Design Science

Co-author network analysis of human-centered design for development

Nancy Li[1], Julia Kramer[2], Pierce Gordon[3] and Alice Agogino[2]

1 *University of California Berkeley, Electrical Engineering and Computer Science, Berkeley, California 94720, USA*
2 *University of California Berkeley, Mechanical Engineering, Berkeley, California 94720, USA*
3 *University of California Berkeley, Energy and Resources Group, Berkeley, California 94720, USA*

Abstract

Human-centered design (HCD) offers a systematic approach to innovation practice, driven by customer research and feedback throughout the design process. Within the community of engineers and researchers who engage in design for global development, interest in HCD has grown in the past decade. In this paper, we examine the human-centered design for development (HCD+D) academic community to better understand the interactions between researchers. By building and evaluating a co-authorship network from a dataset of HCD+D papers, in which the nodes are researchers and the connecting links are co-authorship relationships, we provide a decade-long benchmark to answer a variety of questions about collaboration patterns within this emerging field. Our analysis shows that most HCD+D authors publish few papers and are part of small, well-connected sub-communities. Influential authors that bridge separate communities are few. HCD+D is emerging from disparate disciplines and widely shared scholarship across disciplines continues to be developed. Influential authors in HCD+D play a large role in shaping HCD+D, yet there are few authors that are in a position to connect and influence collaborative research. Our analysis gives rise to several implications including an increased need for cross-disciplinary collaboration and the need for a stronger core of HCD+D practitioners.

Key words: cross-disciplinary collaboration, co-author network analysis, human-centered design, global development

1. Introduction

Human-centered design (HCD) is an approach to developing and leveraging a deep understanding of potential users and stakeholders toward the creation of novel interventions of value to the stakeholder community. In this paper, we consider the six principles of HCD introduced in ISO 9241-210, the international standard of HCD for interactive systems (ISO 2010):

- 'the design is based upon an explicit understanding of users, tasks and environments;

- users are involved throughout design and development;

- the design is driven and refined by user-centered evaluation;

- the process is iterative;

- the design addresses the whole user experience;

- the design team includes multidisciplinary skills and perspectives.'

HCD has recently manifested as a widespread approach to create interventions addressing challenges of global development. Human-centered design for development (HCD+D) is the practice of using HCD approaches to create interventions with an explicit aim to address the multifaceted complex issues of poverty and underdevelopment around the world. For example, an intervention created using an HCD process is a portable ultrasound device for midwives in Ghana (Brunette *et al.* 2010). Another example is a mobile application for patients receiving contraceptive counseling in the United States (Gilliam, Martins & Bartlett 2014).

Given its recent emergence, we aim to understand the current research landscape of HCD+D and find patterns of co-authorship collaboration. A core principle of HCD is to include cross-disciplinary collaboration, given its focus on human needs, technological possibilities, and business viability. The collaborative HCD process invokes designers and researchers to work in teams with other stakeholders and end-users to create useful design interventions. This cross-disciplinarity is of particular importance in HCD+D, where solutions are developed to address pressing and serious social needs. However, no study to our knowledge has aimed to systematically characterize the research community in HCD+D. We aim to fill this gap by quantifying the network of HCD+D authors. Because we are interested in the cross-disciplinary interactions among researchers who engage in HCD+D, we apply social network analysis (SNA) techniques to HCD+D co-author relationships.

In this paper, we present a quantitative co-author network analysis of the community of researchers engaged in HCD+D from a targeted database of 278 authors and 78 papers. Co-authorship of a paper represents a sustained collaboration effort between the co-authors. In contrast to a citation network, the structure of a co-authorship network reveals knowledge and collaboration patterns of the HCD+D scholarly community.

To provide motivation and context for our quantitative analysis, we first present a brief background on design, HCD, and development. We then discuss previous related research used to understand the importance of cross-disciplinary teamwork, particularly in the context of complex research-oriented problems. After presenting background on these motivating research thrusts, we then detail our methodology and our analytical techniques for network analysis. Finally, we present our results along with a discussion of implications for HCD+D scholarly research and directions for further research.

2. Background & motivation

2.1. Design thinking, human-centered design & development

As Ramirez Jr (2011) points out, a large amount of approaches interested in 'design with a social conscience' exist today, including green design, universal (accessible) design, and corporate social responsibility campaigns. The focus in this paper is on human-centered design for development (HCD+D). HCD is viewed as a

particularly useful framework in development because it allows practitioners to gain a deep understanding of customers and stakeholders tied to their design context, and in its ideal form, HCD gives practitioners the freedom to modify any part of the design context toward its betterment. Given its iterative focus on creative thinking, HCD+D stimulates designers to conduct extensive background research and fieldwork, generate a wide swath of ideas, and rapidly build and test these ideas.

Although 'design thinking' (Rittel & Webber 1973; Rowe 1987; Brown & Rowe 2008) and 'human-centered design' are sometimes viewed as synonymous, we choose to focus on 'human-centered design' because the HCD methodology is specific to practices that are driven by stakeholder research and feedback at all stages of the design process. 'Design thinking' is often used more broadly and can include creative thinking or prototyping without significant stakeholder involvement.

IDEO, a global design consultancy firm, formalized their practice of HCD for development with their publication of the Human-Centered Design Toolkit (IDEO 2009). This toolkit explicitly aimed to explain how practitioners can use HCD to 'enhance the lives of people living on less than $2/day' (IDEO 2009; Fuge & Agogino 2015).

We define 'development' using three criteria: (1) work with a community experiencing a form of multidimensional poverty, (2) work with a community experiencing 'institutional voids,' or the absence of supportive intermediary institutions like credit card companies (Palepu & Khanna 2010; Levine, Agogino & Lesniewski 2016; Clarke 2015), or (3) work with a community experiencing a loss of freedoms or capabilities (Sen 1993).

2.2. Cross-disciplinarity in human-centered design

HCD is an inherently cross-disciplinary field, as demonstrated in the formative historical literature on design. Herbert Simon's *Sciences of the Artificial* book (1969) explores artificial intelligence, complexity, and design. In his section on 'The Science of Design,' he argues for 'the role of design in the life of the mind,' and suggests that design is 'a common core of knowledge that can be shared by the members of all cultures.' The seemingly insurmountable division between, in Simon's example, engineers and composers, is bridged by the understanding that design is 'the common creative activity' underlying those (and, of course, other) fields.

Buchanan (1992) echoes Simon's notion, and adds that as liberal arts and sciences have become more specialized, they have contributed to a more fractured research ecosystem. He points out the need for 'integrative disciplines' to cohere narrowly specialized research fields. Buchanan goes on to suggest that design thinking (which, as noted in the previous section, is related to HCD) serves as one such integrative discipline connecting knowledge from arts and sciences.

As corollary to Simon and Buchanan's suppositions, we propose that, by their very being, HCD practitioners are knowledge integrators and cross-disciplinary connectors. Design is a bridge between fields, and design practitioners are physical embodiments of this connection across disciplines. Therefore, as we seek to understand the nature of collaboration in HCD+D, we are necessarily interested in pointing our lens toward cross-disciplinary collaboration.

