The evolution of interdisciplinarity in five social sciences and humanities disciplines: relations to impact and disruptiveness

3 authors:


Hongyu Zhou
University of Antwerp


Raf Guns
University of Antwerp


Tim Engels
University of Antwerp

The evolution of interdisciplinarity in five social sciences and humanities disciplines: relations to impact and disruptiveness

Hongyu Zhou[1], Raf Guns[1], and Tim C. E. Engels[1]

[1] *Hongyu.Zhou@uantwerpen.be, Raf.Guns@uantwerpen.be, Tim.Engels@uantwerpen.be*

Centre for R&D Monitoring (ECOOM), Faculty of Social Sciences, University of Antwerp Middelheimlaan 1, 2020 Antwerp (Belgium)

Abstract

It is generally believed that the tide of interdisciplinarity is rising and becomes increasingly prevalent among various disciplines in natural and biomedical sciences. However, for the social sciences and humanities (SSH) limited evidence supports such a statement from bibliometric perspectives. Also, it has seldom been quantified how interdisciplinarity and its various aspects evolve over time. This paper analyzes the evolution of interdisciplinarity focusing on two aspects, namely process and outcomes, to draw a comprehensive trajectory of interdisciplinarity for five SSH disciplines over 50 years. We find that research in each of these five SSH disciplines is broadening its knowledge base by involving more disciplines yet is at the same time shifting towards further specialization. Interdisciplinarity is found to be positively correlated with citation impact and visibility and becomes stronger as the citation window widens up. The disruptiveness of publications, however, is negatively correlated with the level of interdisciplinarity and increasingly so over time.

Introduction

Interdisciplinary research (IDR) emerges to tackle the complex and societal-pressing problems that cannot be truly resolved by a single discipline (Carayol & Thi, 2005; Frodeman & Mitcham, 2016). To closely monitor and evaluate the supporting initiatives and understanding the status and mechanisms behind IDR, several recent studies are devoted to examining its different aspects (Rousseau et al., 2019), such as input (e.g. disciplinary diversity in team assembly; Schummer, 2004; Zhang et al., 2018), process (e.g. disciplinary diversity in references; Mugabushaka et al., 2016; Porter et al., 2007), outputs (e.g. topic diversity in abstracts; Bu et al., 2020), and outcomes (e.g. research impact; Larivière et al., 2015; Larivière & Gingras, 2010; Szell et al., 2018; J. Wang et al., 2015). Among all, quantitative measurements and indicators to evaluate the intensity of IDR processes, i.e., how interdisciplinary one's research is, is one of the most discussed focal research topics (Wagner et al., 2011). Although criticized as confusing and unsatisfying to achieve universally convergent assessment for precise decision-making (Q. Wang & Schneider, 2019), these indicators can still assist us to get a glimpse of IDR as a social phenomenon and to interpret it from a macro perspective.

One of the frequently referenced macroscopic statements by researchers, policy-makers, and the media is that "science is becoming more interdisciplinary". A consensus seems to have formed that scientists inhabiting dissimilar knowledge bases or mastering different skills have been crossing disciplinary borders and collaborating more frequently and with unconventional partners; this leads to more interdisciplinary and scientifically significant outcomes. Empirical

evidence, however, is still limited to disciplines from STEM and biomedical sciences. Temporal change in IDR in the social sciences and humanities has not been studied so far.
In this study, we analyse the evolution of interdisciplinarity in five SSH disciplines over half a century and examine the possible impact such change in IDR, if any, might produce on the outcome of scientific research. The rest of the paper is organized as follows: we first introduce the dataset and methodology we adopt. The next section presents the results and discussion and the last section concludes.

Data

The Microsoft Academic Graph (MAG) dataset is adopted in our empirical study; previous studies have shown that MAG is a viable source for scholarly communication research in terms of coverage (Paszcza, 2016) and the completeness of citation and metadata (Thelwall, 2017). As Microsoft continues to improve the coverage, design, and accessibility of MAG, it has become one of the most promising bibliographic datasets and is more frequently employed by those who study research dynamics quantitatively (Kong et al., 2020; Ma et al., 2020).

