On the measurement of cognitive interdisciplinarity with OpenAlex's concepts

Giulio Giacomo Cantone[a], Venera Tomaselli[b], Paul Nightingale[a]

[a] Science Policy Research Unit, Business School, University of Sussex, UK
[b] Department of Economics and Business, University of Catania, IT.

1. Introduction

Scientific interdisciplinarity is intuitively understood as a property of scientific production. The more research is interdisciplinary, the more is able to bring together assumptions, perspectives, and results that belong to separate disciplinary traditions (Wagner et al., 2011). Under these premises, the potential benefits of fostering interdisciplinary research (IDR) have been widely discussed. Proponents argue that it stimulates innovation by encouraging the exploration of novel perspectives and facilitating breakthrough discoveries. Reservations against IDR revolve around the difficulty of conducting fair evaluations of IDR activities Seeber et al., 2022).

Empirical evidence on the outcomes of interdisciplinary research presents a mixed picture, with a (possibly misleading) prevalence of positive outcomes reported (Yegros-Yegros et al., 2015). Rafols et al. (2012), demonstrating that IDR is systemically undervalued and possibly underfunded. The intuitive argument is that scientific careers are defined by the choice to maximise the chances to get research funds. If research products too divergent from the core topics of famous scientific journals are poorly evaluated, the intellectual stimulation coming from IDR is actively impeded by the evaluation systems. Indeed, proponents of IDR also advocate for alternative models of research evaluation that do not penalise interdisciplinary forms of innovation (Wagner et al., 2011).

This study highlights novel possibilities to adopt bibliometric data provided by the open database OpenAlex (OpAl). Section 2 presents the state-of-the-art in indicators of interdisciplinarity. These indicators are canonically defined on data of citations, instead in Section 3 we will adapt them for the measurement of IDR through the 'scientific concepts' introduced by OpAl. Along the canonical measure for interdisciplinary *Diversity*, we introduce a novel indicator of *Difformity* as a divergence between the observed disciplinary profile of a published paper and the typical expectation from the journal where it is published.

2. How interdisciplinary is measured

Across the rich literature on the measurement of interdisciplinarity, it is possible to identify two main dichotomies for the operative definition of IDR activities. These are:

• Cognition *vs.* Organisation. It refers to the unit of analysis of IDR. The 'organisation' approach is interested in atypical combinations of authorships. The cognitive approach is much less interested in combinations of authorships, and much more in the transmission of concepts across papers and journals. This can happen directly, with the application of text mining techniques, or, as more frequently seen, comparing lists of references and citations. On a more abstract level, 'cognition' implies that scientific authors, who operate within disciplines, still 'recognise' the relevance or the usefulness of concepts lying outside the typical bounds of their own disciplines (Abramo et al., 2018).

- Integration *vs*. Diffusion. Integration can be defined as the capacity to combine a list of different inputs and arrange them in a coherent way. Commonly, these inputs are usually observed through analytical operations on the features of the reference lists of the papers. Diffusion refers to the capacity of a scientific product to be cited, mentioned, and replicated across different other papers, and it can be also recognised as a dimension of the scientific impact.

Although there have been attempts to provide a more complete operative definition of Integration (Rafols, 2014), this concept is often conflated with the measurement of diversity in the disciplinary profile of the paper. *Diversity* is paradigmatically defined as the interaction of three factors:

$$Variety \times Balance \times Disparity \tag{1}$$

Variety (a.k.a. ecological richness) implies that the authors relate their production to many disciplinary categories, independently of the difference of relevance given to these and to how these categories are similar to each other. Balance (a.k.a. ecological evenness) implies that the variance in the relevance of the categories is minimal. These two factors are usually measurable jointly (or not) with conventional indicators, e.g. Simpson index of repeat Rousseau (2018). The third factor of Disparity measures the similarity across the categories and it is commonly associated with the introduction of a matrix of similarity-dissimilarity in the equation, as in the Rao index of quadratic entropy (Stirling, 2007).

3. Measurement of integration in cognitive interdisciplinarity

We propose that in the bibliometric methodology exists a paradigmatic model of Normal Interdisciplinarity, which, referencing the aforementioned dichotomies is the degree of cognitive integration of a disciplinary profile. This model of measurement is not only the most frequent in literature, but it is also the most justifiable in practical terms. Typically, papers have only a few authors but many references, so reference lists have a higher size. The variation in the size of references across papers also has a finite variance, propriety not holding for lists of citing papers (Diffusion), because the number of citations grows over time.

