UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE AUGUSTIN COURNOT [ED221]

Bureau d'Economie Théorique et Appliquée, UMR 7522

THÈSE présentée par

Núria MORATAL FERRANDO

soutenue le : 28 Février 2019

Pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences de Gestion

The Role of Large Research Infrastructures in Scientific Creativity.

A User-Level Analysis in the Cases of a Biological Database Platform and a Synchrotron.

THÈSE dirigée par :

Mme WOLFF, Sandrine	Maître de Conférences HDR en Sciences Économiques, Université de Strasbourg

---

RAPPORTEURS :

Mme HUSSLER, Caroline	Professeur en Sciences de Gestion, Université Jean Moulin
Mme MERINDOL, Valérie	Professeur en Sciences de Gestion, Paris School of Business

---

EXAMINATEURS :

M.BURGER-HELMCHEN, Thierry	Professeur en Sciences de Gestion, Université de Strasbourg
M. COHENDET, Patrick	Professeur en Affaires Internationales, HEC Montreal
M. CORBEL, Pascal	Professeur en Sciences de Gestion, Université Paris Sud

“You certainly usually find something, if you look, but it is not always quite the something you were after.”

J.R.R. Tolkien, The Hobbit

Merci…

Quel étrange sentiment de mettre un point final à un projet auquel on a consacré cinq ans de sa vie. C'est l'instant propice pour faire le bilan de ces cinq années et de rendre hommage à toutes les personnes qui m'ont soutenu pendant cette période car l'aboutissement de cette thèse n'aurais jamais vu le jour sans le soutien et l'aide de nombreuses personnes.

Tout d'abord je tiens à remercier ma directrice de thèse, Sandrine Wolff, pour m'avoir offert l'opportunité de me lancer dans cette aventure au sein du BETA, pour ses conseils et pour m'avoir guidé dans la réalisation de ce défi. Je remercie également Caroline Hussler, Valérie Mérindol, Pascal Corbel, Patrick Cohendet et Thierry Burger-Helmchen d'avoir accepté de lire et évaluer ma thèse.

Merci à Mireille Matt, pour avoir pris le temps de faire partie de mon comité de suivi et m'avoir donné des bons conseils. Bien évidement je remercie mon co-auteur Moritz Müller qui m'a beaucoup aidé et appris. Je remercie également tous les enseignants-chercheurs de l'axe Innovation du BETA qui depuis mon Master et au cours de cette thèse m'ont formé à la recherche. Je remercie tout particulièrement Julien Pénin, pour m'avoir donné envie, pendant mon Master 2, de faire une thèse ; Patrick Llerena pour m'avoir encouragé à passer ma première année de thèse à Pise et Christophe Lerch qui m'a donné des nombreux conseils pendant la dernière année d'écriture de ma thèse. Merci à Émilie pour ses conseils et son soutien ces derniers mois. Enfin merci à Stefano pour tous ses conseils que tout au long de ces années m'ont aidé dans la réalisation de ce travail et dans mon apprentissage et formation comme chercheur.

Je suis également reconnaissant aux différents organismes qui ont permis la réalisation de cette thèse. Tout d'abord au BETA, qui m'a fourni le support intellectuel nécessaire à la réalisation de cette thèse et pour le financement de colloques et séminaires me permettant de découvrir l'ensemble des facettes que comporte le métier de chercheur. Merci à l'Ecole Doctoral Augustin Cournot de m'avoir fait confiance en m'accordant un contrat doctoral en 2013. Je remercie également l'IUT Louis Pasteur, la Faculté des sciences économiques et de gestion et la Faculté de Psychologie de m'avoir initié à l'enseignement, une activité qui j'aime beaucoup. Mes remerciements vont également à la Scuola Superiore Sant'Anna, à Pise pour m'avoir accueilli pendant une année en Italie où j'ai pu faire des très belles rencontres et j'ai découverte par la première fois l'académique. Enfin, j'adresse toute ma gratitude à Danielle, Géraldine, Virginie, Jane, Nathalie et Danièle, qui m'ont toujours aidé dans des nombreuses démarches administratives et qui ont toujours été patientes et souriantes.