MAG distinguish from other bibliographic databases in that it adopts a bottom-up approach for the field categorization process (K. Wang et al., 2020), as opposed to, for instance, Web of Science, which uses existing journal categories to classify publications (called a top-down or classification-based approach, see Wagner et al., 2011). Different from previous bottom-up classification methods that tend to adopt co-citations, co-words, and/or bibliographic couplings (Wagner et al., 2011), MAG quantifies the semantic distance between two textual paragraphs representing two certain publications and then clusters the retrieved semantic representations to form the basis of concepts, which are de facto fields, domains, or disciplines in practice. Six levels of concepts are clustered automatically on different granularities. The top two levels of concepts (L0 and L1) are manually defined into a unique hierarchical structure to be consistent with most of the categorization systems (K. Wang et al., 2020), where L0 is comprised of 19 fields (e.g., physics, chemistry, and economics) and L1 consists of 294 subfields (e.g., theoretical physics, biochemistry, and macroeconomics).

The disciplines we use as case studies in this paper, i.e. Anthropology, Applied Psychology, Linguistics, Library Science, and Macroeconomics, are situated in the L1 level of MAG's category setting. Therefore, L1 is also used in the categorization of individual publications and their corresponding references to achieve consistency.

We recognized 300,559 journal or conference publications between 1960 and 2009 under the field category Anthropology (62,619), Applied Psychology (116,868), Linguistics (65,557), Library Science (17,678) and Macroeconomics (39,481). Publications labeled for more than one category were assigned to each category.

Method

We investigate the evolution of IDR focusing on two aspects, namely IDR process, and IDR outcome.

IDR Process

Stirling (2007) pointed out that diversity consists of three basic concepts, namely variety, balance, and disparity, each of which is a necessary but insufficient property of diversity as a whole. This notion and their generic indicator of diversity were then introduced and modified by Rafols and Meyer (2010) to Information Science as a quantitative measurement of knowledge integration to infer interdisciplinarity. A significant proportion of research is devoted to devising indicators that integrate two or three factors (dimensions) of diversity to achieve a reliable metric and assess or compare interdisciplinarity for different entities. In this study, we try to work as comprehensively and detailed as possible so that information loss caused by dimension reduction or integration can be minimized. Therefore, to quantify the intensity and evolution of IDR processes, we employed both single-factor (variety, balance, and disparity themselves; Stirling, 2007) and multi-factor measurements: Rao-Stirling (RS) diversity (Rafols & Meyer, 2010; Stirling, 2007), DIV (Leydesdorff et al., 2019), and ${}^{2}D^{s}$ (Zhang, Rousseau, & Glänzel, 2016) that involve two or all three single factors. We believe that the involvement of single-factor measurements may provide more implications that directly point to practical aspects of IDR, for instance, the number of disciplines referenced, and that the multi-factor measurements shed insight into the evolution of IDR from a more comprehensive perspective.

Table 1 provides notations and mathematical definitions of each indicator we employed in this study. Variety (n_c) is operationalized as the number of disciplines referenced for each publication, which reveals information regarding the **broadness of the knowledge base** in this study. Its variant, relative variety (n_c/N), is used in DIV representing variety in a relative scale; this variant is defined as variety divided by the number of categories in total. Balance (B), representing the **evenness of the knowledge base**, is set to be $1 - Gini_c$ where B = 1 indicates maximum evenness and B = 0 shows extreme imbalance. Here, $Gini_c$ represents the Gini coefficient of the distribution of disciplines in references. Disparity captures the average dissimilarity (or distance, explained further on) between every two disciplines referenced for each publication, which can be utilized to examine the cognitive distance and **heterogeneity of the knowledge base**.

Table 1. Selected measures of IDR.

Notation			
n_c	number of disciplines referenced	d_{ij}	dissimilarity between categories *i* and *j*
p_i	proportion of elements in category *i*	N	number of categories in total
x_i	number of references to the i-the category in an ascending order		
Indices			
Variety	n_c	Rao-Stirling (RS)	$\sum_{i,j} d_{ij}(p_i p_j)$
Balance (B)	$1 - Gini_c = \quad 1 - \frac{\sum(2i - n_c - 1)x_i}{\sum_{n_c} x_i}$	DIV	$(n_c/N) * B * D$
Disparity (D)	$\frac{\sum_{i \neq j} d_{ij}}{[n_c * (n_c - 1)]}$	${}^{2}D^{s}$	$1/(1 - RS)$

Three indicators integrating a part of or all the above-mentioned three factors are also employed (i.e., multi-factor indicators as aforementioned). DIV is the multiplication of relative variety, balance, and disparity ranging from zero to one. In terms of RS, this indicator calculates the sum of distances between every two disciplines referenced multiplied by the product of proportions each discipline accounts for in reference. ${}^{2}D^{s}$can be regarded as a variant of RS that possesses greater discriminatory power, satisfying the properties proposed in Leinster and Cobbold (2012). Like RS, this indicator employs similarity among categories instead of disparity directly.