According to Mugabushaka et al. (2016), indicators of interdisciplinarity are historically understood as a procession of 'generations'. Assuming a $X:\{i,j,\dots\}$ system of countable elements or numeric traits, the 1st generation of indicators of diversity consists of measures of entropy of first order (including Gini-type entropies). An accepted indicator of *Diversity* of '2nd generation' is the aforementioned Rao-Stirling index of quadratic entropy:

$$D_{RS}(X) = \sum_{(i,j)} [p(X,i) \cdot p(X,j)]^{\beta} \cdot d_{(i,j)}^{\alpha} \tag{2}$$

where p_i and p_j are couples of the relative frequencies of the elements or the normalised score (i.e. $\frac{\text{score}}{\sum \text{score}}$) of the trait, and $0 < d_{(i,j)} < 1$ is a value for the dissimilarity between i and j. Indeed, $d_{(i,j)}$ acts as prior about the expected value of $p(i) \cdot p(j)$, unconditional to X; low dissimilarity penalises the apportion of the couplet to $D_{R\square}(X)$. α and β are modelling parameters which are canonically set equal to 1 in parametrically naive measurement models[1].

---

[1] Notice that for $p(i) = p(j)$ when $d_{(i,j)} = 1$ (or, alternatively, $\alpha = 0$), Eq. 2 collapses into the Simpson index of repeat Rousseau (2018), also known as Hirschman index. Hence, the naive parameterisation of Rao-Stirling is considered the canonical 2nd generation index.

Eq. 2 has been criticised for two reasons:

1. It lacks the propriety referred in many ways: "trueness", "replication principle" or "composition principle". Trueness means that given n sets $X_1 \dots X_n$, such that their $D_{RS}(X)$ is equal for all of them and all of them have no elements in common, then the diversity of their union is n times $D_{RS}(X)$, i.e. $D_{RS}(\cup(X)) = n \cdot D_{RS}(X)$. As a corollary of the lack of "trueness", D_{RS} has a low discriminant power as a test statistic (e.g. to answer questions as "is X_1 significantly more diverse than X_2?") (Zhang et al., 2016).

2. The Rao index inherits propriety of Simpson's: non-monotonicity to balanced addition. This propriety implies that by adding any new non-empty category ($p_i \neq 0$), the index will differ, and this is even in the case of $p_i = \bar{p}_i$. This propriety contrasts with the monotonic behaviour of the Gini index based on the Lorenz curve [2]. According to authors who debated on this feature, the adoption of Rao-Stirling implies that Variety and Balance are not dual but a unique feature (Leydesdorff et al., 2019; Rousseau, 2019).

Authors such as Leydesdorff et al. (2019) and Rousseau (2019) proposed to adopt a different index based on Gini's index to solve the second issue. On the contrary, Zhang et al. (2016), inspired by the systematisation of the theory of Hill-type measures in Leinster and Cobbold (2012) noticed that adopting the transformation

$$D_{RS}^{(T)} = \frac{1}{1 - D_{RS}} \tag{3}$$

the first issue is solved (Mugabushaka et al., 2016). These advancements constitute the 3rd generation of indexes of *Diversity*. In the following application, we will propose how to expand this framework to welcome new variables of IDR present in OpAl's database.

3. Application

3.1 Sampling Frame

Let a paper be symbolised by X, a journal as K, and an author as A. Papers have many authors, so the fractional contribution to the authorship of an author is $a_A(X) = \frac{1}{N(A_X)}$.

We queried the OpAl catalog (Aria and Le, 2023) for papers published in the years 2018, 2019, 2021, and 2022 in 939 Class A journals in the official list for Disciplinary Area 13 of the Italian National Agency of Evaluation of the University and Research System (ANVUR). These are considered by ANVUR the most relevant international journals for Economics, Business, Management, Finance Statistics, and Demography.

We fetched 31,632 papers with at least one author classified as affiliated with an Italian university. We sampled 64 research units (U) as departments of Economics, Management, Statistics, Business or Finance[3]. Only 7,280 papers have been authored by at least one author affiliated with these 64 research units.

---

[2] In statistical software R, `abdiv::simpson(c(1,1))` != `abdiv::simpson(c(1,1,1))` holds, while `DescTools::Gini(c(1,1))` != `DescTools::Gini(c(1,1,1))` it does not hold.

[3] Two departments from the same university would be counted as different research units. This is a representative and sature sample of departments of Area 13 - ANVUR in Italy. Only small departments have been ecluded from the sample. In some cases (e.g. Bocconi Schools, or University of Calabria) it was convenient to consider an aggregation of departments as a research unit, instead.