C'est avec beaucoup d'émotion que je remercie l'ensemble de mes collègues doctorants. Les longs débats sur des sujets divers et variés que nous avons eus, ces moments de vie que nous avons partagés en dehors du PEGE, tous ces petits à côté ont été des parenthèses précieuses. Je tiens tout spécialement à remercier Vincent, pour ses nombreux encouragements, pour ses conseils et pour être toujours là pour m'écouter. Nos pauses café (presque) chaque après-midi restent des souvenirs précieux. Je remercie également mes amis Thomas, Cyrielle, Rémy, Thierry, Guillaume, Fréderic et Bruno avec qui j'ai passé des moments inoubliables. Merci aussi à Benoit, Arnaud, Catherine, Quoc, Anne, Yann, Laetitia, Deborah, Sofia, Nicolas, Benjamin, Nicolas, Audrey, Nguyen, Marion, Faustine, Emre et Sila avec qui j'ai partagé cette aventure.

Pour finir, un immense merci à tous ceux, qui en dehors du BETA et de l'Université de Strasbourg m'ont soutenu et accompagné. A Marcelo, pour son optimisme et ses conseils. A Aurore pour sa présence et son écoute au cours de cette dernière année. A Nicolas, avec qui j'ai eu et ai toujours des discussions passionnantes malgré la distance. Merci pour tous ces moments inoubliables. A Lauriane, pour nos discussions, sa gentillesse, son écoute et ses conseils. Je suis très reconnaissant envers mes parents, Andres et Lola, et mon frère, Andreu, qui me soutient inconditionnellement. Enfin, un grand merci à Paul, pour son soutien sans faille depuis le début de cette aventure et qui depuis ces longues années m'accompagne et me soutient en toute circonstance.

Overview

Table of Contents

Résumé exhaustif en français

1 Introduction et Emergence de la question de recherche

A l'origine de cette thèse il y a le constat d'une science en changement. Ce changement se caractérise par deux grandes tendances globales.

La première est la dépendance croissante à des grands équipements coûteux et partagés. Le progrès scientifique a toujours été jusqu'à un certain point déterminé par l'évolution des instruments scientifiques. Aujourd'hui, ces instruments sont de plus en plus complexes et sophistiqués. Ils atteignent souvent un coût qui les rend inaccessibles pour une seule université ou laboratoire (Stephan 2012). En conséquence, les instruments clés et d'autres types d'équipements sont de plus en plus financés par des programmes publics et proposés à la communauté scientifique selon un principe d'ouverture à des utilisateurs extérieurs. Dans ce contexte les Infrastructures de recherche (IR) constituent un élément de politique scientifique visant à fournir ces ressources.

La deuxième grande tendance de la science actuelle est la production de données de masse qui sont également très coûteuses à stocker et gérer. La science a toujours produit des données. Très tôt l'observation du monde environnant a conduit à la production de données. Compter, mesurer et conserver ont toujours été des tâches importantes de l'activité scientifique. Cependant, avec les nouveaux instruments de mesure la quantité de données a connu une croissance exponentielle (André, 2014). En effet, plusieurs domaines de recherche ont vu leurs pratiques transformées au cours des dernières décennies par le développement d'outils capables de produire des données de masse : la génomique et ses séquenceurs en sont un bon exemple. Cette croissance rapide des données remet en cause la capacité technologique de les stocker et de les gérer. Elle nécessite en outre des technologies de pointe, des ressources humaines expertes et, par conséquent, un financement très important. Comme pour les grands instruments, les IR et plus précisément les infrastructures numériques (ou e-infrastructures) tendent à jouer un rôle crucial en fournissant aux chercheurs un accès ouvert à des bases de données de haute qualité, très onéreuses à gérer et nécessitant une très grande capacité de stockage.