Four out of six measures (i.e. disparity, RS, DIV, and ${}^{2}D^{s}$) employ the dissimilarity between two categories d_{ij} in their calculation, which is operationalized as (1 - similarity) in practice, as shown to be valid and efficient in Zhang, Rousseau, and Glänzel (2016). The temporal perspective of this paper makes a few modifications to the cosine similarity necessary, that is, the application of a time window on similarity calculation. As the distance or reference strength among disciplines may be changing over time (Frank et al., 2019), potential structural changes to the similarity matrix itself cannot be ignored when performing temporal analysis. To account for this, we construct ten similarity matrices with a five-year time window each. This yields the following equation:

$$d_{ij} = 1 - \frac{R_{ij}+R_{ji}}{\sqrt{\left(TC_i^t + TR_i\right)\left(TC_j^t + TR_j\right)}} \tag{1}$$

where i and j refer to two sets of publications from two different categories published during the period t, R_{ij} denotes the number of times set i publications cite set j publications, TC_i^t denotes the total number of citations set i publications received during the period t, and TR_i denotes the total number of references initiated by papers from set i.

The handling of multi-labeling in the categorization of publications is also a tricky issue when calculating balance, disparity, and RS as counting frequencies of categories for references is required. Upon our examination, more than 50% of the publications in our dataset are labeled with more than one category. To address this issue, we used fractional counting to quantify the relative frequencies. For each multi-labeled reference, we set the frequency of each category labeled to $1/m$ where m equals the number of unique categories associated with this reference and then sum all the category frequencies for each reference to form the category frequency for the publication. For example, a publication with two references A and B, where A is labeled with both “Linguistics” and “Literature”, and B is labeled with “Linguistics” and “Natural Language Processing (NLP)”, will have an overall category frequency distribution as follows: Linguistics 1; Literature 0.5; and NLP 0.5.

IDR outcome

Furthermore, the evolution of IDR outcome is investigated focusing on two aspects, namely **citation impact** and **disruptiveness**. 2-year, 5-year, and 10-year citations are calculated to reflect the academic significance and visibility over time. Disruptiveness, proposed in Funk & Owen-Smith (2017) and applied in Wu, Wang, and Evans (2019), captures the level of disruption a publication contributes by calculating the percentage of the net increase of citing papers it invites to an existing local citation network. As shown in equation (2), for a certain paper P, n_i denotes the number of its forward citations received during time t that did not cite any of P’s backward citations, n_j represents publications published during time t citing both P and its backward citations, and n_k captures publications that only cited P’s backward citations instead of P itself.

$$D_t = \left(n_i - n_j\right)/\left(n_i + n_j + n_k\right) \tag{2}$$

Results and discussion

The evolution of IDR process

Broadness of knowledge base

Figure 2a illustrates the evolution of variety for five disciplines in 50 years (1960-2009). Each subplot represents a discipline and each solid curve shows the probability distribution of variety for all publications in that discipline in the corresponding period. The vertical dashed lines denote the mean of variety for each period. The same color for a solid curve and a dashed line means that they are describing the same period.

