 2010 were included in the regression models. However, the coefficients are not reported as they were not very informative. Standard errors in parentheses.
***$p<0.001$, **$p<0.01$, *$p<0.05$, +$p<0.1$.

5. Discussion

Linear measures of physical distance influence the likelihood a collaboration will form in only one of the buildings we study. In contrast, a new measure of functional proximity, path overlap, has a significant and positive effect in both buildings (M4 and M8). Both functional proximity and physical distance are associated with the early outcomes of collaborations in each building (M12-OUT and M16-OUT). Nonetheless, physical distance has a stronger effect at BLD1 than in BLD2. These differential effects of physical distance and path overlap may result from variation in the science or the scientists or from differences in the spatial layouts of the buildings themselves. More research is necessary, but we suspect that distance matters more in BLD1 both because the occupants are relatively new to the space in the years we observe and because of differences in the spatial topologies of the buildings. BLD1 has an internal atrium that separates labs and offices.[13] The single largest contiguous part of BLD1 (the northern wing) is 428′ long by 86′ wide, giving a length-to-width ratio close to 5. In contrast, BLD2, which has no internal atrium and has a compact central service core, is 223′ long by 117′ wide giving a ratio roughly equal to 2. Therefore, BLD1 is more linear in its topology and longer in terms of actual physical dimensions. Physical distance is sensitive to building layout effects of the kind that occur when two buildings having the same area and number of spaces array those spaces in more square or rectangular layouts (Kabo et al., 2013). These potential effects of building layout provide a plausible explanation for why physical distance exerts a stronger influence in BLD1. The robustness of path overlap to building layout differences makes it useful for comparisons of spatial effects on collaboration and innovation processes across buildings or spatial settings.

The effects of path overlap on the formation of new collaborations are significant and identical for BLD1 and BLD2. These effects are substantial and suggest that subtler dimensions of spatial proximity matter. Intriguingly, these more nuanced spatial effects vary somewhat across BLD1 and BLD2, suggesting that within building micro-level differences in proximity also influence the propensity to collaborate, and highlighting the need for further research.

5.1. Sensitivity analysis

The five types of spaces we use to define overlaps between individuals may vary in terms of the extent to which they are characterized by sedentary, task-related activities, and movement or the fact that some of these spaces are circulatory and are therefore occupied fleetingly as individuals make their way to another space or destination. For example, office spaces might fall on the sedentary end of the spectrum, elevators and stairs on the movement pole, with labs and restrooms likely falling somewhere in between. However, substantively, our measure of overlap captures movement to/from locations, and what a person is doing in a given location is thus irrelevant. In other words, two scientists are just as likely to bump into each other in the hallway between their respective labs as they are to bump into each other between their respective offices. We tested the robustness of path overlap by running models with different specifications (or that use different subsets of the five core types of spaces). Our results indicated that path overlap was robust to differences in functional zone specification in both BLD1 and BLD2. In the latter building, overlaps based on specifications that omitted office spaces were non-significant but only in models that also controlled for physical distance. That is, path overlaps were significant when considered independently. We believe that this is an artifact of the differences in spatial configuration across BLD1 and BLD2, and particularly the location of office spaces (Figs. 3a–4b). These results confirm that the overlap measure captures hallway activity. To the extent that we exclude important destinations (such as offices), we become less accurate in estimating hallway activity, but the remaining hallway encounters (without offices) are still significant. The main exception is that BLD2 is unique in that to the extent the offices are so central that a very large percentage of hallway activity is lost with their removal. So overall, the results do not indicate that offices play a different substantive role in the construction of overlap, but they do suggest the need for further work to understand the relationship between floor plan design and hallway interactions.

From Fig. 3a we see that considering the links from offices to labs in BLD1, there are overlaps between the B–C and A–B dyads. When the offices are removed in Fig. 3b, potential overlaps still exist for the two dyads on the basis of restrooms and elevators. The situation is rather different in BLD2. In Fig. 4a, we see an overlap in the B–C dyad on the basis of the office-lab path, and the A-B and A-C dyads on the basis of the paths to the elevators and (potentially, contingent on gender similarity) the restrooms. Removing the offices in BLD2 (Fig. 4b) has a larger effect on overlap possibilities, reducing them to possible overlaps for the A–B and A–C dyads based on paths to the restroom (and contingent on gender similarity), and a possible overlap for the B–C dyad based on the path to the closest stairs.