In this context, a research unit (U) consists of a list of authors. The sum of their authorship can roughly estimate the nominal total research output of the research unit, i.e. $y_U = \sum a_A(X) \mid A \in U$.

3.2 Methods

In OpAl each unitary record (an X journal, a K journal, etc.) is related to a 'concept' through a c score. Concepts are labels associated with scientific activities and are classified through different levels. At level 1, they work as disciplinary labels. In the context of OpAl, it is convenient to treat disciplinary labels as observable traits (i) of the record.

Even concepts have a score that relates to each other, so for the determination of the aforementioned matrix of similarities we adopted the normalised scores of similarity provided by OpAl:

$$d_{(i,j)} = \begin{cases} 0; & i = j \\ \dfrac{\delta_{ij}}{\max(\delta_{ij})}, \delta_{ij} = \dfrac{c(i,j)}{\sum c(i,j)}; & i \neq j \end{cases} \tag{4}$$

where $c(i,j)$ is the similarity score that OpAl assigns to disciplinary labels i and j.

Each u research unit is evaluated through its average score of *Diversity* and *Difformity*. The second is a dimension still unexplored in literature, which is a measure of how unexpected is the disciplinary profile of a paper within the context of the journal where it is published.

3.2.1 Measurement of Diversity through OpenAlex's concepts

The estimator $p_i = \dfrac{c_{X,i}}{\sum c_{X,i}}$ is considered for the determination of the *Diversity* (Eq. 3) of a paper. The proposed method has the advantage of requiring much less information and computation over the canonical alternative of formulating p_i as the disciplinary proportion of the reference list (Zhang et al., 2018).

Since $D_{RS}^{(T)}$ has the propriety of trueness (or, 'composition'), its linear aggregation is not biased. In this case, in order to balance the difference in research outputs across research units (y_U), we are interested in the average

$$D_{RS}^{(T)}(\bar{U}) = \frac{\sum_{A \in U} D_{RS}^{(T)}(X_A) \cdot a_A(X)}{\sum_{A \in U} a_A(X)} \tag{5}$$

3.2.2 Measurement of Difformity through OpenAlex's concepts

Difformity is the dissimilarity of the X paper from the expectation of the archetypical $\langle X_K \rangle$ paper published in its K_X-journal.

Let the normalised absolute divergence between the portions of disciplinary scores observed by OpAl between X and K_X be

$$\phi(X,i) = \frac{|p(X,i) - p(K_X,i)|}{\sum |p(X,i) - p(K_X,i)|} \tag{6}$$

then the 'true' estimator for *Difformity* ($\Phi(X)$) can be derivated by Eq. 2 and Eq. 3

$$\Phi^{(T)}(X) = \frac{1}{1 - \sum_{(i,j)} \phi(X,i) \cdot \phi(X,j) \cdot d_{(i,j)}} \tag{7}$$

and the estimator for the average *Difformity* of the research unit can be derived by Eq. 5.

4. Results

Figure 1 shows the distribution of the 64 research units across the two dimensions and two-time windows. Papers published in the years 2018 and 2019 are considered a sample of scientific production just *before* COVID-19, while those in 2021 and 2022, *after* it.
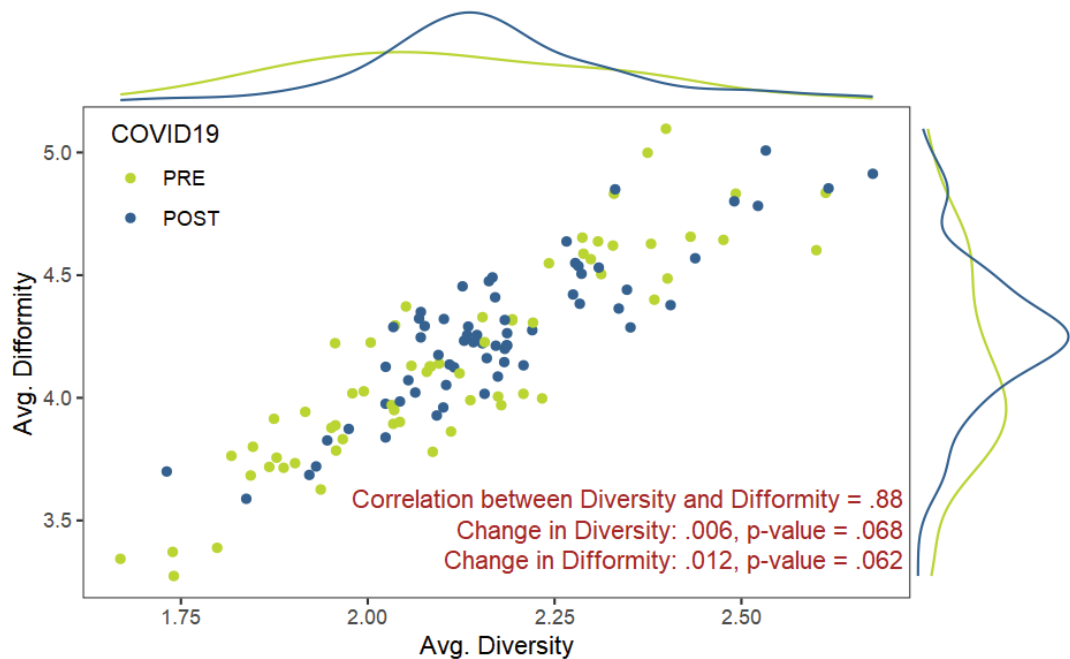

