Liquid Science: How knowledge flows in scientific networks

Simone Belli
Faculty of Political Sciences and Sociology, Complutense University of Madrid, Spain
Yachay Tech, Ecuador

Ernesto Ponsot
School of Mathematical and Computational Sciences, Yachay Tech, Ecuador

Abstract

Nowadays our lives have introduced digital practices in daily activities. This change is also reproduced in professional settings, like scientific ones. Research practices have been modified from this transformation, from communication with other researchers to data analysis. In every discipline, we observed the so-called "digital turn" or the digital transformation in scientific collaboration. This digital transformation represents a new strategy for sharing knowledge between researchers and opening new possibilities for international collaborations.

A survey was designed and applied to researchers within the framework of the H2020 project XXXX. The researchers consulted embrace different disciplines and are located in different universities and research institutes. There are 305 different interviews and 159 variables or observed responses. In order to complement and correctly interpret the information gathered with the survey, selected researchers were interviewed, and some of their opinions are also reported in this work.

The research explores how information and communication technologies in this specific professional setting have affected the way in which research teams are related. In particular, this investigation explores the extent to which scientific networks are composed and how they collaborate thanks to the digital transformation, observing the ways of communication and collaboration of research teams, and the number of their scientific publications.

Keywords: scientific collaborations, digital transformation, research groups, scientific publications, scientific networks.

Introduction

Nowadays our lives have introduced digital practices in daily activities. This change is also reproduced in professional settings, like scientific ones. Research practices have been modified from this transformation, from communication with other researchers to data analysis. In every scientific discipline, we observe the so-called "digital turn". This term identifies a series of new professional practices to access and share knowledge between researchers. Digitalization has opened up new possibilities for researchers in their handling of information and knowledge.

Unlike previous generations, researchers today have a seemingly endless variety of potential scientific collaborations and research networks available through digital tools. Thanks to the Internet, less time and efforts are required to begin and manage a scientific collaboration. Recent studies have shown that Internet-mediated collaboration is generally thought to increase scientific productivity (Duque et al., 2009; Olson et al., 2008).

Implicit in Bauman's ideas of 'liquid modernity' is the suggestion that professional activities are being eroded by the proliferation of extensive 'networks' of professional possibility. Drawing a parallel with the concept of 'liquid love' (Bauman, 2003; Hobbs et al., 2017), the concept of 'liquid science' can also be presented, arguing that is has transformed scientific activities into a type of practices where researchers can do research with colleagues from different parts of the World.

This article explores how information and communication technologies in this specific professional setting have affected the way research teams are related. In particular, this investigation explores the extent to which scientific networks are composed and how they obtain and manage information thanks to digital transformation. Observing the ways of communication and collaboration of research teams, information search and access to the knowledge, and the major trends and changes they have perceived in the last years.

In this paper, we are interested in identifying which are the digital mechanisms that account for success in scientific collaborations and the principles of co-evolution of the digital infrastructures in scientific communities. For this purpose, a survey was designed and applied to researchers within the framework of the XXXX project. The researchers consulted embrace different disciplines and are located in different countries. The objective was to study how the new information and communication technologies have affected the way research groups are related. The first response was recorded on 12/16/2016 and the last on 2/13/2018. Members belong to both European and American countries. There are 305 different interviews and 159 variables or observed responses. As to date there is limited research specifically on digital tools, this study aims to be an exploratory investigation that identifies the various affordances and transformations provided by the technologies, with the intent of also highlighting areas in need of further research.

These are the two hypotheses that have guided the construction of this research:

- *Hypothesis 1.* Scientists working in small and local research teams tend to use analogical tools and to have less scientific collaborations/scientific publications.
- *Hypothesis 2.* Scientists working in big and international research teams tend to use online tools and to have more scientific collaborations/scientific publications.

To test these hypotheses, we will observe the size of the networks of the researchers, its structure (local or global), the frequency of interactions and the modality (analogical or digital tool) in relation to the number of scientific publications.

What follows is a brief review of the existing literature and the study's methodology, and then a more in-depth exploration of emerging patterns of usage and their professional consequences.

Literature Review

Several bodies of literature inform this investigation. The first is the sociological research on digital transformation of intimacy by Hobbs et al (2017) that has inspired this paper. Authors explore the experience of users of digital dating and hook-up apps in order to assess the extent to which a digital transformation of intimacy might be under way. Thanks to this study, it is possible to understand how networks in other contexts, for example professional ones, are influenced by digital technologies that modify the relations between their participants. As noted earlier, drawing a parallel with the concept of "liquid love", Bauman (2003) argues that virtual relationships are increasingly supplanting more fixed and inert 'real' relationships, and that the widespread usage of mediated communication is leading individuals to think more of transient connections than life-long partnership.

The second point is the research on digital transformation conducted in the project XXXX. We have found that scientific networks are expanded through the use of digital technology, leading to an increase in the number of scientific collaborations and scientific production (Author XXXX). Authors XXXX analyze the use of digital technologies of the academic community in Latin America, its functions, limitations, and potentialities. This, with the purpose of building a common framework that allows for the mediation of technologies in the production of knowledge and the development of equitable collaborative networks. For example, to encourage a common open access policy for the Latin America.

The third point of this review section is the change of space where research is conducted; researchers work in online and offline space at the same time. This mixed infrastructure is also called “e-Infrastructure” or digital infrastructure and it refers to research environment in which all researchers have shared access to unique or distributed scientific facilities (data, instruments, computing and communications), regardless of their type and location in the world (Candela, Castelli, and Pagano, 2011; Corti and Fielding, 2016). Renaud (2000) defines infrastructure as the collective structures that enable, enhance, embody and structure research. This concerns network infrastructure, where hardware and software platforms and tools and applications provide creative technology environments. Digital transformation enables people to reduce barriers of location, time, institution, and discipline (Atkins et al., 2003). For Heimeriks, Horlesbergerm and Van der Besselaar (2003), the emergence of digital information, online open database, and ICT have enabled a radical lowering of the costs related to collaboration, communication, and information dissemination within the science system and between knowledge producers and users, with new patterns of communication and collaboration emerged.

The fourth block is composed of effective communication. Leonard, Graham and Bonacum (2004), Manser (2009) and Lingard et al. (2004) showed the importance of effective communication in teamwork to provide appropriate health care. Other studies such as Sexton, Thomas and Helmreich (2000) also examined the repercussions of stress, errors and teamwork in different contexts. In Author XXXX, we have recognized the technical problems occurring in online meeting, through the negative emotions expressed by participants. In this range of negative emotions caused by slow internet connection and technical difficulties, there is focus on stress, concern and social unrest. These negative emotions are causes of affective conflicts in meetings.

These four points introduce us to the concept of liquid science based on the new forms to share knowledge between researchers (and society) as a relevant indicator for the emerging patterns of cooperation, composed by actors involved (nodes), structure of the network (relations), and knowledge (content), based on the model of Heimeriks, Horlesbergerm, and Van der Besselaar (2003). Liquid science is this conjunction of communication, knowledge, and network, and it has increased collaborations between institutions and disciplines.

Methodology

Design and measurements of the Survey

This is a quantitative research consisting of an online survey sent to a multi-stage random sample of European and Latin America and Caribbean university researchers in all disciplines via email from the accounts of the members of the XXXX project (10 scientific institution in Latin American and Caribbean countries, and 9 in European Union) to their scientific networks connection and institutional colleagues.