We found empirical support for the logical deductions above when we ran models where the overlap set of spaces did not include offices. While we do not report these results, it is noteworthy that they are highly suggestive of a mechanism by which the configuration or layout of a building structures dyadic encounters and interactions through its ordering of the topological relationships between types of spaces. In sum, our measure of path overlap captures the extent to which two people are likely to encounter each other in hallways, irrespective of whether they are moving between offices, labs, restrooms, or other spaces. While each of these kinds of spaces has different levels of mobility associated with them, they are all relevant and useful for ascertaining the likelihood of hallway encounters as people move from one space to the next.

5.2. Limitations and future directions

The findings suggest that functional proximity plays an important role in the formation and success of life science collaborations. Nevertheless, some qualifiers to this analysis are necessary. Primarily, and most importantly, this analysis cannot entirely rule out the possibility that the functional proximity between investigators reflects their choices. Scientists who intend to begin collaborating may select offices or labs that are near one another, thus increasing the degree of path overlap. Likewise, investigators who are pursuing potentially complementary research agendas may be positioned nearby one another through organizational processes for space allocation that are consciously designed to foster collaboration. It is for these reasons that we focused on new collaborations and used the two-stage Heckman correction for potential sources of endogeneity in the formation of collaborations. Several additional features of our analysis suggest that neither office selection nor organizational policies drive our findings.

First, the effects we find for path overlap are remarkably stable across two very different buildings that were constructed more than ten years apart. Space allocation processes at this university have varied significantly across time with changes in academic administrations and shifts in the availability and location of cutting edge research space on campus. Second, investigators in BLD1 in 2006 were the first to occupy new construction and thus may have

[13] As noted earlier, we needed to control for the endogeneity presented when people choose to be co-located with the intent that this proximity will boost the likelihood that they will collaborate.