Laudel (2002) suggests that collaboration is defined by collaborative 'research activities,' which are the 'actions that are aimed at the production of new scientific knowledge' (Krohn & Küppers 1990). In Beaver's (2001) study of scientific collaboration, he finds a broad scope of reasons why people collaborate, including access to expertise and resources, improved efficiency and productivity, decreasing one's feelings of isolation, and advancing knowledge and learning. Beaver also notes 'physical location is no longer a barrier to the free and easy exchange of information.'

In this paper, we explore the nature of HCD+D collaboration through a co-authorship social network analysis (SNA), with a focus on cross-disciplinary collaborations. Yang & Heo (2014) provide an overview of three 'levels of integration' (multidisciplinary, interdisciplinary, and transdisciplinary) in cross-disciplinary research collaborations:

> 'Multidisciplinary research refers to research in which researchers from more than one discipline work independently on different areas of a project while remaining within their disciplinary boundaries. Interdisciplinary research is defined as research efforts in which researchers from various disciplines work in partnership on a project using their discipline-specific perspectives. Transdisciplinary research is undertaken by researchers from different disciplines collaborating on a project using a shared framework that integrates various disciplinary approaches into a collective whole.'

For this study, we use the term 'cross-disciplinary' because it serves as a broad umbrella term, with multi-, inter-, and transdisciplinary research approaches having more specific definitions but still falling under the purview of cross-disciplinary research. We do not attempt to characterize research collaborations based on their level of integration, as is done by Yang & Heo (2014).

2.3. Social network analysis

In our attempt to understand the network of researchers who engage in HCD+D, we employ visual and mathematical techniques from SNA, a method to map and measure relationships between people, groups, organizations, computers, or other connected entities (Scott 2013). Using co-authorships to proxy collaboration, we use SNA to quantitatively study the cross-disciplinary interconnections between HCD+D researchers.

SNA has been used to study emerging research disciplines, look at patterns of collaboration, identify key researchers, and study patterns of cross-disciplinarity in various academic communities. For example, Uddin *et al.* (2012) looked at the co-author network of papers on steel structures published since the 1970s to understand the field's evolution and emergence. Yang & Heo (2014) studied the cross-disciplinarity of different research fields in Korea, measuring betweenness, closeness, and eigenvector centrality to identify Materials Science, Biotechnology, and Nanoscience as fields with the most variation in author disciplines. Newman (2001a,b) assessed co-authorship of four databases: MEDLINE (biomedical), the Los Alamos e-Print Archive (physics), SPIRES (high energy physics), and NCSTRL (computer science). They looked at metrics including papers per co-author, co-authors per author, betweenness, and collaboration weight to find the most influential individuals. We use similar metrics in our work to understand the emerging HCD+D field.

Liu *et al.* (2015*a*) proposed a method based on the PageRank algorithm to evaluate the importance of authors in a co-author network; their method measures the influence of each authors' papers over time by their number of citations. Their focus was on determining the importance of authors in co-author networks to better evaluate the impact of a paper by calibrating the paper's influence over time. This method, however, is out of this paper's scope to understand the emerging HCD+D community.

We do consider influential authors as 'brokers' of knowledge between different actors. Haythornthwaite (1996) explained how 'brokerage' relationships (i.e., those relationships that serve to connect otherwise disconnected groups) represent the potential to control the information that flows between others. Brokerage is also measured by the betweenness centrality index, and an actor in a brokerage position is a gatekeeper that filters and shares information among their connections.

Zare-Farashbandi, Geraei & Siamaki (2014) looked at the co-authorship network of articles from the *Journal of Research in Medical Sciences* to assess researchers' willingness to cooperate with other members. Betweenness and closeness scores were used to find the most influential individuals. Liu *et al.* (2015*b*) used the same centrality measurements to look at the digital library research community, ranking author statuses and summarizing the health of the collaboration network.

The evolution of a co-author network can be studied in order to predict the network's future growth (Barabasi, Vicsek & Palla 2007). Barabasi *et al.* (2007) found that small co-author communities are most stable when its members stay for a long time, whereas larger communities are the most stable with high turnover and fluctuation among its members.

Liu *et al.* (2015*b*) illustrate how scientific collaborations are structured. Unlike citation networks, the co-authorship links represent previous engagements in scholarly collaboration. Co-authorship analysis gives a unique opportunity to see not how researchers gather knowledge, as is assumed in a citation network, but how they collaborate with colleagues toward design-based research.

3. Aims

HCD+D requires cross-disciplinary collaboration, and no study to our knowledge has attempted to study cross-disciplinary collaboration in the HCD+D research community. Therefore, in this paper, we survey the landscape of the publishing HCD+D community with a view to answering the following research questions:

- Who is publishing together, and who are the influencers?

- What are the cross-disciplinary and collaboration characteristics of the network?

- How has the co-author network changed over time, year to year?

We use SNA of a co-authorship network in order to find groups of authors publishing together to understand the characteristics of their collaboration, and to determine who the influencers are year by year. In this study, we use co-authorship relationships to proxy collaboration. Co-authorships represent a direct and visible mark of collaboration between two or more authors. While it is

certainly true that not all forms of 'collaboration' are rewarded with co-authorships (as Laudel (2002) points out), we are studying the strong collaborative practices that result in co-authorship. Rarely do designers capture their whole design process; peer-reviewed publications give an opportunity to study the narratives the authors decide to prioritize. By learning about how authors connect in this deep, interactive, narrative process of communicating research, we learn the structure of the co-authorship network of HCD+D practitioners that can be used as a foundation in the future to see how those collaborative connections shape the projects or the field. While we are excluding other forms of collaborations at this point, we intentionally begin our analysis in the published academic research to explore the emergence of HCD+D research and to provide a decade-long benchmark from which to study this evolving new cross-disciplinary field in the future. This study fills a gap in the literature by providing a quantification of cross-disciplinary collaborations present within the first decade of the emerging HCD+D research field.

We measure cross-disciplinary characteristics of the network by classifying papers into focus areas, and looking at the number of different author disciplines represented in each paper. We determine the key influencers of the HCD+D network by analyzing metrics of closeness and betweenness centrality. We also consider 'cut-point' authors, who are authors that connect two separate sub-communities.

This paper contributes to the broader design literature by providing an in-depth study of a particular community of designers. In our study, we clarify the notion of 'cross-disciplinary collaboration' and we employ novel quantitative practices to study cross-disciplinary collaboration. These contributions benefit the design community at large by providing an example of how to study and reflect upon collaborative practices in design, a cornerstone of the HCD process.

The practice of applying design techniques in global development is not new, but the formal scholarly literature around this topic is relatively recent. Therefore, our study can open up a reflective conversation about empirical patterns of current collaboration in HCD+D, with a view to improving cross-disciplinary collaborative practices in the future.