The invitation was then subsequently shared via email and social networks (Twitter, Facebook, LinkedIn, ResearchGate, Academia.edu) by willing network connections in a 'snowballing' fashion (Hobbs et al., 2017). While the 'snowball method' can have epistemological limitations with regards to generating statistically significant representative samples, the research method is nevertheless capable of collecting data indicative of broader social patterns and trends, especially when the survey reaches a broad cohort of participants (Atkinson and Flint, 2003; Denscombe, 2010; Hobbs et al., 2017; Neuman, 2011).

The study population contains researchers, including Ph.D. students and postdoctoral researchers. The questionnaire contained items relating to digital transformation in science, detailed questions about digital communication and scientific production, in addition to questions about their research activities.

In terms of statistics, the methodology includes a combination of descriptive and inferential analyses based on the data collected. For the inferences, the Poisson regression model (Agresti, 2015) is used, in its variant that corrects the over dispersion present in the data, quasi Poisson. This model, a member of the family of generalized linear models (Nelder & Wedderburn, 1972), has a clear application when the response variable is the product of case counting and seeks to determine rates associated with a period of time (here, number of publications per year). All data processing was done in R Statistical Software (R Core Team, 2018).

The objective was to study how the new information and communication technologies have affected the way in which work teams are related through 159 variables or observed responses. For the current analysis, the variable under considerations are in Table 1. The survey consisted of a combination of open-ended, multiple-choice and Likert-scale questions and took approximately 15-20 minutes to complete.

Table 1. Survey variables under consideration for this analysis.

Code	Question	Variable
P004	Country:	Country
P006	Gender:	Gender
P007	Age:	Age
P008	Position (level)	Position
P009	Type of Institution	InsType
P010	Discipline:	Discipline
P012	Do you primarily work alone or as part of a research/project team?	AloneTeam
P013	How many people are there in these teams (typically)?	TeamSize
P014	Where are team members located? [Same institution]	LocSameIns
P015	Where are team members located? [Different institution]	LocDiffIns
P016	Where are team members located? [In other Countries]	LocDiffCou
P017	How often do team members interact (typically)?	FreqInt
P018	How do team members keep in touch? Could you indicate whether the following are 'Essential', 'Used' but not essential, or 'Not Used': [Letters]	ContactLetter
P019	How do team members keep in touch? Could you indicate whether the following are 'Essential', 'Used' but not essential, or 'Not Used': [Phone calls]	ContactPhone
P020	How do team members keep in touch? Could you indicate whether the following are 'Essential', 'Used' but not essential, or 'Not Used': [E-mails]	ContactEmail
P021	How do team members keep in touch? Could you indicate whether the following are	ContactList

Code	Question	Variable
	'Essential', 'Used' but not essential, or 'Not Used': [Discussion mailing lists]	
P022	How do team members keep in touch? Could you indicate whether the following are 'Essential', 'Used' but not essential, or 'Not Used': [Instant messenger (WhatsApp, Line, etc.)]	ContactInsMes
P023	How do team members keep in touch? Could you indicate whether the following are 'Essential', 'Used' but not essential, or 'Not Used': [Videoconferencing (Skype, GTalk, etc.)]	ContactVideoC
P093	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Journal Papers (not-Open Access)]	pubTotal, pubScience
P094	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Journal Papers (Open Access)]	pubTotal, pubScience
P095	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Books (Written or edited)]	pubTotal, pubScience
P096	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Book Chapters]	pubTotal, pubScience
P097	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Reports]	pubTotal, pubScience
P098	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Conference papers]	pubTotal, pubScience
P099	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Non-refereed articles]	pubTotal
P100	How often do you publish? Could you indicate the approximate average number of publications per year of the following types? [Blog posts]	pubTotal

As can be seen in Table 1, only a subset of the variables collected in the survey are used for current analysis. There are three variables that are constructed from others:

- PubTotal: contains the total number of publications reported in the survey.
- PubScience: contains only the number of publications considered as scientific.
- Loc: Contains the grouping of locations of reported team members.

Human ethics

It is responsibility of the XXXX consortium and of the Vice-Chancellor of Research of XXXX University that no harm whatsoever occurs to individuals by virtue of their participation in project activities and data collection processes. These two bodies/entities gave approval to the project in September 2016. In overall terms, the project complies with the following ethical standards for research in social sciences and not involving greater than minimal risks for the subjects:

- The research is designed and applied ensuring transparency and integrity
- The participants in surveys will be informed about the project objectives, methods and intended use of the collected data.
- The information gathered with the participation of human subjects will be treated

confidentially and the anonymity of respondents will be respected.

- The participants will take part in the research voluntarily.

Regular project meetings and other checks by project leaders and the coordinator shall guarantee the compliance with ethical standards before publicizing any research findings. Finally, the XXXX project will make an effort to pursue the integrative concept of Responsible Research and Innovation (RRI) and of Protection of Human Subjects (PHS).

Data Collection

The first response to the survey was recorded on 12/16/2016 and the last on 2/13/2018. Members belong to both European countries and American countries, for a total of 18 countries. The survey had a total of 305 respondents, of whom most, but not all, answered all questions under consideration. In addition, for this research, the responses of those interviewed who indicated that they worked alone were not considered, only those who indicated that they were working as a team or both. In terms of gender, 4 people preferred not to inform (1%), 101 are female (34%) and 200 are male (66%). This leaves a total of 264 interviews to consider. Figure 1 shows the demographic pyramid of the surveyed.

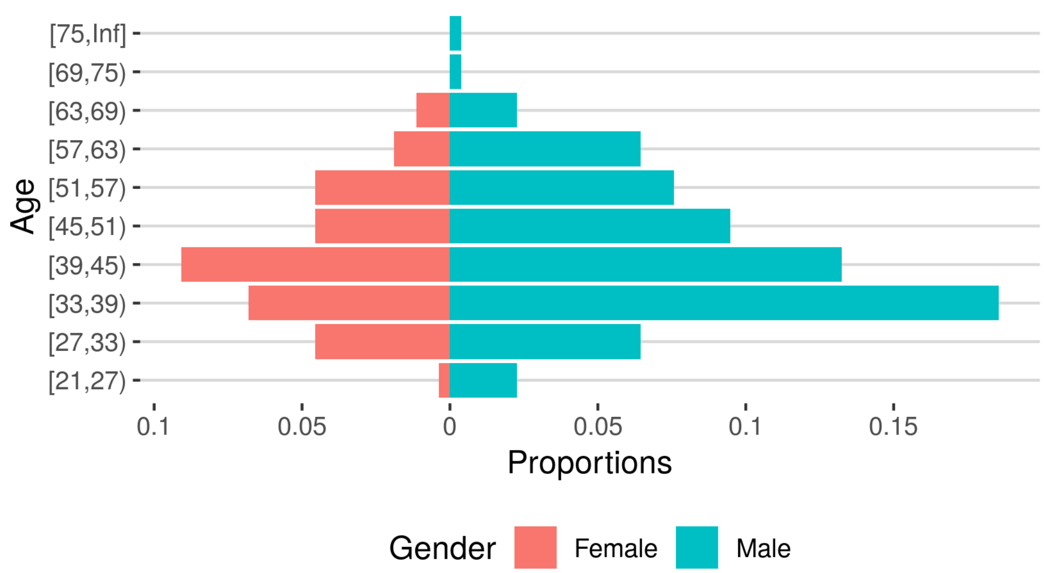

Figure 1. Demographic pyramid of the sample.