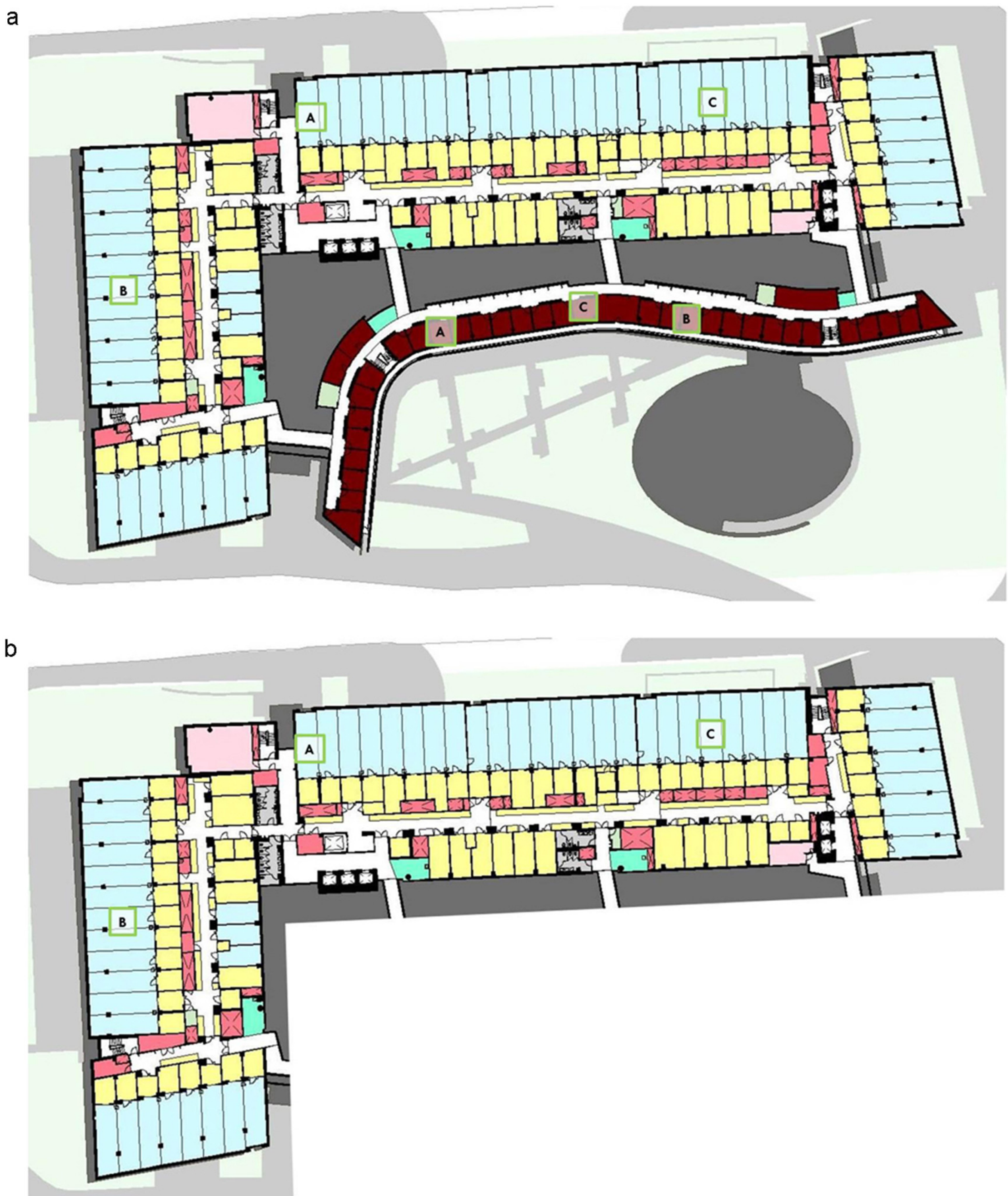

Fig. 3. (a) Offices (in dark red) in BLD1 are arranged linearly along the main façade of the building. The connections between offices and other spaces are entirely via the atrium or skywalks spanning the atrium for the upper floors. (b) Removing offices in BLD1 reduces the number of alternative paths between pairs of other spaces, but does not disconnect them. The four internal hallways – two running roughly east-west and two running north-south – make it possible to navigate the building even in the absence of offices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

exerted greater control over their relative locations, but those who were resident in BLD2 in 2006 occupied spaces in a long established, popular facility with a low vacancy rate. Under such conditions it is difficult to imagine widespread, successful efforts to engineer higher degrees of functional proximity among investigators who intended to collaborate. As a result, we take the similarity in our key findings across buildings to suggest that investigator choice and organizational policy are unlikely to be driving the results.

Finally, our primary measure, path overlap, relies on the relative locations of two types of spaces whose location investigators might be able to influence (labs and offices) and three types of spaces (restrooms, elevators, and stairs) that are fixed, shared features of buildings. In order for investigator preferences to drive proximities as we measure them, pairs of scientists who are not yet collaborators would need to exert control over the allocation of two sets of scarce resources while considering pathways between them and taking into account the complications created by their location relative to fixed design features. While such calculations may be possible, they are not likely to be widespread or particularly effective.

Our measure of path overlap itself rests on several assumptions that should be tested in future research. First, we assume

Fig. 4. (a) Offices (in red) in BLD2 are clustered around the core of the building. The connections between offices and other spaces are on the basis of immediate adjacency, which is a more direction connection than in BLD1. (b) The effect of removing offices in BLD2 is to nearly sever the two wings or ends of the floor plan. This has a significant impact on the overlaps between researchers, and is likely to be even more pronounced in instances where ingress and egress onto the floor is through the two staircases on either end of the floor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

that the most relevant spaces for defining an investigator's functional spaces are his or her offices and labs along with restrooms, elevators, and stairs.[14] We do not consider the possibility that other types of public spaces (e.g. break rooms or kitchens) or scientific locales (e.g. shared instruments or animal colonies) might define functional zones. Similarly, we assume that individuals favor the shortest, most efficient pathways among their various spaces. This assumption too calls for future empirical investigation. Many things might entice people to take longer paths. Friends to be sought out, rivals to be avoided, favored views and decorations, commitments to exercise, or even simply whimsy might alter habitual pathways and with them patterns of overlap. As a result, further research will be necessary to establish deeper insight into the relationship between specific features of the built environment and patterns of occupancy and movement. Moreover, future work should address the likelihood that social dimensions such as power and authority, and differences associated with age, status, scientific field, or gender might shape both the kinds of spaces people occupy and the paths they walk.