4. Data and method

4.1. Data collection

Our dataset consists of 78 HCD+D papers, written by 247 authors, and published during HCD+D's foundational decade between 2004 and 2014. To obtain this dataset, we conducted a Google Scholar search over a set of terms related to HCD+D (i.e., 'human-centered design' plus 'developing countries,' 'developing economies,' 'developing world,' 'global development,' 'global inequality,' 'international development,' 'low-income,' 'low-resource,' 'poverty,' 'resource-limited,' or 'third world'). This search gave us an initial set of 1441 papers that included any of the above keyword pairs in the title, abstract, keywords, or main text.

While there are many other frameworks similar to HCD (e.g., design thinking, human–computer interaction, etc.), we limit this analysis to only focus on papers that specifically say they engage in HCD work in order to narrow the focus to the

community of researchers engaged in customer- or stakeholder-driven design and evaluation.

We then systematically filtered down the set of papers using filters that (1) excluded papers with no citations, if published before 2014; (2) excluded books; (3) excluded papers not available in English or not accessible to us online; (4) excluded papers that were not peer-reviewed (e.g., theses, dissertations, or policy briefs) and papers that were not discussing work held in an actual research site (e.g., theory papers); and (5) excluded papers where the authors themselves were not actually engaged in the design work (e.g., a literature review). Therefore, what is included in our set of papers are archival peer-reviewed papers written in English that serve as practical examples of researchers engaging in an HCD+D approach. The papers in our set are all on-the-ground projects where the authors themselves engaged in the work. What is not included in our set of papers are theory papers, second-hand accounts, and projects not explicitly engaged in 'development' work, as we have defined previously.

A more detailed description of our process, including the rationale for these filters, is explained in our working paper (Gordon *et al.* 2017). The full list of papers included in our analysis is available at www.tinyurl.com/hcddpub lications.

4.2. Methods for representing and analyzing the co-authorship network

The names and titles in our dataset were pre-processed to replace all special characters with acceptable ASCII characters. All author names were translated into a common first initial, last name representation (e.g., J. Smith). We found no two authors shared both first initial and last name in our sample. We manually collected the disciplines of each author, as represented by their listed affiliation on each publication, to perform separate analyses on author disciplines represented in each focus area of HCD+D papers.

To build a co-authorship matrix, pairs of co-authors were extracted from the data and parsed using Python. The number of contributions for each pair of co-authors was incremented each time the pair published a paper in the dataset. This matrix was used to generate the visual graph shown in Figure 1, using Python's JSON library, the d3.js JavaScript library, and HTML.

Density, clustering coefficient, network diameter, largest connected component, betweenness centrality, closeness centrality, and authors who are 'cut-points' were identified as important metrics for this research. Density measures the cohesion of the network, while the clustering coefficient tells us the tendency of authors to collaborate with each other. The network diameter and the size of the largest connected component give us a sense of the longest path in the network – showing how quickly authors can communicate with each other through their links – and the biggest group of collaborators. Betweenness and closeness centralities measure an individual author's social status in the network, and we use this to help find influential researchers in HCD+D. Betweenness indicates the number of times an author is located between any two other authors in the network on the shortest path; authors with high betweenness are good connectors. Closeness measures the average length of the shortest paths between an individual author and all other authors; authors with high closeness are located in the 'best' positions of the
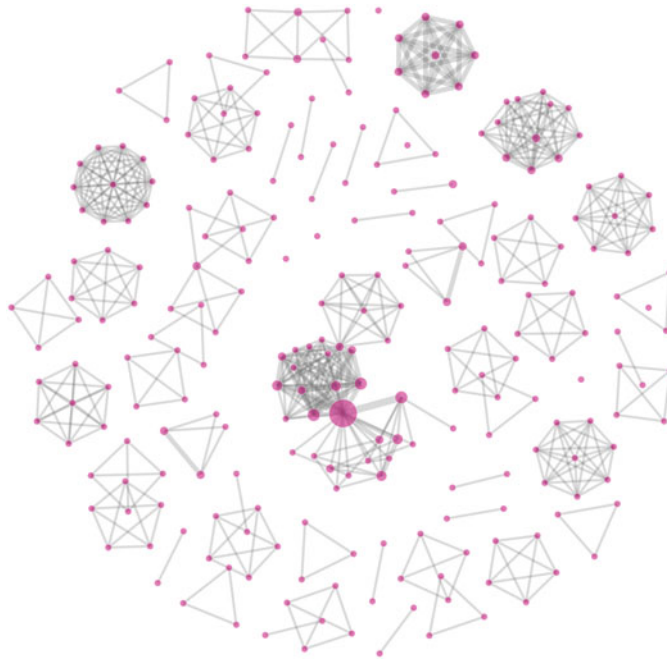
Figure 1. HCD+D network of authors who published between 2004 and 2014. Each node is a unique author, and edges between nodes represent a co-authorship relationship, weighted by the number of co-authored publications.

network in terms of quickly reaching all other authors. Finally, we also identify cut-point authors. These authors are part of a connected community such that their removal causes the community to become disconnected. We consider these cut-point authors influential for helping glue the overall network.

These metrics were collected using Python's NetworkX library. A more detailed and mathematical description of these metrics can be found in the appendix A.

5. Results

Using the co-authorship matrix, we generated a visual network with weighted and undirected links (since co-authorship is bi-directional), shown in Figure 1. Each node represents a unique author and node sizes are weighted by the number of publications the author has made in total. The edge between a pair of nodes represents co-authorship, and the edge thickness is weighted by the number of publications co-authored by that pair. Figure 1 shows the fragmentation and sparsity of the HCD+D network.

To assess the cross-disciplinarity of our HCD+D network, we listed unique 'disciplines' of each author, looking at their listed departmental affiliation at the time they published each paper. Table 1 shows how a variety of authors had slightly different disciplines in name, though we collapsed these disciplines into general clusters.

We then assessed the areas the authors of these disciplines tended to work in. We found eleven 'focus areas' by considering the goals of four global development organizations and initiatives: the United States Agency for International

Table 1. List of HCD+D author disciplines

Discipline cluster	Affiliation as named in the papers
Design	HCD, art, architecture, design, industrial design, engineering design
Information technologies	IT, information systems, system sciences, informatics, information science, computing
Business	Economics, management, business
Health	Healthcare, medicine, life sciences, radiation sciences, pharmacy, public health
Environmental studies	Agriculture, urban, water & sanitation, environmental science
Communications	Journalism, media, computing & digital media
Education	Education policies
Advocacy	Social welfare, deaf community
Humanities	Anthropology, sociology, linguistics, political science, English
Engineering science	Biomedical, mechanical, materials science, industrial, computer science
Development studies	Public policy, international affairs, global studies, development

Development (USAID) (USAID n.d.), the United Kingdom's Department for International Development (DFID) (Gov.UK n.d.), the Sustainable Development Goals (SDG) proposed in 2015 by the United Nations (United Nations 2015), and the Millennium Development Goals (MDG) proposed in 2000 by the United Nations (United Nations 2000). We pooled the areas of work from these organizations and grouped them into like categories. The resulting eleven focus areas are shown in each row of Table 2 below. We then classified each paper of our dataset into the appropriate focus area(s). Table 2 shows the percentage of authors from each discipline that wrote papers in each of these focus areas.