Plusieurs facteurs peuvent nous amener à penser que les IR sont des lieux favorables à la créativité. En effet les IR fournissent un accès à une technologie de pointe pour faire de la recherche et elles sont ouvertes à toute la communauté scientifique. On observe dans tous les textes de politique scientifique en Europe que l'objectif de ces IR est de réussir à produire une science d'excellence. Enfin il existe plusieurs exemples de grandes découvertes scientifiques faites grâce à l'une de ces IR. Ceci suggère que ces infrastructures sont des endroits propices à la créativité scientifique.

Cependant les moyens par lesquels les IR favorisent la créativité n'ont pas été étudiés. Pour traiter cette question, cette thèse est organisée de la façon suivante : Le Chapitre 1 introduit le contexte empirique, le cadre théorique et les choix méthodologiques. La problématique se décline en deux sous-questions de recherche. D'abord nous nous demandons *comment les IR peuvent-elles contribuer à la créativité scientifique de leurs utilisateurs.* Puis nous nous interrogeons sur : *comment mesurer cet impact ?* Pour répondre à la première question le Chapitre 2 utilise une étude de cas sur un très grand instrument de recherche, un synchrotron. Le Chapitre 3, quant à lui, étudie le cas d'une grande plateforme de bases de données biologiques. Enfin, pour répondre à la deuxième question le Chapitre 4 développe des mesures d'impact de créativité.

2 Chapitre 1 : Contexte empirique, cadre théorique et méthodologie

2.1 Cadre théorique

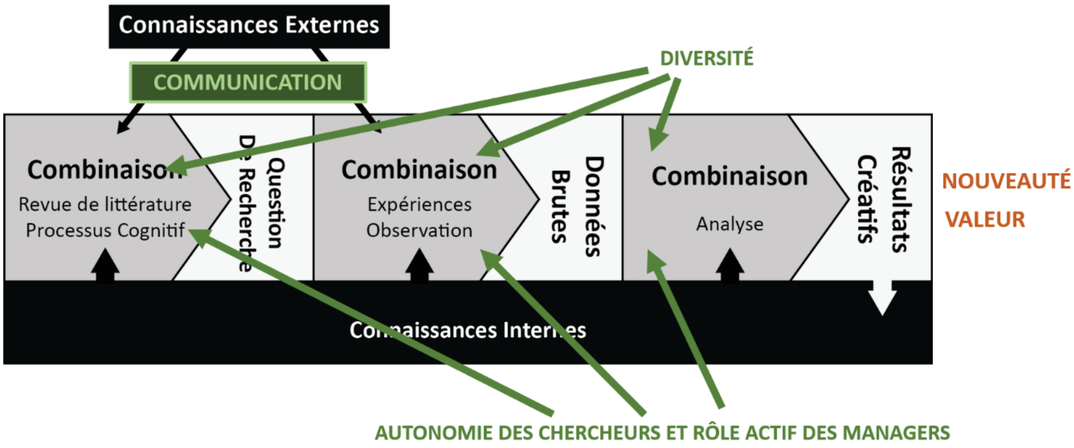
En premier lieu il est important d'introduire la notion de créativité scientifique. Quand il s'agit de l'étude de la science, nous proposons de définir la créativité comme la production de connaissances et de capacités qui sont nouvelles et qui ont de la valeur. Ces deux critères, la nouveauté et la valeur sont des critères en conflit car en science les idées très novatrices sont souvent rejetées et considérés comme bizarres ou improbables (Staw 1995). Ceci s'explique souvent par la résistance des paradigmes en place face aux paradigmes émergents (Kuhn 1962, Merton 1973). Une conséquence possible est que la science devienne moins novatrice car les chercheurs décident de ne pas prendre de risques et préfèrent exploiter des voies de recherche déjà ouvertes (Heinze 2009). En lien avec ce qui précède, on observe une division, dans l'étude de la créativité scientifique, entre ceux qui vont se concentrer sur l'étude de l'impact et ceux qui, conscients des problèmes de reconnaissance de la science novatrice, vont se concentrer sur la nouveauté. Dans cette thèse nous nous efforçons de toujours considérer les deux attributs de la science créative.