Lastly, we anticipate that our study could benefit from additional fine-grained data on spatial use patterns and actual interaction behaviors. Potential sources of these data include ethnographic and observational data. These data would enable us to further refine the path overlap concept by addressing questions such as how organizational culture affects the likelihood that people leave their office doors open and increase the likelihood of serendipitous encounters. Data on actual behaviors and patterns of spatial use would enable further refinement of path overlap as a measure for predicting tie formation and collaboration success. Location tracking data, for example, allow the possibility of adding a temporal element to the path overlap mechanism. These logical extensions hold the promise of significantly improving our theoretical understanding of spatial proximity and its role in the formation of new relations, as well as how it impinges on the outputs of the dyads thus formed.

[14] As we note in the sensitivity analysis section, removing offices from the spaces that constitute the functional zone changes the significance of path overlap in BLD2, a finding that we think is attributable to the clustering of offices around the building core in BLD2 in contrast to the linear, more dispersed locations of offices in BLD1.

6. Conclusion and implications

The work we present here takes important steps to measure and test how functional, naturalistic conceptions of proximity shape the formation and outcomes of relationships. Our work proposes specific measures targeted to studies of collaboration and workplace design.

Efforts to establish systematically how ties form and what different mechanisms of formation mean for the structure and outcomes of larger social systems are becoming ever more important. While much work has focused on inter-organizational relationships (Powell et al., 2005; Sorenson and Stuart, 2008), new attention is being paid to the dynamics of interpersonal networks (Zaheer and Soda, 2009). Physical and social conceptions of space alike have played important roles in both types of analysis (Kossinets and Watts, 2006; Liu, 2010; Owen-Smith and Powell, 2004; Whittington et al., 2009).

Little work has paid nuanced attention to the role of the built environment in tie formation. Yet people live and interact in physical space. Organizations, likewise, pursue their goals, form partnerships, and compete in particular places. As a result, a full effort to develop comprehensive theories of the collective dynamics of networks both in and outside of the workplace will benefit from greater attention to the spatial environments in which many dyadic connections form. We expand on a classic study of interpersonal tie formation using new methods from architecture to offer one route to a sociospatial science of network dynamics. Emphasizing how people's habitual patterns of movement through built space create or hinder passive encounters and how those encounters become progressively deeper interactions offers fertile ground for future work on the collective dynamics of networks.

Our research has important implications for studies of the dynamics and outcomes of scientific collaboration. As we suggest in our introduction, research on the importance of physical space for collaboration formation and effectiveness is mixed. Particularly in the early stages of research, where individuals are more or less actively prospecting for new ideas, and when success relies on the coordination of highly tacit work, physical proximity is especially important. Our findings suggest a new way to address the role of physical proximity in the direction and outcomes of spatially embedded dyadic collaborations.

Finally, our findings suggest new directions for efforts to design, renovate, and allocate research space in settings concerned with innovation. While care should be taken to identify and account for differences in broad fields of research, the effects we find for path overlap in these two buildings imply that spatial layouts influence collaboration and potentially discovery more than has generally been recognized. Homans characterized work environments as constituting of interrelated social, physical, and technical factors (Homans, 1950). This interrelatedness should matter to designers because it suggests that work environments function best when spatial layouts are configured in response to what is known about the organizational processes and structures that impinge on task performance and social interactions. The reality, however, is that spatial layouts are typically designed on the basis of untested assumptions about the links between physical space and human behavior (Porteous, 1971). Our findings furnish evidence for the actual mechanisms that govern social interactions in the context of scientific collaboration, in the process affirming the role of spatial layout in making possible overlaps in the functional zones of individuals (Festinger et al., 1950). These overlaps can be viewed as precursors to chance encounters and social interactions between individuals. Workplace design goes beyond the mere provision of buildings and furniture. Indeed it implies an intervention into a social system (Goodrich, 1982). Our findings have the potential to shift the conversation away from well-worn debates like the one on "open versus closed offices" and toward more nuanced models of sociospatial proximity such as the functional zone.