From this table, we see that papers written about inclusive infrastructure, food security, and economic inclusion projects were the most cross-disciplinary, with an even spread of authors from different disciplines. In our dataset, each paper represents an average of 1.78 different disciplines. We calculated the network's degree of cross-disciplinary collaboration (DCC) to be 0.56. Qiu's (1992) work found that journal publication communities tend to have low degrees of cross-disciplinary collaboration (less than 0.1), but there has been little other work to calculate DCC in other publishing communities. This figure is calculated from the number of papers authored by researchers from different disciplines divided by the total number of papers (Qiu 1992).

$$\text{DCC} = \frac{\text{Number of Cross-disciplinary Papers}}{\text{Total Number of Papers in Network}}.$$

We computed 'network metrics' as shown in Table 3 and 'structural metrics' as shown in Table 4. In the following section we discuss findings from both types of metrics.

On average, an author in our network published 1.22 papers. Figure 2 shows the distribution of the number of papers written by authors in our network. 213 authors published a single paper within our dataset, while only four authors published four or more papers. These low numbers may be due to the fact that

Table 2. Papers written by authors of different disciplines (design, information technology, business, health, environmental studies, communication, education, advocacy, humanities, and development studies) in different focus areas. The numbers listed in the cells are percentages with the exception of the "Total" column, which represents the count of the number of papers in each focus area

Paper focus	Design	Info tech	Business	Health	Env	Comm	Education	Advocacy	Humanities	Engineering	Dev studies	Total number of papers
Poverty and inequality	18.92	8.11	0	59.46	2.70	0	0	0	0	10.81	0	37
Inclusive infrastructure	25.00	21.55	0	25.00	2.59	0.86	3.45	0	2.59	16.38	2.59	116
Economic inclusion	33.33	18.52	0	7.41	0	3.70	0	0	7.14	29.63	0	27
Food security	11.54	11.54	11.54	30.77	7.69	0	7.69	0	3.85	15.38	0	26
Education	18.60	4.65	0	51.16	2.33	2.33	0	2.33	0	16.28	2.33	43
Global health	23.96	16.67	0	38.54	1.04	2.08	1.04	0	2.08	13.54	0	96
Global partnership	38.46	0	0	15.38	0	0	0	0	0	46.15	0	13
Water and sanitation	33.33	26.67	0	13.33	0	0	0	0	0	26.67	0	15
Environmental sustainability	33.33	44.44	0	0	0	11.11	0	0	11.11	0	0	9
Governance, human rights, conflict	0	62.50	0	37.50	0	0	0	0	0	0	0	8

Table 3. Summary of network metrics

Network metric	Value
Total authors	247
Total papers	78
Average papers published per author	1.22
Average authors per paper	3.85
Average co-authors per author (degree)	4.44
Average degree of cross-disciplinary collaboration (DCC)	0.56
Average disciplines represented per paper	1.78

Table 4. Summary of structural metrics

| Structural metric	Value
Density	0.018
Diameter	4
Largest component size	34
Average component size	4.26
Average clustering coefficient	0.82
Average path length	1.04
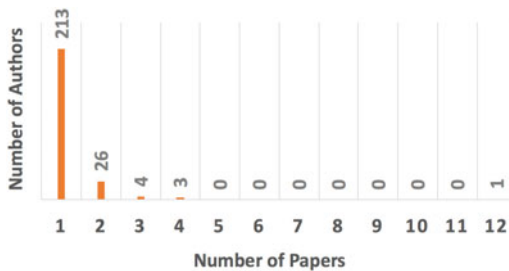
Figure 2. Distribution of the number of papers published per HCD+D author (average = 1.22).

HCD+D is an emerging field and that there are limited publication venues with this focus.

Figure 3 shows the distribution of the number of co-authors (an author's degree) in the HCD+D network. Most authors co-authored with up to 6 other people, with the average being 4.44 other authors.

The papers in our network have an average of 3.85 authors, and the network distribution is shown in Figure 4. Papers written by five or more authors are rare, with 11 being the maximum number of co-authors in a single paper.

11/24
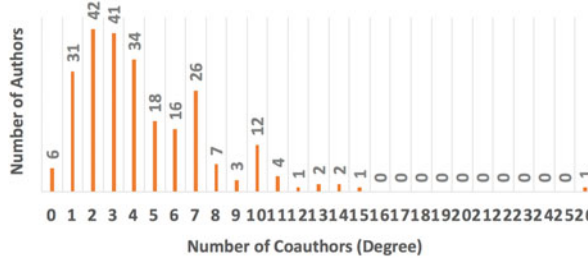
Figure 3. Distribution of the number of co-authors per HCD+D author (i.e., author's degree) (average = 4.44).
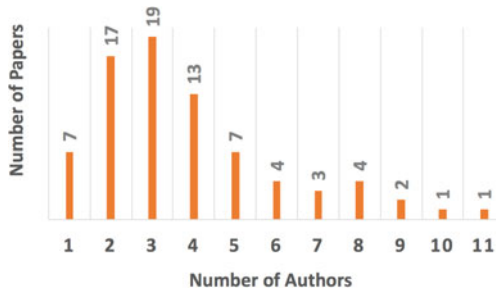

Figure 4. Distribution of the number of HCD+D co-authors per paper (average = 3.85).

We also look at the overall network structure and present metrics quantifying how authors are organized in the whole of the network. We have called these 'structural metrics' (Table 4).

The density of our co-author network is 0.018, representing a very sparsely connected network. This sparsity is visible in Figure 1. The diameter of the largest connected component in our network is 4, while the average shortest path length between any pair of authors is 1.04. The small average path length is due to the overall network being composed of small disconnected cliques. The highly fragmented nature of this field implies that it is difficult for information to flow throughout the network.

The clustering coefficient of the entire network is defined as the average clustering coefficient of all nodes in our network. The average clustering coefficient of our network is 0.82, and the distribution of authors' local clustering coefficients can be seen in Figure 5. The average value being nearly one indicates that authors tend to form persistent publishing cliques.