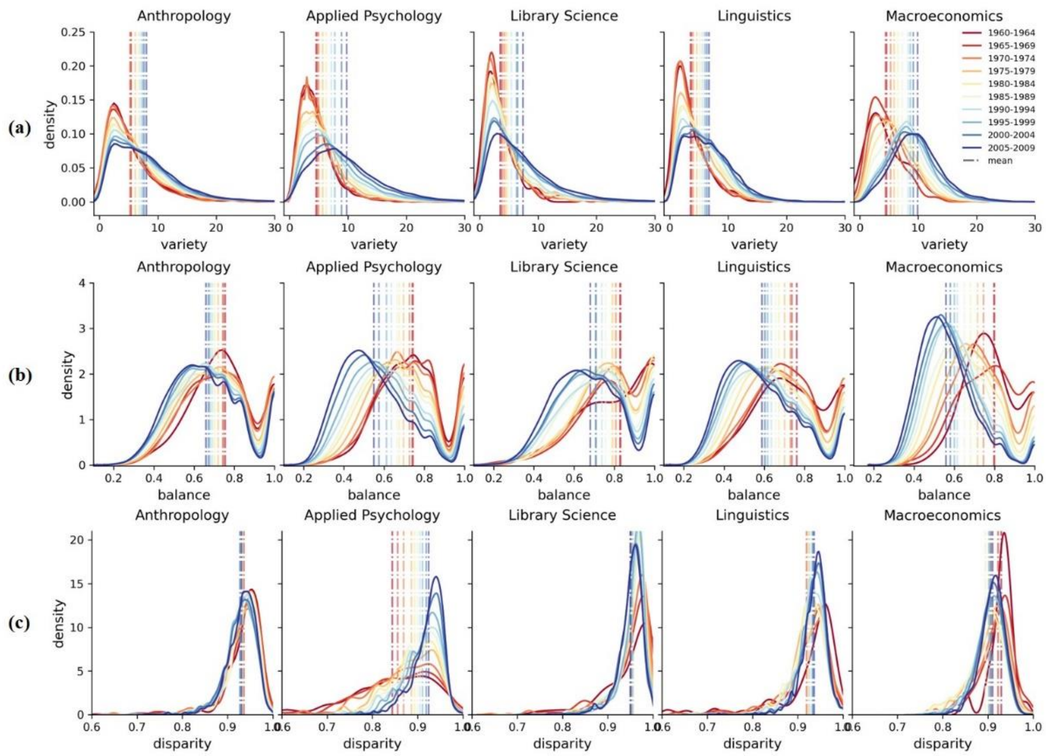

Figure 2. Evolution of variety, balance, and disparity in Anthropology, Applied Psychology, Library Science, Linguistics and Macroeconomics.

The dashed curves show sustained growth in terms of the central tendency, i.e. mean values, of variety for all five disciplines, but also illustrate a few discipline-wise differences. Applied Psychology and Macroeconomics, for instance, possess the biggest rise (from 5 to 10), which makes them the most interdisciplinary fields in terms of the broadness of the knowledge base. A rather moderate increase is associated with Linguistics and Anthropology. On the other hand, the growth of mean variety in Applied Psychology and Library Science seems to be accelerating as the gap between adjacent mean values (i.e., the distance between two adjacent vertical, dashed lines) keeps widening up over time. Such a phenomenon is missing or nebulous in the other disciplines.

The dominance of low variety publications has weakened over time as opposed to the rise of high variety publications. The rise of right tails is significant in all five disciplines which indicates an increasing percentage of high variety publications in each year. On the other hand, the right-shifted peak (mode) can also be spotted in Applied Psychology and Macroeconomics and illustrates that an obvious shift towards a diverse knowledge base for research in recent years.

The aforementioned findings capture the ever-increasing broadness of the knowledge base for research in SSH, more significantly in application-oriented disciplines. Furthermore, such multi-discipline-sourced studies have become the main force of science production. Clearly, researchers in SSH are more eager and proactive to absorb external knowledge or skills to advance their own study.

Evenness of knowledge base

In this paper, balance (B) is operationalized as $1 - Gini_c$ where B = 1 indicates the maximum evenness (e.g., an array [2,2,2,2] that has the same value for all elements has a value of balance equal to one) and B = 0 indicates the maximum imbalance. If the allocation of references to each category is evenly distributed, the authors have absorbed knowledge equally from various disciplines and treat each discipline as an equally significant part of the knowledge base to their study, hence indicating diversity and interdisciplinarity. On the contrary, if references are primarily devoted to one or a few disciplines, we can assume that certain specializations or sets of specializations take place in the study, yielding a low level of balance and less interdisciplinarity. A special and contradictory case occurs for papers that only have one category (discipline) in their references. This type of publications show a high specialization and low balance; yet, based on the formula of calculating B, its balance equals one which indicates maximum diversity. To handle this issue, in this paper, we exclude all papers with variety equal to one when examining balance. As such, 9.09% of publications for Linguistics, 4.08% for Applied Psychology, 6.84% for Anthropology, 9.62% for Library Science, and 1.57% for Macroeconomics are excluded.

As shown in Figure 2b, the distribution of balance exhibits multiple peaks and skewness in distribution. Publications with B =1 account for the right-most peak in all curves. The other peaks are associated with publications having a high balance in the earlier period and a low balance in recent years. It is also worth noticing that the skewness of the distribution also changes over time: In Linguistics, Applied Psychology, and Macroeconomics, the distribution is left-skewed in earlier years and shifted to right-skewed in recent years.