One area where applying the functional zone concept may assist designers is in better predictions of how layouts may promote or hinder specific social interactions, thus allowing architects to avoid unintended negative consequences of workplace design. Consider the hypothetical case where restrooms are one of the key markers or framers of individuals' functional zones in a specific workplace. Placing the men's and women's restrooms on different ends or sides of the office may significantly lower the likelihood of social interactions between those of different genders. Knowledge work, such as the scientific collaborations in this study, is contingent on interactions and the resultant exchange of information and other resources between individuals. In the scenario where restrooms are separated by gender, we would expect that there would be fewer collaborations between people of the opposite sex given that the spatial layout would diminish the opportunities for mixed sex dyads to have unscripted social interactions at the workplace.

Building on a rich history of qualitative and descriptive research that takes physical space seriously, we propose new methods to characterize the built environment and track collaborations and their outcomes in administrative data that create exciting new directions for research in several fields. Such work will form the basis of what we hope will become a robust sociospatial network science of innovation that can make fundamental contributions to social science research on networks while offering concrete and potentially useful insights to designers, administrators, and scientists themselves.

Acknowledgements

Funding for this project was provided by the National Science Foundation (Grants # 1158711 and 1262447) and the following units at the University of Michigan: the Office of the Vice-President for Research (OVPR), the Social Science Annual Institute (SSAI), the Institute for Research on Labor, Employment, and the Economy (IRLEE), the Institute for Social Research (ISR), and the North Campus Research Complex (NCRC). We thank Jerry Davis and Rick Price for invaluable feedback.

References

Abramovsky, L., Simpson, H., 2011. Geographic proximity and firm-university innovation linkages: evidence from Great Britain. Journal of Economic Geography 11, 949–977.

Agneessens, F., Wittek, R., 2012. Where do intra-organizational advice relations come from? The role of informal status and social capital in social exchange. Social Networks 34, 333–345.

Agrawal, A., Kapur, D., McHale, J., 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data. Journal of Urban Economics 64, 258–269.

Allen, T.J., 1977. Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information Within the R&D Organization. MIT Press, Cambridge, MA.

Allen, T.J., Henn, G., 2007. The Organization and Architecture of Innovation: Managing the Flow of Technology. Butterworth-Heinemann, Burlington, MA.

Appel-Meulenbroek, R., 2009. Knowledge sharing in research buildings and about their design. In: Koch, D., Marcus, L., Steen, J. (Eds.), 7th International Space Syntax Symposium. KTH, Stockholm.

Atkins, D.E., Droegemeier, K.K., Feldman, S.I., Garcia-Molina, H., Klein, M.L., Messerschmitt, D.G., Messina, P., Ostriker, J.P., Wright, M.H., 2003. Revolutionizing Science and Engineering Through Cyberinfrastructure. National Science Foundation.

Backhouse, A., Drew, P., 1992. The design implications of social interaction in a workplace setting. Environment and Planning B: Planning and Design 19, 573–584.