Pour répondre à la problématique, cette thèse introduit un cadre théorique qui est nourri et inspiré par plusieurs courants de littérature. Tout d'abord nous proposons d'étudier la création des connaissances scientifiques comme un processus collectif. En effet, plusieurs auteurs en sociologie de la science considèrent que la science est un processus collectif (Callon 1988, Latour 1999), et qu'elle est de plus en plus faite en équipe car les articles à auteur unique sont de plus en plus rares (Li et al., 2013). Pour élaborer le cadre théorique de cette thèse, nous considérons aussi que ce processus est non linéaire et qu'il utilise des connaissances externes et internes à tout moment, comme cela a été proposé par plusieurs auteurs en science du management pour l'étude des processus d'innovation (Garud et al., 2013).

En outre, nous soutenons l'idée que ce processus de création scientifique est caractérisé par une activité combinatoire dans laquelle des nouvelles connaissances sont créées à partir de combinaisons de connaissances existantes, ainsi que l'ont décrit plusieurs économistes de l'innovation tels que Nelson et Winter (1982). Ces combinaisons, quand elles sont rares ou très rares, vont favoriser la créativité comme le montrent plusieurs chercheurs en sociologie comme Uzzi et al (2013). En psychologie aussi, il est admis que l'union de cadres de pensée différents et jamais mis ensemble auparavant débouche sur des résultats créatifs comme il a été proposé par Koestler (1964).

Enfin, il est important de prendre en compte que l'occurrence des nouvelles combinaisons favorisant la science créative dépend des facteurs organisationnels dans lesquels la science se fait (Simonton 2004, Amabile 1988). Plus précisément, des auteurs venant d'une multitude de disciplines différentes telles que le management, la psychologie ou la sociologie, observent que les facteurs suivants vont être propices à la créativité. Premièrement, ils relèvent l'existence d'une communication active et efficace des individus à l'intérieur et avec l'extérieur de l'organisation (Heinze et al 2007), communication qui est elle-même souvent facilitée par l'existence de communautés (Brown and Duguid, 1991). Ensuite, ils notent l'importance de la diversité, à plusieurs niveaux, comme par exemple la diversité de profils des équipes ou bien la diversité des méthodes (Hollingsworth 2002). Enfin, ces auteurs mettent en avant le rôle de l'autonomie des chercheurs pour prendre des décisions sur la manière dont ils vont mener leur recherche" ou sur les questions pertinentes à étudier, ainsi que l'investissement des managers dans l'activité scientifique (Heinze 2007, Amabile 1988).

Ce travail de revue de littérature nous permet d'aboutir au modèle conceptuel suivant :



Ce model traduit un processus qui se passe en plusieurs étapes, qui repose sur l'utilisation en continu de connaissances internes et externes et sur l'existence d'une série de facteurs favorables tels que : la communication, la diversité, l'autonomie des chercheurs et l'investissement des managers dans l'activité scientifique. Ces conditions favorables vont permettre d'avoir des résultats nouveaux et utiles, c'est à dire des résultats créatifs. Le premier objectif de cette thèse est de comprendre par quels moyens les IR favorisent la créativité, et ce modèle conceptuel doit nous permettre d'analyser les mécanismes et les facteurs expliquant la créativité scientifique en général, avant de l'appliquer au cas des IR, puis d'aborder, dans la dernière partie de la thèse, la question de la mesure de cette créativité.