The figure shows that those responding to the survey are mostly male, with ages between 33 and 45 years old. Moreover, female scientists face ‘patrifocal’ constraints in some developing world cultures that limit their access to new technologies and the ability to network outside the domestic arena (Palackal et al., 2006). Age is also a social factor that differentiates access to social networks (generally favoring experience and tenure) and technology adoption and use (generally favoring youth). In combination, unequal access to network and technology resources along regional, gender and age dimensions can be grouped together under the phrase ‘geo-social asymmetries’ (Duque et al., 2009).

Detailed demographic information was collected from the research participants, because it is one of the main objective of the XXXX project. Table 2 shows that researchers from 18 different countries were surveyed, 67% of the respondents work in an Ecuadorian university, followed by Mexican universities (14%) and Spanish universities (5%). There is an obvious bias towards Ecuador, because the principal researcher is based in this country. All other countries represent 13% of the sample. For regions, we have 1% from USA, 9% from European Union, and 90% from Latin American and Caribbean Countries.

Table 2. Distribution of the sample for countries.[1]

[1] In the tables that follow, ’f’ refers to the absolute frequency, ’F’ to the cumulative absolute frequency, ’rf’ to the relative frequency (or proportion) and ’rF’ to the relative accumulated frequency.

Country	f	F	rf	rF
01. Ecuador	178	178	0.670	0.670
02. Mexico	38	216	0.140	0.820
03. Spain	13	229	0.050	0.870
04. Italy	6	235	0.020	0.890
05. Argentina	5	240	0.020	0.910
06. Brazil	4	244	0.020	0.920
07. Cuba	3	247	0.010	0.940
08. USA	3	250	0.010	0.950
09. Chile	2	252	0.010	0.950
10. Colombia	2	254	0.010	0.960
11. Peru	2	256	0.010	0.970
12. Venezuela	2	258	0.010	0.980
13. Czech Republic	1	259	0.000	0.980
14. France	1	260	0.000	0.980
15. Germany	1	261	0.000	0.990
16. Netherlands	1	262	0.000	0.990
17. Portugal	1	263	0.000	1.000
18. UK	1	264	0.000	1.000

The majority of participants work in universities (92%), where they are professors (68%), graduate students (13%), and post docs (5%). The specialties are shown in Table 3.

Table 3. Specialties of the interviewees.

Specialty	f	F	rf	rF
01. Technological Science	43	43	0.160	0.160
02. Life Sciences	35	78	0.130	0.300
03. Mathematics	20	98	0.080	0.370
04. Medical Sciences	16	114	0.060	0.430
05. Pedagogy	15	129	0.060	0.490
06. Sociology	15	144	0.060	0.550
07. Agricultural Sciences	14	158	0.050	0.600
08. Economic Sciences	14	172	0.050	0.650
09. Physics	14	186	0.050	0.700
10. Science Of Arts And Letters	13	199	0.050	0.750
11. Science Of Earth And Space	13	212	0.050	0.800
12. Geography	11	223	0.040	0.840
13. Astronomy And Astrophysics	10	233	0.040	0.880
14. Chemistry	9	242	0.030	0.920
15. Politic Science	7	249	0.030	0.940
16. Psychology	7	256	0.030	0.970
17. Other	3	259	0.010	0.980
18. History	2	261	0.010	0.990
19. Science (Field Not Defined)	2	263	0.010	1.000
20. Earth Sciences	1	264	0.000	1.000

Note from Table 3 that there is a balance between the surveyed who embrace scientific or technical careers and humanistic or artistic careers.

Results

Scientific Production

In this first section of our results, we will focus on the scientific production of the researchers that have answered all the questions of the survey. Two response variables were constructed, considering the sum of the scientific production measured in publications count terms. A first (pubTotal) that considered all publications, regardless of their type (from “P093” to “P0100”, Table 1). The category “All publications” contains these types of documents: Journal Articles, Books, Book Chapters, Reports, Conferences Articles, Non-arbitrated articles, and Publication in Blogs.

The second (pubScience) that takes into account only those publications considered scientific ones (from “P093” to “P098”, Table 1). This category contains these types of documents: Journal Articles, Books, Book Chapters, Reports, and Conferences Articles.

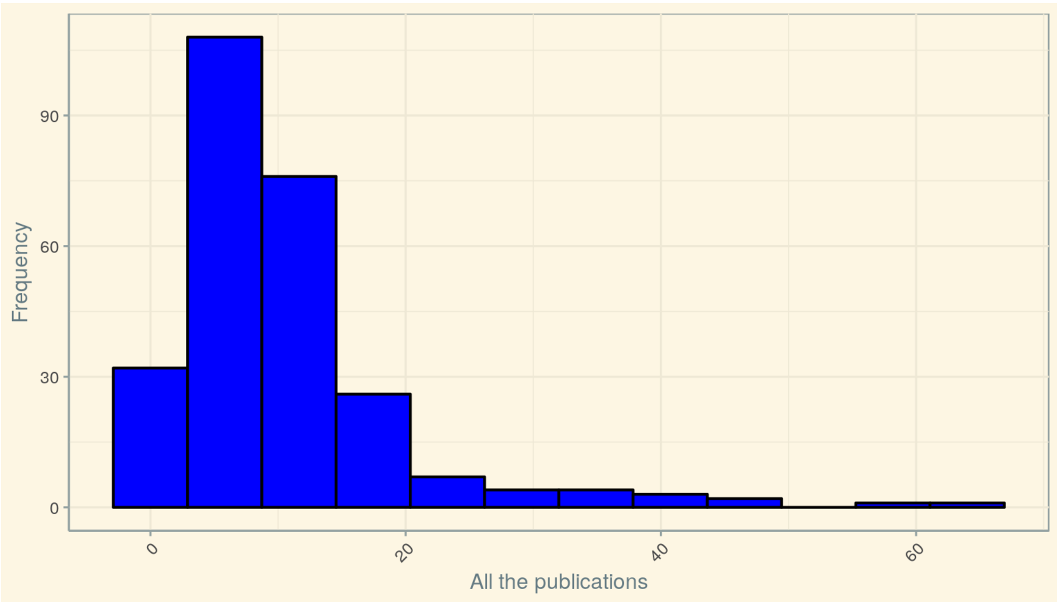

Figure 2. Total production reported in the sample.