Contrary to variety, the mean of balance, as shown in vertical dashed lines, exhibits a decreasing trend for all disciplines. Macroeconomics achieved the largest drop of around 30.1% for 50 years, followed by Applied Psychology with a 26.0% decrease. The smallest drop can be found in Anthropology whose values for mean balance are reduced only by 12.6% during the same period. If we recall the results from the last section, we can see that disciplines which decreased the most in balance also achieved the biggest increase in terms of variety. One interpretation is that disciplines that tend to acquire knowledge from more peer disciplines do not always treat them as equally significant or relevant partners. Even though more disciplines are invited to their knowledge base, researchers from Applied Psychology and Macroeconomics are still heavily and unevenly reliant on a few disciplines.

The decreasing trend of balance is tightly related to the specialization of research (Foster et al., 2015). The formation of sub-disciplines or research topics might result in clusters of disciplines that are always regarded as the most significant knowledge base and more frequently and intensely referenced together thus yielding a dominant knowledge combination. Such tendency indicates that SSH researchers tend to have a clearer and more strategic agenda in terms of how to situate their research and how to learn from their peer scientists. Furthermore, we suppose that the increasing richness in knowledge base (i.e., variety) might as well be associated with the decrease in evenness (i.e., balance) to some extent. The occurrences of new disciplines (an increase of the value of variety) in the knowledge base might be naturally weak in intensity and proportions, which leads to an imbalanced knowledge base.

Heterogeneity of knowledge base

Disparity is operationalized as the mean distance among each pair of disciplines referenced. If only one discipline is referenced in a certain publication, its value of disparity is undefined as no "distance" can be defined or calculated. Therefore, we intentionally exclude this set of publications since there is no added information other than the decrease of the share of mono-source publications, which was already discussed in previous sub-sections

Figure 2c shows the evolution of disparity for our three selected disciplines, where one can see that the range of disparity looks similar, from ~0.60 to 1.00. Yet, the change of mean disparity varies fundamentally among all five. We observe an increase in Applied Psychology and Linguistics (9.6%, and 1.5%, respectively). Macroeconomics, on the other hand, has experienced a minor decrease in terms of mean disparity over the selected period. The other two disciplines, namely Anthropology and Library Science, fluctuate and remain at a similar level of disparity throughout the 50-year time window.

Besides changes in the mean values, we would also like to discuss the change in overall distributions of disparity, for which Applied Psychology exhibited more dramatic changes than the others. In the case of Applied Psychology, two clear and significant changes can be spotted, namely the formation of high peaks and the weakening in the left tail. Both indicate that researchers in Applied Psychology are referencing more "remote" or previously less connected disciplines to constitute their knowledge bases.

The decrease regarding the proportion of low variety publications illustrates a general tendency in Applied Psychology that researchers from this domain are aiming to voluntarily absorb and borrow knowledge or skills from more "remote" disciplines and the cognitive distance within their knowledge base continues to widen.

Integrated IDR trends

We calculated the evolution of integrated IDR using three indicators, namely DIV, ${}^{2}D^{s}$, and Rao-Stirling (RS) to investigate how the diversity of references as a whole evolves over time. The distribution over time for ${}^{2}D^{s}$ and five disciplines are shown in Figure 3, where dots represent mean values and bars indicate their 99% confidence intervals. Different periods are denoted as various colors in which darker colors represent more recent periods. Similar temporal trends are also found for RS and DIV.

It is obvious that publications from all studied disciplines in Social Sciences and Humanities are, at the aggregate level, becoming increasingly interdisciplinary over time, with a few distinctions for each. Linguistics was and continues to be the least interdisciplinary one among all five. Applied Psychology experienced the largest increase in the last 50 years. Although failing to gain as much growth as the others, Anthropology retains its leading position in interdisciplinarity. For the most recent 5 year period 2005–2009, Library Science is found to be the most interdisciplinary among all five.