2.2 Choix méthodologiques

Pour répondre à la première des sous-questions de recherche nous avons adopté une méthodologie qualitative d'étude de cas et l'avons utilisée dans les chapitres 2 et 3. Le Chapitre 2 étudie le cas d'un synchrotron, et le Chapitre 3 celui d'une plateforme de bases des données biologiques. Ce choix méthodologique a été motivé par le fait qu'il s'agit de très grandes IR et donc de cas emblématiques susceptibles de bien illustrer l'avenir de la science tout en permettant de bien identifier les mécanismes sous-jacents à la créativité. Cette recherche repose sur une démarche abductive qui consiste à procéder par des aller-retours entre des éléments théoriques et des faits empiriques issus du terrain. Ce dernier a été exploré à l'aide d'une quarantaine d'interviews. Tous ont été transcrits en verbatim, puis analysés avec un codage par thèmes venant du cadre théorique et aussi avec des thèmes qui ont émergé du terrain.

La deuxième méthode mise en avant dans cette thèse consiste en un développement d'une mesure de créativité et en l'application de la méthode ainsi développée. C'est l'objet du Chapitre 4, qui propose une mesure d'impact des IR sur la créativité de leurs utilisateurs. Cette méthode est développée à partir d'une recherche bibliographique pour connaitre les différentes mesures de valeur et de nouveauté, et en les adaptant aux besoins de cette recherche. La méthode est ensuite appliquée au cas du synchrotron à travers l'analyse de deux lignes de lumière séparément. Un total de 761 articles utilisant le synchrotron a été comparé à plus d'un million d'articles publiés dans les mêmes domaines et les mêmes années. Ces articles ont été téléchargés depuis le Web of Science avec, à chaque fois, toute l'information disponible, telle que : auteurs, mots clés, *abstract* ou citations.

3 Chapitre 2 : Les grands instruments en tant que facilitateurs du processus créatif des utilisateurs : le cas d'un synchrotron

3.1 Cadre général

Ce chapitre se concentre sur l'étude d'un grand instrument de recherche, un synchrotron, qui représente un cas exemplaire de l'un des deux grands changements que connaît la science aujourd'hui : l'utilisation d'équipements lourds, complexes et très coûteux. L'objectif de ce chapitre est de comprendre de quelle façon et par quels moyens un tel équipement peut faciliter le processus créatif. Un synchrotron est un accélérateur de particule. L'accélération génère une onde électromagnétique qui est collectée à différents endroits appelés lignes de lumière. A chacune d'entre elles le faisceau lumineux est transformé avec des miroirs et autres outils optiques afin de sélectionner une gamme de longueurs d'onde. La lumière transformée est ensuite utilisée pour examiner des échantillons ou conduire des expériences. Deux lignes de lumière sont étudiées spécifiquement dans la thèse, celle de diffraction par rayons x, qui est une technologie relativement mature, d'une part, et celle de spectroscopie par infrarouge, qui est une technologie plus récente en synchrotron[1], d'autre part. Dans ces deux cas, et aussi pour toutes les autres lignes de lumière du synchrotron, ce sont des utilisateurs extérieurs qui se déplacent jusqu'à l'instrument avec leurs échantillons, pour faire des expériences pendant quelques heures ou quelques jours, et ensuite repartir dans leurs laboratoires. Dans ce chapitre nous étudions le cas des utilisateurs académiques, issus de la recherche publique.

Nous cherchons à identifier les facteurs organisationnels et les mécanismes favorables à la créativité que note modèle conceptuel propose. Pour cela nous étudions les pratiques mises en œuvre par le personnel de l'IR, les pratiques des utilisateurs et les diverses manifestations de la créativité observées sur les deux lignes de lumière du synchrotron.