Figure 1. Average *Diversity* and *Difformity* before and after 2020 in 64 research units.

Maybe after COVID-19 research units focused more scientific production on a diversified research agenda, possibly as feedback from high-quality journals accepting different papers from their own disciplinary canon. However, evidence is still not sufficient to claim that Italian research in high-quality journals became on average more diversified after COVID-19.

There is a significant reduction in the dispersion of *Diversity* (F-Levene: 5.69, p-value = $.019$) and *Difformity* (F-Levene: 8.25, p-value = $.005$). Another observed change is the following: before COVID-19 no significant correlation is observed between research output (y_U) and *Diversity*. However, in years after COVID-19 there is an ambiguous positive correlation between the two (Pearson corr. coefficient: $.23$, p-value = $.06$). Considering that y_U depends on the size of the departments, even the increase in average research output before and after 2020 (6.95 more authorships, p-value = $.051$) is ambiguous, too.

These results should be calibrated with the credibility of OpAl's assessment of discipline. In Figure 1, *Difformity* is usually around double of *Diversity*. This is by-effect of the addition of *Richness* of the journal's discipline in Eq. 7. In other words: in many cases, disciplinary labels diverge between journal and paper. This could be a signal that OpAl's level 1 concepts are not an accurate representation of a disciplinary reality in scientific papers. A suggested development is to consider the taxonomy at level 0 to improve robustness of results.

References

Abramo, G., D'Angelo, C.A., Zhang, L. (2018). A comparison of two approaches for measuring interdisciplinary research output: The disciplinary of authors vs the disciplinary diversity of the reference list, *Journal of Informetrics*, **12**(4), pp. 1182–1193.

Aria, M., Le, T. (2023). *OpenAlexR: Getting Bibliographic Records from 'OpenAlex' Database Using 'DSL' API*.

Leinster, T., Cobbold, C. A. (2012). Measuring diversity: the importance of species similarity, *Ecology*, **93**(3), pp. 477–489.

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*, **13**(1), pp. 255–269.

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), pp. 593–607.

Rafols, I. (2014). Knowledge integration and diffusion: Measures and mapping of diversity and coherence. In Y. Ding, R. Rousseau, D. Wolfram (eds.). *Measuring Scholarly Impact: Methods and Practice*, pp. 169–190. Springer International Publishing, Cham.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management, *Research Policy, Elsevier*, **41**(7), pp. 1262–1282.

Rousseau, R. (2018). The repeat rate: from Hirschman to Stirling, *Scientometrics*, **116**(1), pp. 645–653.

Rousseau, R. (2019). On the Leydesdorff-Wagner-Bornmann proposal for diversity measurement, *Journal of Informetrics*, **13**(3), pp. 906–907.

Seeber, M., Vlegels, J., Cattaneo, M. (2022). Conditions that do or do not disadvantage interdisciplinary research proposals in project evaluation, *Journal of the Association for Information Science and Technology*, **73**(8), pp. 1106–1126.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, **4**, pp. 707–719.

Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T., Boyack, K.W., Keyton, J., Rafols, I., Borner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature, *Journal of Informetrics*, **5**(1), pp. 14–26.

Yegros-Yegros, A., Rafols, I., D'Este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity, *Plos One*, **10**(8): e0135095. Publisher: Public Library of Science.

Zhang, L., Rousseau, R., Glanzel, 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), pp. 1257–1265.

Zhang, L., Sun, B., Chinchilla-Rodriguez, Z., Chen, L., Huang, Y. (2018). Interdisciplinarity and collaboration: on the relationship between disciplinary diversity in departmental affiliations and reference lists, *Scientometrics*, **117**(1), pp. 271–291.