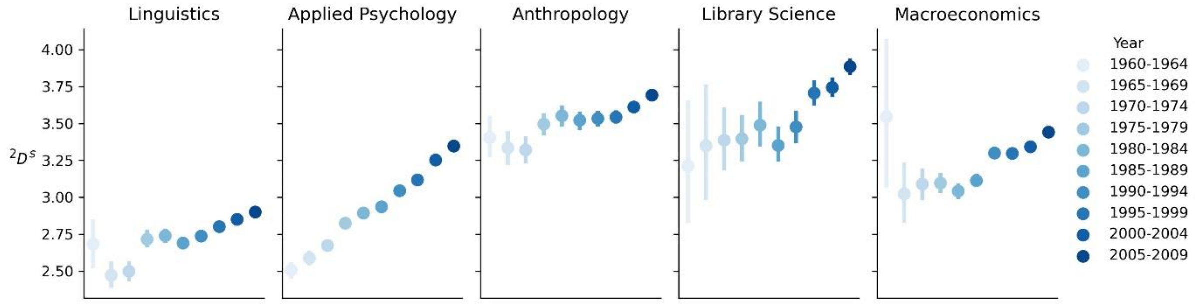

Figure 3. The evolution of integrated IDR measurements.

What should be additionally pointed out is that interdisciplinarity is not ever-increasing for all disciplines and all periods. In the mid-1960s, Linguistics and Anthropology exhibit a temporary drop in the level of interdisciplinarity on average, same for linguistics, Anthropology, and Library Science in 1985-1989. On the other hand, we can also observe certain synchronicity in change for certain periods, which suggests that although the developments in interdisciplinarity are realized by each individual discipline itself with their own pace or characteristics, they might be synchronously facilitated or hindered by a certain historical factor which could be academically related or otherwise.

The evolution of IDR outcome

Academic significance and visibility

To examine the potential impact of interdisciplinarity on academic significance and visibility, we determined citation impact with a 2- (C_2), 5- (C_5), and 10-year-long (C_{10}) citation window. The Spearman coefficient is calculated for each citation indicator and IDR indicator pair, as shown in Table 2. All correlations are statistically significant at the 0.1% level.

Table 2. Correlation between IDR measures, citation impact, and disruptiveness.

	Variety	*Balance*	*Disparity*	*RS*	*DIV*	$^2D^s$
C_2	.317	-.323	.094	.165	.229	.165
C_5	.395	-.406	.110	.209	.286	.209
C_{10}	.423	-.438	.117	.223	.306	.223
D_2	-.232	.277	-.048	-.092	-.147	-.092
D_5	-.276	.338	-.054	-.106	-.172	-.106
D_{10}	-.293	.363	-.055	-.109	-.180	-.109

Five out of six indicators turned out to be positively correlated with all three citation indicators, which indicates IDR's potential positive effect in terms of alleviating publications' academic significance and visibility. Balance, on the other hand, which measures the level of specialization in knowledge base, is negatively correlated with citation impact. This suggests that research may benefit from a more specialized perspective or knowledge base than focus-lacking ones. Variety is most positively correlated with citations among all.

Furthermore, the correlation becomes stronger as the citation window widens up which holds for all six IDR measurements. A possible interpretation of this observation is that the benefit interdisciplinarity might produce on citations requires time to form and may not be achieved overnight.

Disruption or consolidation

As explained in the "Methodology" section, the disruptiveness index (Wu et al., 2019) captures the proportion of net increase of new citing papers a publication invites to the local citation network. For instance, the publication will be regarded as disruptive with $D_t > 0$, if the number of its citing papers that also cite its references triumphs the number of citing papers that do not. In this case, the citing papers only acknowledge this paper's contribution rather than its intellectual forebears which can be a sign of disruption. The counter case would be publications with $D_t < 0$, which means more of their citing papers also acknowledge their forebears, yielding consolidation.

This index has many advantages. Firstly, it possesses the potential capacity of recognizing scientific works that are commonly regarded as innovative and break-through in the field. A large-scale empirical study devised by Wu et al. (2019) found that papers that directly contribute to Nobel prizes tend to show high levels of disruptiveness while review articles are normally associated with low disruptiveness. What's more, it can be more robust in resisting malicious citing behavior that is initiated by strategic considerations in one's career rather than the quality of the work.

The correlations between disruptiveness and all of the IDR measurements are shown in Table 2. D_2, D_5, and D_{10} denote disruptiveness within a 2-year, 5-year, and 10-year time window, respectively. All correlations are significant at the 0.1% level. It seems that all IDR indicators except balance are negatively correlated with disruptiveness, which suggests interdisciplinary publications tend to consolidate fields or knowledge rather than disrupt them. Another interpretation is that research whose knowledge base is more diverse might eventually lead to scientific outcomes that are more likely to back up or build on prior knowledge. On the other hand, as captured by the correlation between balance and disruptiveness, a more specialized knowledge base could be beneficial in the production of disruptive research.