Component size is the number of nodes within a connected sub-community. The network's largest connected component included 34 nodes, representing 14 percent of the total network. The whole co-authorship network examined in this paper is not a connected graph but rather 58 total disconnected components (or sub-communities), with an average component size of 4.26 nodes per sub-community. This figure is consistent with another author network study conducted on the ASME Design Automation Conference (DAC) Network, which found that most clusters have three to four individuals (Halasz 2015).
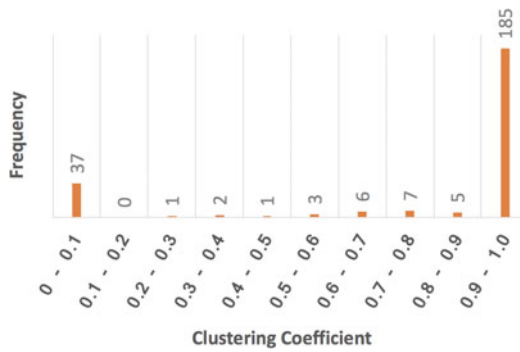
Figure 5. Distribution of HCD+D authors' clustering coefficients (average = 0.82).
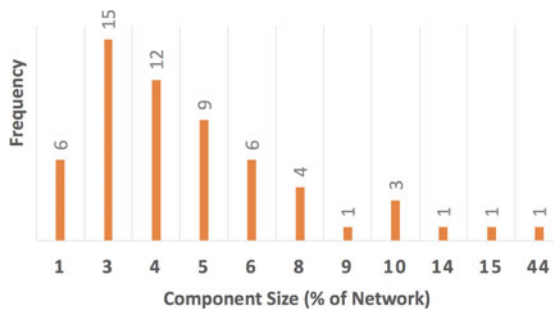
Figure 6. Distribution of connected component (sub-community) sizes (average = 4.26).

The distribution of connected component sizes is shown in Figure 6. The most frequently occurring sub-community sizes are three and four authors.

5.1. Influential authors

In this section, we identify the most important nodes within our network using measures of closeness and betweenness centrality. Figure 7 shows that the great majority of authors have low betweenness (i.e., they do not lie on the shortest paths between other pairs of authors and thus may not have the ability to regulate information flow to others in the network).

Only a few prominent authors hold a position as the most common mutual co-author between their connections. Three authors are at the 99th percentile for high betweenness (see Table 5).

Another measure of influence is closeness centrality; authors with high closeness centrality are good at propagating information throughout the network, as information originating from them reaches others quickly (Newman 2001c). Figure 8 suggests that few authors in the network are located in central positions. There are only four authors in our network at the 98th percentile for high closeness; all of these authors also have high betweenness scores (see Table 5). The uneven distribution indicates that most authors cannot quickly obtain or disseminate information to the rest of the HCD+D research community.
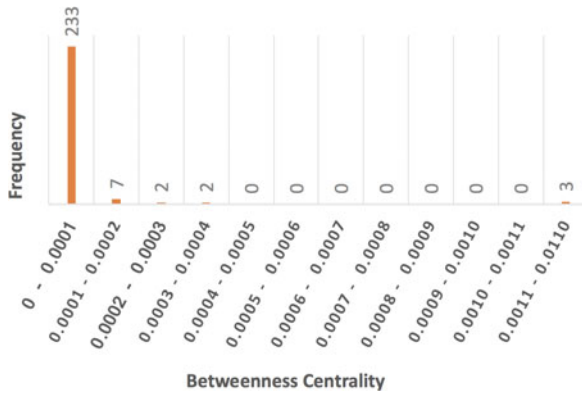
13/24

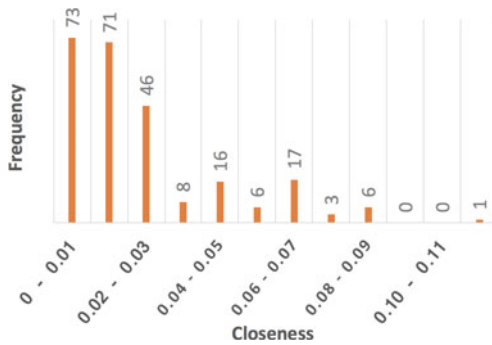**Figure 7.** Distribution of HCD+D authors' betweenness centrality.

Figure 8. Distribution of HCD+D authors' closeness.

Table 5. Analysis of influential authors. Author identifier numbers refer to the communities these authors are in (e.g. Author 2a is a member of sub-community 2) or the communities these authors connect (e.g. Author 1-2 connects sub-community 1 to sub-community 2)

Author identifier	Cut point (y/n)	Betweenness score	Closeness score	DCC
Author 1–2	y	0.0054	0.0851	1.00
Author 2a	n	0.0004	0.0835	1.00
Author 2b	n	0.0004	0.0851	1.00
Author 2–3	y	0.0105	0.1107	0.75
Author 3–4	y	0.0012	0.0681	0.75
Author 5–6	y	0.0002	0.0203	1.00

In addition to closeness and betweenness, we look for authors who are 'cut-points'; cut-point authors are those who connect two separate sub-communities together. The removal of these cut-point authors would result in two separate

14/24
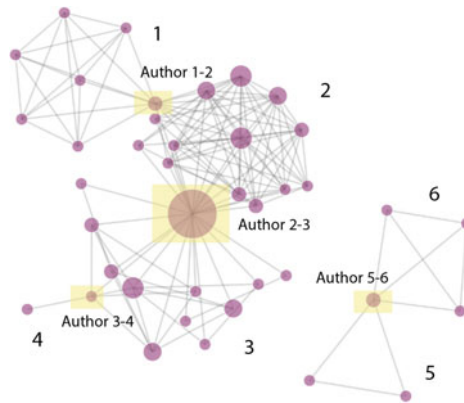
Figure 9. Cut-point authors in the HCD+D network. Each sub-community (Table 5) is labeled with a number.

communities, so we consider the cut-point authors to be influential connectors who help glue the overall network together. As seen in Figure 9, there are only four cut-point authors; one of these authors is highly cited with overall impact factors h-index = 69 and i10-index = 186.

In Table 5, we have gathered all influential authors with high betweenness, high closeness, or status as a cut-point author, and compare their centrality scores as well as their DCC.

The network average for authors' degree of cross-disciplinarity (DCC) is 0.56, and influential authors all score higher than the network average. This suggests that the influential authors in HCD+D tend to be more cross-disciplinary than the network as a whole. We do note that the general area of focus that these influential authors engage in is ICT (information, communication, technologies), which is a very cross-disciplinary area involving researchers in computer science, communication, design, mechanical engineering, and health sciences.

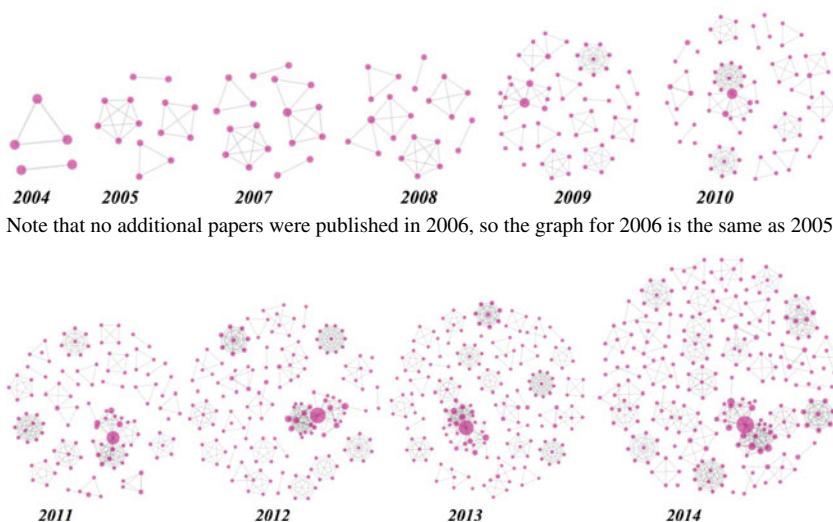## 5.2. Co-authorship networks over time

The papers examined in our dataset were published over the decade between 2004 and 2014. To understand how the co-authorship network has evolved over time, we graphed data from each year (Figure 10) and calculated structural and network metrics for each of these years (Table 6). Each year's co-authorship network includes authors who published in that year and in all preceding years. For example, 2007's co-authorship network consists of authors who published papers in 2004, 2005, 2006, and 2007. Note that no papers in our dataset were published in 2006.