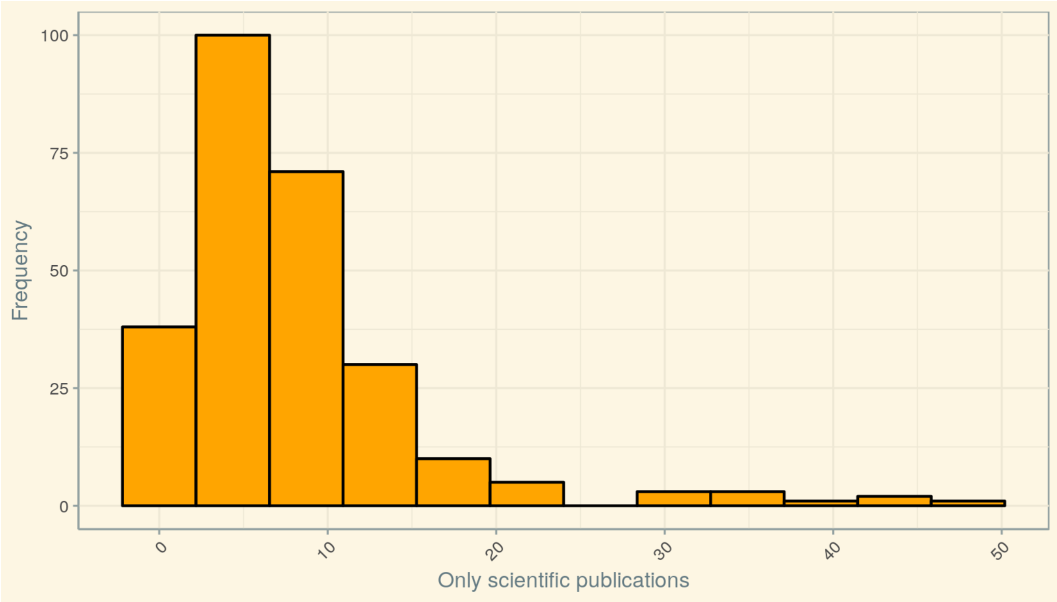

Figure 3. Scientific production reported in the sample.

Figures 2 and 3 contain the histograms of the total production and the scientific production, respectively. As expected, the general structure of both variables is very similar, favoring a number of publications around 10 in the first case, and 5 in the second.

We have done these two categories to map not only the scientific publications of researchers, but also their dissemination activity, that becomes more important every day to spread science in society. With this distinction, we argue that the concept of liquid science is something that goes from the academic environments to arrive to the rest of society.

As it can be observed, both variables suggest a Poisson process in their origins. On the other hand, the Poisson distribution requires the equality between the mean and the variance, that is, if $X \sim Poisson(\lambda) \Rightarrow E[X] = \lambda = V[X]$.

In the data set the estimated mean for the total production is 10.06 and its estimated variance is 86.93. For scientific production, the estimated mean and variance are 7.99, 53.49, respectively. So, we note over-dispersion problems. For this reason, Quasi-Poisson regression models are proposed for its analysis.

Table 4 contains the different levels of each variable considered, in case it is not of a numerical type. The first level is the reference level and is marked in blue.

Table 4. Variable levels for production analysis.

Variable	Levels
Country	Argentina, Brazil, Chile, Colombia, Cuba, Czech Republic, Ecuador, France, Germany, Italy, Mexico, Netherlands, Peru, Portugal, Spain, UK, USA, Venezuela
Gender	Female, Male, Prefer not to say (excluded)
Position	Academic technician, Associate Professor / Reader, Director, Doctoral student, Head of Hospital Teaching, Lecturer / Fellow, Master, Masters student, Occasional Professor, Official Radiation Protection, Post Doc / Assistant, Private company worker, Professor, Project Manager , Research Assistant, Research technician, Researcher, Scientific Director, Senior Lecturer / Senior Fellow, Senior Researcher, Student
InsType	Government, Private Sector, Public Research Institute, University
Discipline	Agricultural Sciences, Astronomy And Astrophysis, Chemistry, Earth Sciences, Economic Sciences, Geography, History, Legal Sciences And Law, Life Sciences, Linguistics, Mathematics, Medical Sciences, Other, Pedagogy, Physics, Politic Science, Psychology, Science (Field Not Defined), Science Of Arts And Letters, Science Of Earth And Space, Sociology, Technological Science
AloneTeam	Alone (excluded), Both, Team
LocSameIns	FALSE, TRUE
LocDiffIns	FALSE, TRUE
LocDiffCou	FALSE, TRUE
FreqInt	Bi-weekly, Daily, Less often, Monthly, No answer, Twice a week, Weekly
ContactLetter	Essential, No answer, Not Used, Used
ContactPhone	Essential, No answer, Not Used, Used
ContactEmail	Essential, No answer, Not Used, Used
ContactList	Essential, No answer, Not Used, Used
ContactInsMess	Essential, No answer, Not Used, Used
ContactVideoC	Essential, No answer, Not Used, Used
Age	(Numeric)
pubTotal	(Numeric)
pubScience	(Numeric)
TeamSize	(Numeric)
Loc	DiffIns, DiffIns/DiffCou, FALSE, SameIns, SameIns/DiffCou, SameIns/DiffIns, SameIns/DiffIns/DiffCou

A first, purely additive, quasi-Poisson regression model to study the association between the response variable (pubTotal) and the remaining explanatory variables is as follows:

$$M_1: \log(\lambda_1) = Country+Gender+InsType+Position+ AloneTeam+Discipline+Loc+FreqInt+ContactLetter+ ContactPhone+ContactEmail+ContactList+ContactInsMess+ ContactVideoC+Age+TeamSize$$

Where λ_1 is the rate of total publications per year.

Table 5 contains the analysis of deviance of the model factors, with terms added sequentially (first to last). The reference level in each case is the first. In red are indicated the factors that are significant at $\alpha = 0.1$ (that is, with 90% confidence).

Table 5. Analysis of deviance for total production (M_1).[2]

Factor	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			263	1731.060	
Country	17	171.405	246	1559.655	0.023
Gender	1	2.410	245	1557.245	0.512
InsType	3	7.378	242	1549.867	0.726
Position	18	201.360	224	1348.507	0.007
AloneTeam	1	1.991	223	1346.516	0.552
Discipline	19	230.268	204	1116.248	0.002
Loc	5	51.746	199	1064.502	0.101
FreqInt	5	29.716	194	1034.786	0.381
ContactLetter	2	2.558	192	1032.228	0.796
ContactPhone	2	43.902	190	988.326	0.020
ContactEmail	2	1.614	188	986.712	0.866
ContactList	2	39.291	186	947.421	0.030
ContactInsMess	2	29.254	184	918.167	0.074
ContactVideoC	2	51.209	182	866.958	0.010
Age	1	10.106	181	856.853	0.180
TeamSize	1	20.648	180	836.204	0.055

Now, scientific production refers to the number of publications per year that take into account only those considered as scientific production, so that if λ_2 is the rate of annual scientific publications (pubScience), a second purely additive model proposed is:

$$M_2: \log(\lambda_2) = \text{Country+Gender+InsType+Position+AloneTeam+Discipline+Loc+FreqInt+ContactLetter+ContactPhone+ContactEmail+ContactList+ContactInsMess+ContactVideoC+Age+TeamSize}$$

Table 6. Analysis of deviance for scientific production (M_2).