Bechky, B.A., 2003. Sharing meaning across occupational communities: the transformation of understanding on a production floor. Organization Science 14, 312–330.
Berry, H., Guillen, M.F., Zhou, N., 2010. An institutional approach to cross-national distance. Journal of International Business Studies 41, 21.
Borgoni, R., Billari, F.C., 2002. A Multilevel sample selection probit model with an application to contraceptive use. In: Society, I.S. (Ed.), Proceedings of the XLI Conference of the Italian Statistical Society. Cleup Editrice, Milan, Italy, pp. 251–254.
Brown, J.S., Duguid, P., 1991. Organizational learning and communities of practice: toward a unified view of working, learning and innovation. In: Lesser, E.L., Fontaine, M.A., Slusher, J.A. (Eds.), Knowledge and Communities. Butterworth-Heinemann, Oxford, pp. 99–121.
Cairncross, F., 2001. The Death of Distance: How the Communications Revolution is Changing Our Lives. Harvard Business School Press, Boston.
Collins, H.M., 1974. The TEA set: tacit knowledge and scientific networks. Science Studies 4, 165–185.
Collins, H.M., 1992. Changing Order: Replication and Induction in Scientific Practice. University of Chicago Press, Chicago.
Crane, D., 1972. Invisible Colleges; Diffusion of Knowledge in Scientific Communities. University of Chicago Press, Chicago.
Currarini, S., Jackson, M.O., Pin, P., 2009. An economic model of friendship: homophily, minorities, and segregation. Econometrica 77, 1003–1045.
Currarini, S., Jackson, M.O., Pin, P., 2010. Identifying the roles of race-based choice and chance in high school friendship network formation. Proceedings of the National Academy of Sciences 107, 4857–4861.
Daraganova, G., Pattison, P., Koskinen, J., Mitchell, B., Bill, A., Watts, M., Baum, S., 2012. Networks and geography: modelling community network structures as the outcome of both spatial and network processes. Social Networks 34, 6–17.
Festinger, L., Schachter, S., Back, K.W., 1950. Social Pressures in Informal Groups; A Study of Human Factors in Housing, 1st ed. Harper, New York.
Finholt, T.A., Olson, G.M., 1997. From laboratories to collaboratories: a new organizational form for scientific collaboration. Psychological Science 8, 28–36.
Galison, P., 1997. Image and Logic: A Material Culture of Microphysics. University of Chicago Press, Chicago.
Gieryn, T.F., 2002. What buildings do. Theory and Society 31, 35–74.
Gieryn, T.F., 2008. Laboratory design for post-fordist science. Isis 99, 796–802.
Goodrich, R., 1982. Seven office evaluations: a review. Environment and Behavior 14, 353–378.
Hargadon, A.B., Douglas, Y., 2001. When innovations meet institutions: edison and the design of the electric light. Administrative Science Quarterly 46, 476–501.
Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475–492.
Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.
Hillier, B., 1999. The need for domain theories. Environment and Planning B: Planning and Design 26, 163–167.
Hillier, B., 2007. Space is the Machine_Part 1. Space is the Machine. Space Syntax, London, UK.
Hillier, B., Hanson, J., 1984. The Social Logic of Space. Cambridge University Press, Cambridge, UK.
Hillier, B., Penn, A., 1991. Visible colleges: structure and randomness in the place of discovery. Science in Context 4, 23–50, http://dx.doi.org/10.1017/S0269889700000144, 23M.
Homans, G.C., 1950. The Human Group. Harcourt, Brace, New York.
Kabo, F., Hwang, Y., Levenstein, M., Owen-Smith, J., 2013. Shared paths to the lab: a sociospatial network analysis of collaboration. Environment and Behavior (forthcoming).
King, G., Zeng, L., 2001. Logistic regression in rare events data. Political Analysis 9, 137–163.
Knorr-Cetina, K., 1999. Epistemic Cultures: How the Sciences Make Knowledge. Harvard University Press, Cambridge, MA.
Kossinets, G., Watts, D.J., 2006. Empirical analysis of an evolving social network. Science 311, 88–90.
Kraut, R., Egido, C., Galegher, J., 1988. Patterns of contact and communication in scientific research collaboration. In: Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work, ACM, Portland, OR, United States, pp. 1–12.
Lave, J., Wenger, E., 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge [England]; New York.
Lee, K., Brownstein, J.S., Mills, R.G., Kohane, I.S., 2010. Does collocation inform the impact of collaboration? PLoS ONE 5, e14279.