The number of authors and papers increased between 2004 and 2014 because these metrics are all cumulative. From Table 6, we note that the average number of co-authors per paper has steadily increased from 2.5 to 3.85 by 2014, a sign that researchers became more willing to collaborate with one another over time. The largest component size increased from three to 34 over this decade, as has the diameter from one to four, meaning that the largest part of the network gained cohesion. A closer look at the authors within the largest component reveals that these authors are affiliated with the ICT sub-discipline; this may represent the

Table 6. Network and structural metrics over time

Metric	2004	2005	2007	2008	2009	2010	2011	2012	2013	2014
Total authors	5	14	18	22	60	79	118	159	190	247
Total papers	2	4	6	7	21	28	39	52	59	78
Avg. papers/author	1.00	1.00	1.06	1.05	1.13	1.15	1.19	1.28	1.25	1.22
Avg. authors/paper	2.50	3.50	3.17	3.29	3.24	3.25	3.62	3.92	4.03	3.85
Avg. co-authors/author	1.60	2.86	2.67	2.73	3.3	3.85	4.47	4.63	4.74	4.44
Density	0.400	0.220	0.160	0.130	0.056	0.049	0.038	0.029	0.025	0.018
Diameter	1	1	2	2	3	3	3	3	4	4
Avg. path length	1.00	1.00	1.08	1.07	1.09	1.08	1.06	1.04	1.05	1.04
Avg. component size	2.05	3.50	3.60	3.67	4.00	4.16	4.37	4.42	4.63	4.26
Largest component size	3	5	6	6	11	20	24	28	34	34
Avg. clustering coefficient	0.60	1.00	0.74	0.79	0.82	0.78	0.81	0.83	0.84	0.82
Network DCC	0.50	0.75	0.67	0.71	0.57	0.61	0.64	0.69	0.68	0.56

Note that no additional papers were published in 2006, so the graph for 2006 is the same as 2005.

Figure 10. Visual of HCD+D networks from 2004 to 2014.

higher maturation of the HCD+D community in the use of ICT in development and more associated avenues for publication within ICT for development work.

The density of the network has decreased steadily since 2004, beginning from 0.4 and reaching 0.018 by 2014. As the landscape of HCD+D research grows in size (adding new author nodes), fewer new edges (representing co-authorship) are added in comparison, leading to rapidly decreasing density.

Figure 10 below shows the co-author network for each year of study.

Note that no additional papers were published in 2006, so the graph for 2006 is the same as 2005.

6. Discussion

The present study examined the co-authorship network of 247 authors that published 78 papers in HCD+D during its foundational decade, 2004 to 2014. Analysis of the network using Python's NetworkX library showed that the community has a high clustering coefficient of 0.82. In the ASME DAC network (Guo *et al.* 2017), a similar design co-author network over a similar time period (2002–2015), the clustering coefficient increased from 0.63 to 0.79 (Guo *et al.* 2017); this is consistent with our network. However, our wide range of clustering coefficients (Figure 5) indicates that the average is driven by only a few well-connected authors who are more broadly affiliated with other universities and departments. For example, with a focus on ICT for development, the cut-point authors and those with high closeness and betweenness in Table 5 are all inter-related through their ties to the University of Washington or the University of California, Berkeley, or both.

The number of authors per paper in the ASME DAC network ranges from 2.20 to 2.67, while in the HCD+D network it ranges from 2.50 to 4.03 over time, signifying that researchers in HCD+D have tended to become much more collaborative in this field.

The HCD+D network has low density and only 1.8% of the total possible number of relationships were actualized by 2014. This density figure is higher than that of the Software Engineering and Knowledge Engineering co-authorship network (0.0059) (Kharboutly & Gokhale 2015), but much lower than that of the medical sciences co-authorship network (0.081) (Zare-Farashbandi *et al.* 2014).

Despite density that has steadily decreased from 2004 to 2014, from 0.40 to 0.018, the network's clustering coefficient remained steady near 0.80. Each year, new disjoint sub-communities of co-authors were added to the network, rather than new co-authorship links being added between old and new authors, which does little to increase the cohesiveness of the growing network. This may represent the addition of new disciplinary clusters of co-authors being added over time who are not aware of the work of previous authors in different fields.

Assessment of the visual network showed that the authors of HCD+D form 58 separate collaboration sub-communities. The largest component in the network comprised of 34 authors, though this component makes up only 14% of the overall network. This figure is much lower than the size of the largest components studied by Newman (2004), and Zare-Farashbandi *et al.* (2014), which made up 82 to 92 percent of the overall network. The HCD+D community in comparison is highly fragmented, with many pairs of authors not connected to the rest of the network. It more closely resembles the high energy physics co-author network and computer science co-author network, where the largest component makes up 1.1% and 1.9% of the entire network, respectively. A possible explanation is that HCD+D has a poor coverage of its subjects; it overlaps with many other traditional academic disciplines, and as a result some authors may consider HCD+D a sideline or small subset of their overall body of scholarship.

Typically, a network that has a large clustering coefficient coupled with a small average shortest path length (0.82 and 1.04) implies that the network has small-world properties, which is true for the ASME DAC network (Guo *et al.* 2017). In such a network, any random pair of authors are connected by a relatively short chain of acquaintances. However, in a highly fragmented network such as the HCD+D community, pairs of authors in disjoint communities would have difficulty reaching each other. This may be due to the cohesion and maturity of the network. We note that as a cohesive community, the DAC within ASME is more mature in that they have been holding separate conferences for over 40 years, while HCD+D is still an emerging field. The HCD+D scholarly community has only recently become a scholarly community, and the first HCD+D papers were published in 2004. HCD+D has room to mature and cohere as it continues to grow.

We also assessed the cross-disciplinarity of the papers and authors in our network. Author disciplines were noted for each paper, and we discovered that the average paper has researchers coming from 1.78 different disciplines.