Factor	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			263	1325.939	
Country	17	118.495	246	1207.444	0.055
Gender	1	0.429	245	1207.015	0.753
InsType	3	8.459	242	1198.556	0.584
Position	18	164.671	224	1033.885	0.004
AloneTeam	1	3.315	223	1030.570	0.383
Discipline	19	159.593	204	870.977	0.009
Loc	5	43.119	199	827.858	0.078
FreqInt	5	18.530	194	809.328	0.513
ContactLetter	2	2.620	192	806.708	0.740
ContactPhone	2	29.378	190	777.329	0.034
ContactEmail	2	3.712	188	773.618	0.653
ContactList	2	35.485	186	738.132	0.017
ContactInsMess	2	16.601	184	721.532	0.148
ContactVideoC	2	37.423	182	684.109	0.014
Age	1	9.276	181	674.833	0.144
TeamSize	1	22.789	180	652.044	0.022

Comparing the results of tables 5 and 6, it is important that the “Loc” factor is not significant when explaining the total publications, but it explains the scientific production. It is clear that the location of researchers in a team work has an impact on scientific production and not on all the publications. This activity can be done individually

[2] Hereafter, Df: Degrees of freedom. Resid: Residual. Dev: Deviance. Pr(>Chi) [or Pr(>|t|)]: p-value.

for the researcher with local and national means. Consequently, we will concentrate on scientific publications exclusively to analyze how liquid science is distributed in the scientific networks.

On the other hand, in Table 6, Gender, InsType, AloneTeam, FreqInt, ContactLetter, ContactEmail, ContactInsMess and Age factors are not significant, and will be excluded in the search for a more parsimonious model. This does not mean that they are not important, but that nowadays they do not seem to make a difference in scientific production, e. g. male or female essentially produce in the same way, and email and other forms of contacts excluded are valued by all of them in similar ways. Then the new model for pubScience is:

$$M_3\text{: } \log(\lambda_3) = \text{Country+Position+Discipline+Loc+ContactPhone+ContactList+ContactVideoC+TeamSize}$$

Table 7 contains the analysis of variance of the model, now considering the levels of the factors versus the reference level (see Table 4).

Table 7. ANOVA for scientific production (M_3).

Factor	Estimate	Std. Error	t value	Pr(>|t|)	exp(Est.)
(Intercept)	1.974	1.026	1.924	0.056	7.199
CountryBrazil	0.664	0.585	1.137	0.257	1.943
CountryChile	0.642	0.735	0.874	0.383	1.900
CountryColombia	0.215	0.762	0.282	0.778	1.240
CountryCuba	0.087	0.741	0.117	0.907	1.091
CountryCzech Republic	0.720	1.377	0.523	0.602	2.054
CountryEcuador	0.333	0.507	0.656	0.512	1.395
CountryFrance	-0.110	1.283	-0.086	0.932	0.896
CountryGermany	-0.155	0.938	-0.165	0.869	0.856
CountryItaly	0.701	0.591	1.187	0.237	2.016
CountryMexico	0.756	0.530	1.426	0.155	2.131
CountryNetherlands	-0.011	1.179	-0.009	0.993	0.989
CountryPeru	1.567	0.700	2.240	0.026	4.792
CountryPortugal	-0.871	1.340	-0.650	0.516	0.418
CountrySpain	0.849	0.543	1.562	0.120	2.336
CountryUK	0.452	1.082	0.418	0.677	1.571
CountryUSA	0.166	0.714	0.233	0.816	1.181
CountryVenezuela	1.062	0.722	1.473	0.142	2.894
PositionAssociate Professor / Reader	0.028	0.833	0.034	0.973	1.029
PositionDoctoral student	-0.127	0.840	-0.151	0.880	0.880
PositionHead of Hospital Teaching	-0.131	1.298	-0.101	0.920	0.877
PositionLecturer / Fellow	-0.742	0.881	-0.843	0.400	0.476
PositionMaster	-0.114	1.238	-0.092	0.927	0.892
PositionMasters student	-1.254	1.366	-0.918	0.360	0.285
PositionOccasional Professor	-0.130	0.909	-0.143	0.887	0.878
PositionOfficial Radiation Protection	-0.171	1.260	-0.136	0.892	0.843
PositionPost Doc / Assistant	-0.265	0.858	-0.309	0.758	0.767

Factor	Estimate	Std. Error	t value	Pr(>|t|)	exp(Est.)
PositionPrivate company worker	-1.132	1.696	-0.667	0.505	0.322
PositionProfessor	-0.239	0.836	-0.286	0.775	0.787
PositionProject Manager	-17.195	2648.047	-0.006	0.995	0.000
PositionResearch Assistant	0.133	0.926	0.143	0.886	1.142
PositionResearch technician	-3.282	1.562	-2.101	0.037	0.038
PositionResearcher	-0.297	0.868	-0.343	0.732	0.743
PositionScientific Director	-1.681	1.290	-1.303	0.194	0.186
PositionSenior Lecturer / Senior Fellow	-0.075	0.921	-0.081	0.935	0.928
PositionStudent	-17.839	1650.388	-0.011	0.991	0.000
DisciplineAstronomy And Astrophysis	0.503	0.400	1.256	0.210	1.653
DisciplineChemistry	-0.388	0.397	-0.978	0.329	0.678
DisciplineEarth Sciences	0.104	0.827	0.126	0.900	1.110
DisciplineEconomic Sciences	-0.201	0.329	-0.610	0.542	0.818
DisciplineGeography	-0.447	0.386	-1.159	0.248	0.640
DisciplineHistory	1.076	0.871	1.236	0.218	2.934
DisciplineLife Sciences	-0.200	0.250	-0.799	0.425	0.819
DisciplineMathematics	-0.257	0.287	-0.893	0.373	0.774
DisciplineMedical Sciences	-0.166	0.302	-0.550	0.583	0.847
DisciplineOther	0.739	0.398	1.858	0.065	2.093
DisciplinePedagogy	-0.266	0.300	-0.886	0.377	0.766
DisciplinePhysics	-0.214	0.310	-0.691	0.490	0.807
DisciplinePolitic Science	0.257	0.354	0.726	0.469	1.293
DisciplinePsychology	0.176	0.354	0.497	0.620	1.193
DisciplineScience (Field Not Defined)	-0.152	0.594	-0.255	0.799	0.859
DisciplineScience Of Arts And Letters	-0.313	0.337	-0.930	0.353	0.731
DisciplineScience Of Earth And Space	-0.437	0.329	-1.329	0.185	0.646
DisciplineSociology	-0.179	0.301	-0.595	0.552	0.836
DisciplineTechnological Science	0.002	0.234	0.009	0.993	1.002
LocDiffIns/DiffCou	0.087	0.210	0.414	0.679	1.091
LocSameIns	-0.027	0.159	-0.172	0.864	0.973
LocSameIns/DiffCou	-0.194	0.311	-0.622	0.535	0.824
LocSameIns/DiffIns	-0.386	0.219	-1.766	0.079	0.679
LocSameIns/DiffIns/DiffCou	-0.064	0.163	-0.390	0.697	0.939
ContactPhoneNot Used	-0.358	0.188	-1.905	0.058	0.699
ContactPhoneUsed	-0.274	0.129	-2.127	0.035	0.760
ContactListNot Used	0.106	0.180	0.587	0.558	1.112
ContactListUsed	0.337	0.166	2.026	0.044	1.400
ContactVideoCNot Used	-0.546	0.191	-2.854	0.005	0.579
ContactVideoCUsed	-0.188	0.117	-1.605	0.110	0.828
TeamSize	0.022	0.009	2.439	0.016	1.023

Table 7 is produced by fitting a Quasi-Poisson regression model, since we assume that there is over dispersion in the sample. We add a column *exp(Est.)* with the exponentiation of the parameters estimates. The dispersion parameter (Φ) is estimated in 4.31, so the estimated variance is $7.99 \times 4.31 = 34.44$, near that one estimated directly from the data (53.49). The model goodness of fit is poor, as can be seen from the residual deviance (715.12) and residual degrees of freedom (197). Nevertheless, this is to be expected since only a very small subset of all the variables that would explain the variability is used.