Liu, C.C., 2010. A spatial ecology of structural holes: scientists and communication at a biotechnology firm. In: Academy of Management Annual Meeting Proceedings, pp. 1–6.
Lynch, M., 1985. Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory. Routledge & Kegan Paul, London; Boston.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annual Review of Sociology 27, 415–444.
Mechanic, D., 1962. Sources of power of lower participants in complex organizations. Administrative Science Quarterly 7, 349–364.
Mizruchi, M.S., Marquis, C., 2006. Egocentric, sociocentric, or dyadic? Identifying the appropriate level of analysis in the study of organizational networks. Social Networks 28, 187–208.
Mok, D., Wellman, B., 2007. Did distance matter before the Internet? Interpersonal contact and support in the 1970s. Social Networks 29, 430–461.
Monge, P.R., Kirste, K.K., 1980. Measuring proximity in human organization. Social Psychology Quarterly 43, 110–115.
Moodysson, J., Jonsson, O., 2007. Knowledge collaboration and proximity. European Urban and Regional Studies 14, 115–131.
Ocasio, W., 1997. Towards an attention-based view of the firm. Strategic Management Journal 18, 187–206.
Olson, G.M., Olson, J.S., 2000. Distance matters. Human–Computer Interaction 15, 139–178.
Olson, J.S., Teasley, S., Covi, L., Olson, G., 2002. The (currently) unique advantages of collocated work. In: Hinds, P.J., Kiesler, S. (Eds.), Distributed Work. Academic Press, Cambridge, MA, pp. 113–136.
Owen-Smith, J., 2001. Managing laboratory work through skepticism: processes of evaluation and control. American Sociological Review 66, 427–452.
Owen-Smith, J., Powell, W.W., 2004. Knowledge networks as channels and conduits: the effects of spillovers in the Boston Biotechnology Community. Organization Science 15, 5–21.
Penn, A., Desyllas, J., Vaughan, L., 1999. The space of innovation: interaction and communication in the work environment. Environment and Planning B: Planning and Design 26, 193–218.
Pitts, F.R., 1978. The medieval river trade network of Russia revisited. Social Networks 1, 285–292.
Polanyi, M., 1967. The Tacit Dimension. Anchor Books, Garden City, NY.
Porteous, J.D., 1971. Design with people: the quality of the urban environment. Environment and Behavior 3, 155–178.
Powell, W.W., White, D.R., Koput, K.W., Owen-Smith, J., 2005. Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences. American Journal of Sociology 110, 1132–1205.
Powell, W.W., Koput, K.W., Smith-Doerr, L., 1996. Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. Administrative Science Quarterly 41, 116–145.
Rivera, M.T., Soderstrom, S.B., Uzzi, B., 2010. Dynamics of dyads in social networks: assortative, relational, and proximity mechanisms. Annual Review of Sociology 36, 91–115.
Sailer, K., McCulloh, I., 2012. Social networks and spatial configuration – how office layouts drive social interaction. Social Networks 34, 47–58.
Schumpeter, J.A., 1942. Capitalism, Socialism, and Democracy. Harper & Brothers, New York and London.
Serrato, M., Wineman, J., 1999. Spatial and communication patterns in research & development facilities. In: 2nd International Space Syntax Symposium, pp. 11.11–11.18.
Sorenson, O., Stuart, T.E., 2008. Bringing the context back in: settings and the search for syndicate partners in venture capital investment networks. Administrative Science Quarterly 53, 266–294.
Tomz, M., King, G., Zeng, L., 1999. ReLogit: Rare Events Logistic Regression. Department of Government, Harvard University, Cambridge, MA.
Van de Ven, W.P.M.M., Van Praag, B.M.S., 1981. The demand for deductibles in private health insurance: a probit model with sample selection. Journal of Econometrics 17, 229–252.
Weiser, M., Gold, R., John Seely, B., 1999. The origins of ubiquitous computing research at PARC in the late 1980s. IBM Systems Journal 38, 693–696.
Whittington, K.B., Jason, O.-S., Powell, W.W., 2009. Networks, propinquity, and innovation in knowledge-intensive industries. Administrative Science Quarterly 54, 90–122.
Wineman, J.D., Kabo, F.W., Davis, G.F., 2009. Spatial and social networks in organizational innovation. Environment and Behavior 41, 427–442.
Zaheer, A., Soda, G., 2009. Network evolution: the origins of structural holes. Administrative Science Quarterly 54, 1–31.