Individual authors' performances in the network were also analyzed using our 'cut-point author' definition, and authors' betweenness and closeness centrality scores. We found that the vast majority of authors lie on the periphery of the network. Only a few prominent authors with high closeness are strategically placed at central positions within the network, from which they can disseminate information quickly throughout the network. Likewise, only five authors with high betweenness are able to serve as 'connectors' between other groups of individuals. Analyzing their degree of cross-disciplinarity shows that these authors

all score higher than the network DCC average of 0.56, indicating that influential authors are also better collaborators with authors outside of their own disciplines. Co-author networks are highly susceptible to fragmentation from the removal of individuals with the highest betweenness scores, and the HCD+D network would be no exception (Holme *et al.* 2002).

6.1. Limitations

The analysis presented in this paper used a relatively small dataset of 78 papers. The method employed to gathering papers that fit our search criteria involved inputting search terms in Google Scholar, and conducting a manual protocol analysis of papers. Our dataset is biased by the search term keywords 'human-centered design' which does not include other design related keywords, and excludes papers that might be doing design work without saying it explicitly. An implication of our narrow keyword search is that we are potentially missing papers where researchers are engaging in HCD+D work, but do not use the phrase 'human-centered design.' This is an explicit choice we made in conducting this analysis, and we consider this work to be our first pass at understanding HCD+D from those who say they engage in HCD+D. In the future, our methodology can be expanded to include other design terms, such as 'design thinking' or 'human–computer interaction.'

Assessing cross-disciplinarity is non-trivial, and in this study, we have presented several metrics and findings that show the state of cross-disciplinary collaboration in HCD+D. We have only looked at authors' departmental affiliations to monitor their disciplines, and we know that the department may not directly represent an author's expertise area.

We used co-authorships as a proxy of research collaboration, but recognize that collaboration in HCD+D is not an exclusively scholarly venture and, in fact, much HCD work is done outside of academia entirely. However, the purpose of this research is to better understand how HCD+D has evolved over time as a scholarly endeavor as represented by research publications.

7. Conclusions and future research

As we have presented in the Background and Motivation section, HCD+D must be cross-disciplinary in order to work effectively. Both design and development are cross-cutting fields and HCD+D must maintain this cross-cutting nature. The dispersion of the HCD+D network, and the lack of closely connected core, suggests that there is no singular guiding community of researchers all working together and strongly shaping the HCD+D research agenda.

There are very few authors who are cut-points, have high betweenness, or have high closeness, leaving the network susceptible to breaking down if these authors eventually leave the publishing scene (of great concern as one of the most influential authors in the network has recently died). We also see that the influential authors do have a higher degree of cross-disciplinarity than the network as a whole, which supports our notion that those with influence are also those who connect researchers working in different fields together.

In the future, we plan to look into ways to automate data collection, leading to a larger dataset and more expansive network analysis results. We also intend to look into the factors that affect how communities form within the HCD+D network,

such as geography or discipline, and discover ways to increase collaboration between researchers located in different locations across different disciplines. We might be able to assess the nature of collaboration by looking at how individuals collaborate over time, and seeing whether or not an individual's co-authors end up providing pathways to different authors and different knowledge areas.

Future research will track the evolution of HCD+D as it becomes embedded into the scholarship of more disciplines and in cross-disciplinary venues, such as the *Journal of Development Engineering* and conferences with special sessions on resource-sensitive, sustainable, or global development design (e.g., ASME's Engineering for Global Development Research Forum (ASME 2017) and the Design Society's (2017) focus on 'resource-sensitive design' at the International Conference on Engineering Design 2017). As more publication outlets with a specific focus on developing economies appear, we can continue the analysis presented in this paper to understand how the HCD+D network continues to evolve over time. As HCD+D is a relatively new design field, increasing its connections to scholarship in the larger design community has the potential of generally increasing the cross-disciplinary collaborations in both design and development practice. Increased cross-disciplinarity is an important vehicle for increased knowledge generation and synthesis across design, engineering, and development fields.

Acknowledgments

Appendix A

A.1. Network analysis terminology

We define some commonly used terms from network analysis. We draw upon Estrada and Knight's textbook (2015) in developing these definitions.

Node A fundamental unit from which graphs are formed.

Edge A connection (link) between two nodes in the graph.

Network (graph) A collection of nodes and edges.

Path A finite sequence of edges that connect a sequence of nodes distinct from one another.

Connected component A network is considered connected if there is a path connecting any two nodes in the network. Therefore, a connected component is a subset of nodes in the graph that are connected. Every pair of nodes in a connected component must have a path connecting them. In the case of an isolated node without any connecting edges, the node forms its own connected component of size one.

Degree The degree of a node measures the number of edges touching a particular node. In our network, the degree of a node measures the amount of authors an individual has collaborated with.

Clique Also known as a complete graph, a clique is a collection of vertices where any one vertex is connected to every other.

Density The ratio between the number of edges that exist and the maximum possible number of edges that can exist. Density is calculated by the following formula:

$$D = \frac{2 * E}{N * (N - 1)}$$

where E and N are the total number of edges and the total number of nodes, respectively.

Diameter The length of the shortest path between two farthest nodes in our network. The diameter provides a sense of how spread out the network is, and represents the resistance to the flow of information (Fuge *et al.* 2014).

Clustering coefficient Measures how well a node's neighbors are connected to one another. The formula for the clustering coefficient is given below:

$$C = \frac{N}{M}$$

N is the number of existing edges between the node's neighbors and M is the number of maximum such edges. A node will always have a clustering coefficient between 0 and 1.

Betweenness The proportion of shortest paths between all pairs of nodes that pass through a given author, which measures the given author's ability to control the flow of information in the network (Newman 2001*c*). Authors with high betweenness centrality are efficient at gaining and sharing resources from different groups (Kharboutly & Gokhale 2015). The formula for betweenness centrality is given by the following equation:

$$B(v) = \sum_{j,k \neq v} \frac{g(j, v, k)}{g(j, k)}$$

where $g(j, v, k)$ is the number of shortest paths from author j to author k that pass through author v, and $g(j, k)$ are total number of shortest paths from author j to author k.

Average path length The average number of edges in the shortest path between all possible pairs of connected nodes in the network. In a disconnected graph, we take the average path length within each connected component.

Small-world property Networks where the average shortest path distance between nodes increases proportionally to the number of nodes in the network:

$$L \propto \log N$$

while the clustering coefficient is not small. In these types of networks, any pair of nodes are linked by a small chain of edges.

'Farness' of a node The sum of its distances from all other nodes it is connected to.

Closeness of a node The reciprocal of farness. The formula is given by the following equation:

$$C(v) = \sum_{i \neq j} \frac{1}{d(i,j)}$$

where $d(i,j)$ is the distance between two authors i and j. (By convention, if i and j are not connected, closeness is set to be 0.) The more central a node is, the lower its total distance from all other nodes will be. Authors with high closeness are also network influencers, being at the most 'center' of the network, information originating from them propagates throughout the network the fastest (Newman 2001*c*).

References

ASME 2017 Engineering for Global Development Research Forum, International Design Engineering Technical Conference (IDETC2017). https://www.asme.org/events/idetc cie/program/egd-forum.