Similar to correlation with citation, IDR also seems to be more correlated with disruptiveness in a longer citation window. This suggests that disruption or consolidation one might produce in certain research domains will become more evident as time passes by.

An interesting question about the relationship between disruptiveness and interdisciplinarity is why the out-of-the-box thinking that interdisciplinary research strives for leads to consolidation while the more specialized publications turn out to be more capable of disrupting. A possible interpretation is that interdisciplinary publications tend to investigate emerging topics that are possibly heading towards the formation of a newborn field. The exploratory work these IDR research conducted can be conceived as creative or unorthodox in their creation but as consolidating when the topic emerges to discipline after a while. On the other hand, specialized scientific research is normally diving into specific topics that already have a solid foundation of previous knowledge and skill set. Limited room for consolidation can be found which yields more possibilities for disruptive outcomes.

Conclusions

This paper has examined the evolution of IDR in five fields from Social Sciences and Humanities (SSH), namely Anthropology, Applied Psychology, Linguistics, Library Science, and Macroeconomics over a 50-year-long time window. The IDR processes, embodied by disciplinary diversity in references, and outcomes, characterized by citation impact and disruptiveness, are studied and analyzed from a temporal perspective. We find that research from SSH is absorbing knowledge from an increasing number of increasingly "remote" disciplines, which leads to growth in the overall level of interdisciplinarity. Yet, the increasing level of specialization observed in the knowledge base is also significant which could be influenced by the formation or emergence of specific research topics. On the other hand, most of the IDR measurements and citations are found to exhibit positive correlations, which are strengthened as the citation window widens. Most of the IDR measurements and disruptiveness, however, are negatively correlated which becomes more evident with a longer time window.

There are many interesting implications based on our findings. The increasing trend of interdisciplinarity found in SSH further strengthens the evidence of the statement "science is becoming interdisciplinary". This changing phenomenon calls for a necessity for us to reexamine the research paradigm in many disciplines and the evaluation system for various entities. Many research evaluations, for instance, peer review in journals, are conducted by specialized domain experts who might not have the relevant expertise for all interdisciplinary papers that integrates knowledge outside the discipline silos. The selection of a competent evaluation panel could be a somewhat challenging endeavor in the era of interdisciplinarity and should be handled seriously in a scientific way. Furthermore, we observe that the absolute value of the correlation between the number of citations and some interdisciplinarity indicators increases as the length of the citation time windows raises; and so does the degree of disruptiveness. This hints that interdisciplinary research needs time to show its potential, which is consistent with Wang et al., (2015). Previous studies have shown that interdisciplinary research is often encouraged by science policies, but they are insufficiently supported under current funding structures (Bromham et al., 2016). Thus, one implication for science policy decision-makers and funding providers is that assessing interdisciplinary research requires a longer time window.

Limitation and future work

This article has several limitations. For instance, only five disciplines in SSH are investigated which limits us to draw a comprehensive understanding for all SSH disciplines regarding the evolution of interdisciplinarity. In addition, the relationship between interdisciplinarity and research outcomes (citation and disruption) is analyzed in a simple way as a first step. A more sophisticated methodology that involves other confounding variables is required. In future studies, we will expand our current analysis to more disciplines in SSH and adopt advanced models to thoroughly uncover the interaction between interdisciplinarity and research outcomes.

Acknowledgments

This investigation has been made possible by the financial support of the Flemish Government to ECOOM, among others. The opinions in the paper are the authors' and not necessarily those of the government.

References

Bromham, L., Dinnage, R., & Hua, X. (2016). Interdisciplinary research has consistently lower funding success. *Nature*, *534*(7609), 684–687. https://doi.org/10.1038/nature18315

Bu, Y., Li, M., Gu, W., & Huang, W. (2020). Topic diversity: A discipline scheme-free diversity measurement for journals. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24433

Carayol, N., & Thi, T. U. N. (2005). Why do academic scientists engage in interdisciplinary research? *Research Evaluation*, *14*(1), 70–79. https://doi.org/10.3152/147154405781776355

Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and Innovation in Scientists' Research Strategies. *American Sociological Review*, *80*(5), 875–908. https://doi.org/10.1177/0003122415601618