All the factors in Table 7 are significant because at least one of its levels is significant. Then, we can interpret the estimators obtained from the model. Because the model fits in a logarithmic scale, the explanation is clearer by returning to the linear scale, which is (including the intercept):

M_3: log(λ_3) = Country+Position+Discipline+Loc+ContactPhone+ContactList+ContactVideoC+TeamSize

Then, in mathematical terms:

$$log(\lambda_3) = \beta_0 + \beta_1\ x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8$$

where the estimated parameters are the β_i (i=0,1,…,8) and the values of the variables, x_i. The β_i are really vectors, because they accompany categorical variables that have one parameter for each level of the factor (minus the reference), except β_8 because TeamSize is a numerical variable. Then

$$\lambda_3 = exp\{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8\}$$

And if the other variables remain constant, the effect of the *i*-th variable on the mean λ_3 is $exp\{\beta_i x_i\} = \left(e^{\beta_i}\right)^{x_i}$and every unit of increment in the variable x_i produces a multiplicative effect of e^{β_i} in the response λ_3.

In the next sub-sections we analyze each of the significant factors (levels) in relation to scientific publications productivity.

The country: Figure 4 shows the box plot of pubScience by country. The highest median values in the sample correspond to Peru, Spain, and Venezuela, however, only the number of scientific publications in Peru is significantly different from the rest according to our model (M_3). This is because the reported with respect to Spain and Venezuela has a high variance, but not the reported in relation to Peru.

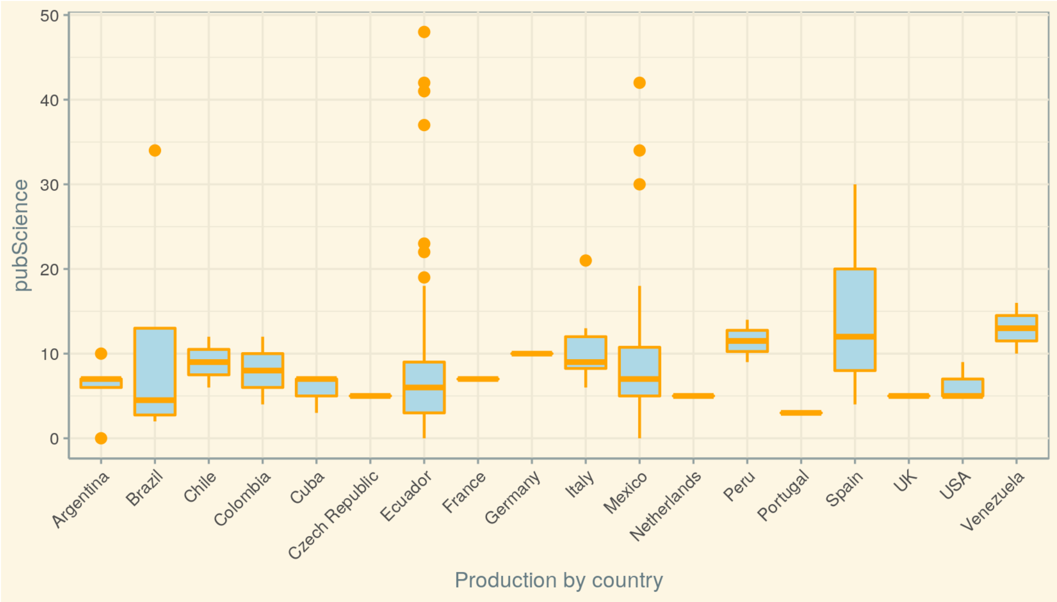

Figure 4. Box plot of scientific publications by country.

According to Table 7, the coefficient associated with Peru is 1.567, exponentiating 4.79, then a researcher from Peru has an estimated 4.8 times more the number of scientific publications than a researcher from Argentina (the reference level).

Position: The only significant level with respect to the position is Research technician. Its estimated coefficient is negative (-3.282), exponentiating 0.038, then research technicians have 0.04 times the number of publications that academic technicians (the reference level) and also that the rest of positions.

Discipline (Specialty): According to Table 7, the only significant level for discipline is “Other”. The scientific production is greater than the rest, but surely this is because in this category there are several different disciplines together. Consequently, the only interesting conclusion in this case is that there do not seem to be significant differences in the scientific production of the disciplines explicitly considered.

Location: In Figure 5 the proportions of the research teams are shown, according to whether their members belong to the same institution (SameInst), to a different institution (DiffInst), to institutions of other countries (OtherCountry) or its combinations.

The majority of their networks is located in the same institution (36%), 24% of the participants have colleagues in the same or in different institutions of the same country or in different countries. 20% of the sample report colleagues only from different institutions. The rest has similar interpretations. These proportions are useful to know the distribution of the research teams, and to know if the group is international or not.

Figure 5. Location of team members (proportions).

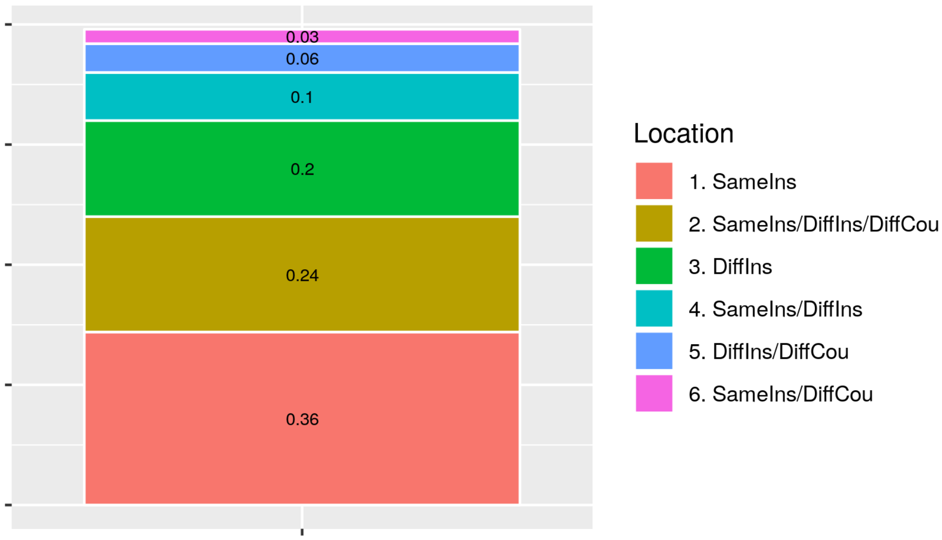

Figure 6 shows the box graphs for pubScience of each location level.

Figure 6. Location of team members (boxplot).

The medians and variability displayed in Figure 6 do not show many differences for each level of location, however it is clear that those researchers who reported working mainly alone have the lowest productivity of the sample, while those who reported work both alone and with other institutions and other countries show the highest productivity.

According to our model, SameIns/DiffIns is the only factor that is significant. Its coefficient is -0.386, exponentiating 0.679, then the team that includes colleagues from the same institution and different institutions, has 0.68 times more the scientific production than those teams that include colleagues from different institutions only. Given the importance of these variables for the study, Table 8 contains similar explanations for the rest of the Location levels.