Barabasi, A. L., **Vicsek, T.** & **Palla, G.** 2007 Quantifying social group evolution. *Nature* **446**, 664–667.

Beaver, D. D. 2001 Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics* **52** (3), 365–377.

Brown, T. & **Rowe, P. G.** 2008 Design thinking. *Harvard Business Review* **86** (6), 252. doi:10.5437/08956308X5503003.

Brunette, W., **Gerard, W.**, **Hicks, M. A.**, **Hope, A.**, **Ishimitsu, M.**, **Prasad, P.**, **Anderson, R. E.**, **Borriello, G.**, **Kolko, B. E.** & **Nathan, R.** 2010 Portable antenatal ultrasound platform for village midwives. In *ACM DEV '10 Proceedings of the First ACM Symposium on Computing for Development*, 1(23).

Buchanan, R. 1992 Wicked problems in design thinking. *Design Issues* **8** (2), 5–21.

Clarke, A. 2015 Design for development, ICSID and UNIDO: the anthropological turn in 1970s design. *Journal of Design History*, epv029. http://doi.org/10.1093/jdh/epv029.

Design Society 2017 International Conference on Engineering Design with a theme of 'Resource-Sensitive Design'. http://iced17.org/.

Estrada, E. & **Knight, P. A.** 2015 *A First Course in Network Theory*, 1st edn. Oxford University Press.

Fuge, M., **Tee, K.**, **Agogino, A.** & **Maton, N.** 2014 Network analysis of collaborative design networks: a case study of OpenIDEO. *ASME Journal of Computing and Information Science in Engineering* **14** (2).

Fuge, M. & **Agogino, A.** 2015 Pattern analysis of IDEO's human-centered design methods in developing regions. *ASME Journal of Mechanical Design* **137** (7).

Gilliam, M. L., **Martins, S. L.** & **Bartlett, E.** et al. 2014 Development and testing of an iOS waiting room 'app' for contraceptive counseling in a Title X family planning clinic. *American Journal of Obstetrics and Gynecology* **211**, 481, e1–8.

Gordon, P., **Kramer, J.**, **Moore, G.**, **Yeung, W.** & **Agogino, A.** 2017 A systematic review of human-centered design for development in academic research. Working Paper 16-1201, *BEST Lab, UC Berkeley*, http://tinyurl.com/HCDDcharacterizationpaper.

Gov.UK n.d. Department for International Development. Retrieved January 1, 2016, from https://www.gov.uk/government/policies?organisations[]=department-forinternatio nal-development.

Guo, T., **Xu, J.**, **Sun, Y.**, **Dong, Y.**, **Davis, N. E.** & **Allison, J. T.** 2017 Network analysis of design automation literature. In *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.

Haythornthwaite, C. 1996 Social network analysis: an approach and technique for the study of information exchange. *Library & Information Science Research* **18** (4), 323–342.

Halasz, Z. 2015 Using clustering to further understand the design automation conference community. *University of Illinois at Urbana-Champaign*.

Holme, P., **Kim, B. J.**, **Yoon, C. N.** & **Han, S. K.** 2002 Attack vulnerability of complex networks. *Physical Review* E **65** (5), 056109.

IDEO 2009 Human-centered design toolkit. http://www.designkit.org/resources/1 (accessed 9/2017).

ISO (International Organization for Standardization) 2010 Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems, https://www.iso.org/standard/52075.html (accessed 9/2017).

Kharboutly, R. E. & **Gokhale, S. S.** 2015 Social analysis of the SEKE co-author network. In *Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering (SEKE)*, pp. 237–243. KSI Research Inc.

Krohn, W. & **Küppers, G.** 1990 Science as a self-organizing system outline of a theoretical model. In *Selforganization*, pp. 208–222. Springer.

Laudel, G. 2002 What do we measure by co-authorships? *Research Evaluation* **11** (1), 3–15.

Levine, D. I., **Agogino, A. M.** & **Lesniewski, M. A.** 2016 Design thinking in development engineering. *International Journal of Engineering Education* **32** (3B), 1396–1406.

Liu, J., **Li, Y.**, **Ruan, Z.**, **Fu, G.**, **Chen, X.**, **Sadiq, R.** & **Deng, Y.** 2015*a* A new method to construct co-author networks. *Physica* A **419**, 29–39.

Liu, X., **Bollen, J.**, **Nelson, M. L.** & **Van de Sompel, H.** 2015*b* Co-authorship networks in the digital library research community. *Information Processing and Management* **41**, 1462–1480.

Newman, M. E. J. 2004 Co-authorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America* **101** (1), 5200–5205.

Newman, M. E. 2001*a* Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review* E **64** (1), 016131.

Newman, M. E. 2001*b* Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review* E **64** (1), 016132.

Newman, M. E. J. 2001*c* Who is the best connected scientist? A study of scientific co-authorship networks. *Physical Review* E **64**, 337–370.

Palepu, K. G. & **Khanna, T.** 2010 Winning in emerging markets: a road map for strategy and execution. In *Harvard Business School Press Books*, 1. http://doi.org/10.1080/1051 712X.2012.666487.

Qiu, L. 1992 A study of interdisciplinary research collaboration. *Research Evaluation* **2** (3), 169–175.

Ramirez, M. Jr 2011 Designing with a social conscience: an emerging area in industrial design education and practice. In *Proceedings of the 18th International Conference on Engineering Design (ICED)*, pp. 39–48. Design Society.

Rittel, H. W. J. & **Webber, M. M.** 1973 Dilemmas in a general theory of planning dilemmas in a general theory of planning. *Policy Sci* **4** (2), 155–169.

Rowe, P. G. 1987 *Design Thinking*. MIT Press.

Scott, J. 2013 *Social Network Analysis*, 3rd edn, chap. 1. Sage Publications.

Sen, A. 1993 Capability and well-being. *Ethics* **1** (December), xi, 453. 58 http://doi.org/10.1093/0198287976.001.0001.

Simon, H. A. 1969 *The Sciences of the Artificial*.

Uddin, S., **Hossain, L.**, **Abbasi, A.** & **Rasmussen, K.** 2012 Trend and efficiency analysis of co-authorship network. *Scientometrics* **90**, 687–699.

United Nations 2000 Millennium Development Goals. Retrieved January 1, 2016, from http://www.un.org/millenniumgoals/.

United Nations 2015 Sustainable Development Goals. Retrieved January 1, 2016, from https://sustainabledevelopment.un.org/sdgs.

USAID n.d. What We Do. Retrieved January 1, 2016, from https://www.usaid.gov/what-we-do.

Yang, C. H. & **Heo, J.** 2014 Network analysis to evaluate cross-disciplinary research collaborations: the Human Sensing Research Center, Korea. *Science & Public Policy (SPP)* **41** (6).

Zare-Farashbandi, F., **Geraei, E.** & **Siamaki, S.** 2014 Study of co-authorship network of papers in the Journal of Research in Medical Sciences using social network analysis. *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences* **19** (1), 41–46.