Frank, M. R., Wang, D., Cebrian, M., & Rahwan, I. (2019). The evolution of citation graphs in artificial intelligence research. *Nature Machine Intelligence*, *1*(2), 79–85. https://doi.org/10.1038/s42256-019-0024-5

Frodeman, R., & Mitcham, C. (2016). New Directions in Interdisciplinarity: Broad, Deep, and Critical. *Bulletin of Science, Technology & Society*. https://doi.org/10.1177/0270467607308284

Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management Science*. https://doi.org/10.1287/mnsc.2015.2366

Kong, X., Zhang, J., Zhang, D., Bu, Y., Ding, Y., & Xia, F. (2020). The Gene of Scientific Success. *ACM Transactions on Knowledge Discovery from Data*, *14*(4), 1–19. https://doi.org/10.1145/3385530

Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, *61*(1), 126–131. https://doi.org/10.1002/asi.21226

Larivière, V., Haustein, S., & Börner, K. (2015). Long-Distance Interdisciplinarity Leads to Higher Scientific Impact. *PLOS ONE*, *10*(3), e0122565. https://doi.org/10.1371/journal.pone.0122565

Leinster, T., & Cobbold, C. (2012). Measuring diversity: The importance of species similarity. *Ecology*, *93*, 477–489. https://doi.org/10.2307/23143936

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2018.12.006

Ma, Y., Mukherjee, S., & Uzzi, B. (2020). Mentorship and protégé success in STEM fields. *Proceedings of the National Academy of Sciences*, *117*(25), 14077–14083. https://doi.org/10.1073/pnas.1915516117

Mugabushaka, A.-M., Kyriakou, A., & Papazoglou, T. (2016). Bibliometric indicators of interdisciplinarity: The potential of the Leinster–Cobbold diversity indices to study disciplinary diversity. *Scientometrics*, *107*(2), 593–607. https://doi.org/10.1007/s11192-016-1865-x

Paszcza, B. (2016). *Comparison of Microsoft Academic (Graph) with Web of Science, Scopus and Google Scholar*. https://doi.org/10.13140/RG.2.2.21858.94405

Porter, A. L., Cohen, A. S., David Roessner, J., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics*, *72*(1), 117–147. https://doi.org/10.1007/s11192-007-1700-5

Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, *82*(2), 263–287. https://doi.org/10.1007/s11192-009-0041-y

Rousseau, R., Zhang, L., & Hu, X. (2019). Knowledge Integration: Its Meaning and Measurement. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer Handbook of Science and Technology Indicators* (pp. 69–94). Springer International Publishing. https://doi.org/10.1007/978-3-030-02511-3_3

Schummer, J. (2004). Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics, 59*(3), 425–465. https://doi.org/10.1023/B:SCIE.0000018542.71314.38

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface*, *4*(15), 707–719. https://doi.org/10.1098/rsif.2007.0213

Szell, M., Ma, Y., & Sinatra, R. (2018). A Nobel opportunity for interdisciplinarity. *Nature Physics*, *14*(11), 1075–1078. https://doi.org/10.1038/s41567-018-0314-6

Thelwall, M. (2017). Microsoft Academic: A multidisciplinary comparison of citation counts with Scopus and Mendeley for 29 journals. *Journal of Informetrics*, *11*(4), 1201–1212. https://doi.org/10.1016/j.joi.2017.10.006

Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., Rafols, I., & Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2010.06.004

Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and Impact: Distinct Effects of Variety, Balance, and Disparity. *PLOS ONE*, *10*(5), e0127298. https://doi.org/10.1371/journal.pone.0127298

Wang, K., Shen, Z., Huang, C., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft Academic Graph: When experts are not enough. *Quantitative Science Studies*, *1*(1), 396–413. https://doi.org/10.1162/qss_a_00021

Wang, Q., & Schneider, J. W. (2019). Consistency and validity of interdisciplinarity measures. *Quantitative Science Studies*, *1*(1), 239–263. https://doi.org/10.1162/qss_a_00011

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9

Zhang, L., Rousseau, R., & Glänzel, W. (2016). Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. *Journal of the Association for Information Science and Technology*, *67*(5), 1257–1265. https://doi.org/10.1002/asi.23487

Zhang, L., Sun, B., Chinchilla-Rodríguez, Z., Chen, L., & Huang, Y. (2018). Interdisciplinarity and collaboration: On the relationship between disciplinary diversity in departmental affiliations and reference lists. *Scientometrics*, *117*(1), 271–291. https://doi.org/10.1007/s11192-018-2853-0