Table 8. Location of the team members' analysis.

Level	exp(Est.)	Comment
DiffIns/DiffCou	1.091	The team that includes colleagues from different institutions and countries, has 1.091 times the scientific production that those teams that includes colleagues from different institutions only.
SameIns	0.973	The team that includes colleagues from the same institution only, has 0.97 times the number of scientific publications that those teams that includes colleagues from different institutions only.
SameIns/DiffCou	0.824	The team that includes colleagues from the same institution and different countries, has 0.82 times the scientific production that those teams that includes colleagues from different institutions only.
SameIns/DiffIns/DiffCou	0.939	The team that includes colleagues from the same institution, different institutions and countries, has 0.94 times the scientific production that those teams that includes colleagues from different institutions only.

Then, it is interesting that the sample consulted reports a higher productivity if working in teams with different institutions only or different institutions and different countries. This affirms the thesis that teamwork, among different institutions and different countries, increases scientific productivity.

Contact means: The way we propose to analyze the scientific networks and how these networks work is to examine the construction of the links between researchers and the use of different technologies, analogical and digital. For Latour (1988), the construction of this link should not be understood as purely symbolic or abstract, but rather as something material (like physical infrastructure, laboratories, offices, and computer networks). Researchers and analogical-or-digital devices constitute this network of a scientific community where everything is constantly under mediation and cooperation.

By focusing on the tools to communicate with the colleagues, we can observe how these have changed in the last years. Researchers have passed from analogical to digital tools to communicate with the rest of colleagues. The form to communicate in professional settings, and not, have changed strongly with the increase of the numbers of national and international networks. In Figure 7, it can be possible to observe how researchers communicate with the rest of the nodes of their networks using analogical tools (Letters, Phone Calls) and digital tools (Emails, Mailing Lists, Instant messaging, Video conference).

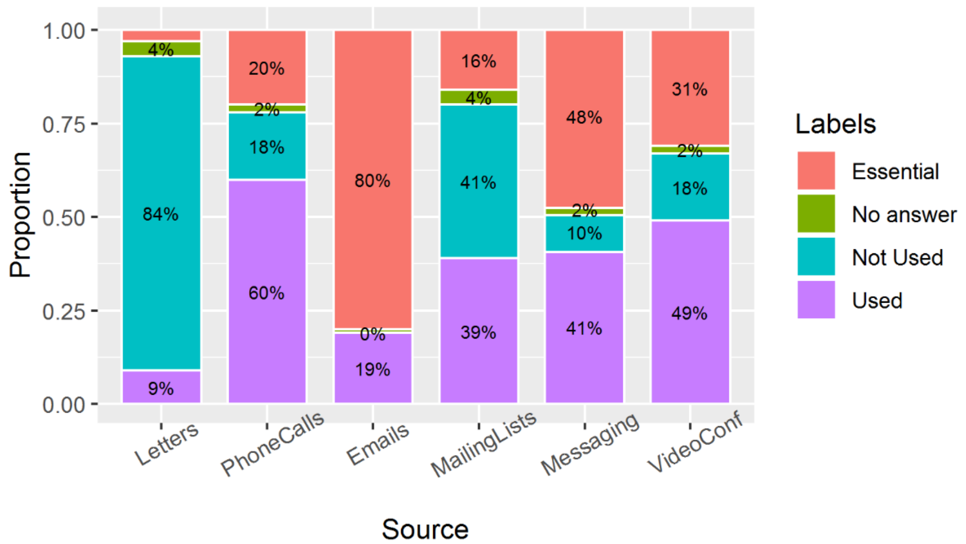

Figure 7. All the mechanisms of contact the team members.

In Figure 7, it can be observed how letters are not used by 84% (n=256) of the researchers, converting it in an obsolete tool. A majority of researchers (80%, n=244) agree that emails are essential in their professional tasks. This great agreement among the interviewees explains why these factors are not significant in our model.

In the last years messaging services via mobile have increased their impact to communicate between colleagues. For researchers, this service is considered essential and useful (89%, n=269).

Among the factors that are significant (see Table 7), we have (the reference level is "essential"):

- Phone: Respondents who say they have not used it have 0.7 times the scientific production than those who consider it essential. And those who say they have (but is not essential) has 0.76 times the scientific production than those who consider it essential. Clearly the use of the telephone is essential.
- Lists: Interviewees who say they have not used them have 1.112 times the scientific production than those who consider it essential. And who say they have (but are not essential) have 1.4 times the scientific production than those who consider it essential. Clearly the use of lists is necessary but not essential.
- Video: Respondents who say they have not used them have 0.58 times the scientific production than those who consider it essential. And those who say they have (but is not essential) have 0.83 times the scientific production than those who consider it essential. Clearly the use of the video conference must be considered essential.

*The size of the network***:** One important question of the analysis is to understand if the researcher works in a collective way, if they prefer to work in a research team or alone. Only a small number of researchers, 6% (n=19) work alone, and are excluded from this analysis. The size of their research groups is between 3 and 6 persons (63%), 34% of the researchers work in groups that have more than 6 researchers (see **Figure 8**).

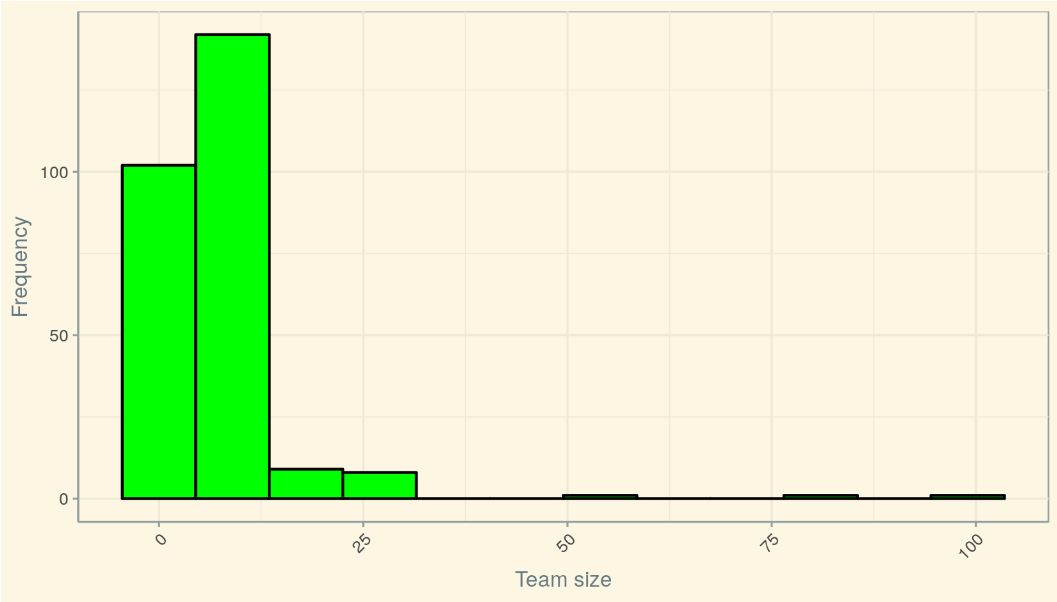

Figure 8. Size of the research groups.

According to our model, each member that is added to the team increases 1.023 times the number of scientific publications. The importance of working in a team is evident.

Discussion

We have observed several features of how liquid science flows in scientific networks thanks to our results that can give a response to our two initial hypotheses. The digital transformation in science has changed and modified the practices between researchers, influencing directly their research.

We have observed that more productive researchers work in medium and big research groups to achieve important results, and in consequence, to increase the number of scientific publications. Members of these research groups must come from different institutions and different countries to have advances in science. Heterogeneity of a research group is a determining variable to observe if the team can have a positive record or not.

It is surprising that in a digital era where the communication is instant and run to the screens, personal telephone conversations are considered essential to have great professional relations with the colleagues. Of course that emails and instant messaging are considered essential too, but phone conversations are needed to negotiate and to be clear between partners in order not to incur in misunderstandings caused by a text message. The use of this type of communication increases the number of scientific publications and scientific collaborations.

Video conferences are an evolution of phone conversations, where the body and facial expressions help to add more details to the voice, and for this it is considered essential to be more productive in science. This digital infrastructure helps researchers to share and distribute scientific facilities regardless of their type and location in the world. Liquid science has increased collaborations between researchers, institutions and countries, creating a digital scientific infrastructure. In this new era, the effective and multimodal communication in teamwork is essential.

Conclusion

The most representative changes that digital communication in scientific networks has provided to the scientific community are mainly two: 1) The efficiency related to communication, and, 2) The information and knowledge flow.

The exploratory and explanatory findings offered by this study suggest that researchers view the digital transformation of science as welcome intermediaries to create and manage scientific networks. Unlike the argument advanced by Bauman, digital tools and digital communication more broadly are not ‘liquefying’ the face-to-face communication, physical meetings, and conferences. Indeed, the data suggest that a majority of researchers continue to value and seek these social phenomena, and are merely using the technology as a means to pursue meaningful partnerships. This study’s participants felt they have more scientific possibilities than previous generations, and that the technologies give them greater agency with regards to pursuing and meeting potential scientific collaborations. The concept of 'networked individualism' (Rainie and Wellman, 2012) is readily applicable to studies of scientific networks of researchers as individuals become responsible for their professional practices within a broadened social network environment. These digital tools provide a scientific network that enhances the researcher's activity in what is called 'liquid science'. This 'liquid science' gives them access to an extensive network of scientific possibilities. These scientific networks enhance a researcher's capacity to build new relationships with researchers from other parts of the world. Liquid science brings new opportunities in the researcher's life.

The majority of this study's participants believed that technology merely enhanced their abilities to work in scientific networks, increasing the number of their scientific publications.

Somewhat superficial and not specific, traditional views on the difference between analogic and digital tools is still largely prevalent. However, the conclusions that can be drawn from our analysis should be limited for three reasons. First, the lack of representatives of different countries; the research was carried out mainly in only three of them for the survey (Ecuador, Mexico, and Spain). Although, we have chosen these countries according to the availability of researchers that collaborate with the members of XXXX project. Thus, further research is needed. Second, the analysis related only to the uses of researchers from different disciplines and at different stages of their careers. These researchers can adopt and shape not shared perceptions of the digital tools that they use. Finally, data do not allow to test if these scientific networks generate some positive results to the research, or they are not competitive like other scientific networks, for example not build on digital technology. The project has bias that can conduct the researcher to respond in a positive way to the introduction of digital technology in their scientific activities and networks.

What conclusions can be drawn from these findings? As discussed in the opening sections, some researchers see expanding their scientific networks thank to the digital tools, establishing professional relationships with other countries. Second, researchers want to increase the number of their scientific collaboration but also they need analogical meetings with others researchers, where face-to-face relations can accomplish the lack of this in digital communication. Thus, empirical evidence suggests that researchers do not end with the analogical communication with their networks and desire to continue to move forward in this mixed world where they can take advantage of both situations, promoting international collaborations.

Funding

This project has been developed within the framework of the research and innovation project Horizon 2020 of the European Union under a grant agreement no XXXX.

References

Agresti, Alan (2015). Foundations of Linear and Generalized Linear Models. John Wiley & Sons. New Jersey, USA.

Atkins, D., K. Droegemeier, S. I. Feldman, H. Garcia-Molina, M. L. Klein, D. G. Messerschmitt, P. Messina, J. P. Ostriker, & M. H. Wright. Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Washington, DC: National Science Foundation, 2003.

Atkinson, R. and J. Flint (2003) 'Sampling, Snowball: Accessing Hidden and Hard-to-reach Populations', pp. 274–80 in R.L. Miller and J.D. Brewer (eds) The A–Z of Social Research. London: Sage.

Author XXXX

Bauman, Z. (2003) Liquid Love: On the Frailty of Human Bonds. Cambridge: Polity.

Candela, L., Castelli, D., & Pagano, P. History, evolution and impact of digital libraries. E-publishing and digital libraries: legal and organizational issues. In Iglezakis, I. Synodinou, T. Kapidakis, S. (eds.), E-Publishing and Digital Libraries. New York: Information Science References, 2011, pp. 1-30.

Denscombe, M. (2010) The Good Research Guide: For Small-scale Social Research Projects, 4th edn. Maidenhead: McGraw-Hill.

Duque, R. B., Shrum, W. M., Barriga, O., & Henríquez, G. (2009). Internet practice and professional networks in Chilean science: Dependency or progress?. *Scientometrics*, *81*(1), 239-263.

Heimeriks, G., Hoerlesberger, M., & Van den Besselaar, P. Mapping communication and collaboration in heterogeneous research networks. *Scientometrics* 2003; 58(2): pp. 391-413.

Hobbs, M, Owen, S., and L. Gerber (2017) Liquid Love? Dating apps, sex, relationships and the digital transformation of intimacy. Journal of Sociology 53(2): 271-284.

Latour, B. (1988). The Pasteurization of France. Cambridge, MA: Harvard University Press.

Leonard, M., Graham, S., & Bonacum, D. (2004). The human factor: the critical importance of effective teamwork and communication in providing safe care. BMJ Quality & Safety, 13(suppl 1), i85-i90.

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G. R., Reznick, R., ... Grober, E. (2004). Communication failures in the operating room: An observational classification of recurrent types and effects.Quality and Safety in Health Care,13, 330-334.

Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. Acta Anaesthesiologica Scandinavica, 53, 143-151.

Nelder, J., and R. W. M. Wedderburn (1972) Generalized linear models. Journal of The Royal Statistical Society, A. 135: 370–384.

Neuman, W.L. (2011) Social Research Methods: Qualitative and Quantitative Approaches, 7th edn. Boston, MA: Pearson.

Olson, G. M., Zimmerman, A., Bos, N. (2008) (Eds), Science on the Internet. Cambridge: MIT Press.

Palackal, A., Anderson, M., Miller, P., Shrum, W. (2006), Internet equalizer? gender stratification and normative circumvention in science, Indian Journal of Gender Studies, 14 : 231–257

R Core Team (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rainie, L. and B. Wellman (2012) Networked: The New Social Operating System. Cambridge, MA: MIT Press.

Renaud, M. The Challenge of Building Infrastructure in the Social Sciences. Social Sciences for a Digital World: Building Infrastructure and Databases for the Future. Paris: OECD, 2000.

Sexton, J. B., Thomas, E. J., & Helmreich, R. L. (2000). Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. Bio-Medical Journal, 320, 745-749.