3.2 Résultats

L'analyse du cas du synchrotron fournit plusieurs éléments qui montrent la pertinence de notre cadre théorique à travers la présence des facteurs favorables, des mécanismes et des résultats créatifs attendus. Le tableau suivant résume ces résultats :

Facteurs Favorables	Multidisciplinarité des scientifiques du synchrotron (davantage dans le cas de la ligne infrarouge)
	Variété et multidisciplinarité des utilisateurs (davantage dans le cas de la ligne infrarouge)
	Projets multidisciplinaires (davantage dans le cas de la ligne infrarouge)
	Autonomie des chercheurs du synchrotron sur leur recherche et la gestion de l'instrument, et investissement des responsables dans l'activité scientifique (davantage dans le cas de la ligne infrarouge)
Mécanismes	Communication effective basé sur la confiance

---

[1] Bien que la spectroscopie infrarouge ne soit pas une technique récente dans l'absolu, elle l'est dans le cas de ses applications en synchrotron. Au moment de notre étude, elle est utilisée depuis seulement une dizaine d'années et est considérée comme particulièrement novatrice.

	Petites collaborations et partenariats de long terme qui se développent en continu grâce à cette confiance
Résultats créatifs	Développement technologique : nouvelles techniques et méthodes
	Nouvelles connaissances : articles publiés dans de bonnes revues

Nous constatons que les éléments identifiés dans le tableau ci-dessus sont toujours présents avec beaucoup plus d'intensité pour la ligne infrarouge, c'est-à-dire celle qui met en oeuvre une technologie récente, que pour la ligne de rayons x, qui constitue une technologie plus mature.

Parmi tous les résultats, deux sont à souligner. En premier lieu, nous pouvons observer, surtout pour la ligne infrarouge, une grande multidisciplinarité des utilisateurs, ce qui permet à ces derniers de faire des rencontres inattendues et variées : sont susceptibles de s'y croiser, par exemple, des médecins, des biologistes, des physiciens, des chimistes, etc. En second lieu, un facteur très important est l'autonomie des chercheurs du synchrotron, car ils utilisent cette autonomie pour s'investir dans la recherche faite par leurs utilisateurs et pour encourager les rencontres et collaborations. De plus, nous remarquons que l'activation de tous les autres mécanismes et facteurs favorables, tels que variété et multidisciplinarité, dépendent de la présence première de cette autonomie et de l'investissement des chercheurs du synchrotron.

Concernant les résultats créatifs on observe principalement de la création des connaissances publiées dans des bonnes revues. Plusieurs des acteurs interrogés affirment que en utilisant ce synchrotron ils sont surs de savoir que leur article sera accepté dans des revues très importants telles que Science ou Nature. On observe aussi de la créativité dans le développement de nouvelles techniques, méthodes ou des nouveaux outils. En effet utilisateurs et employés de l'IR travaillent ensemble pour développer les technologies nécessaires pour pouvoir répondre aux nouvelles questions de recherche et aux nouveaux défis de la science.

Ceci nous amène aussi à un résultat créatif inattendu, la genèse d'une nouvelle communauté scientifique. En effet, dans le cas de la ligne infrarouge, l'engagement pro-actif des scientifiques du synchrotron pour accroître la variété de leurs utilisateurs, pour qu'ils travaillent ensemble, développent de nouvelles méthodes et ouvrent de nouvelles voies de recherche a débouché sur la naissance d'une nouvelle communauté, c'est-à-dire un groupe de personnes qui communiquent en continu de façon formelle et informelle. Cette communication se base sur des liens de confiance et ils travaillent sur un nouveau domaine de recherche qu'ils ont en commun. Plus concrètement, il s'agit d'une communauté de médecins et biologistes qui font de la recherche médicale en utilisant des méthodes de physique.

Tous ces résultats étant liés à la nature physique du synchrotron qui lui permet d'être un lieu de rencontre, il est important d'étudier aussi des IR de nature différente. C'est pourquoi le Chapitre 3 se concentre sur une infrastructure numérique, une plateforme de bases de données biologiques.

4 Chapitre 3 : De grandes bases de données biomédicales en tant que moteurs de créativité : le cas de l'industrie pharmaceutique.

4.1 Cadre général

Ce chapitre se concentre sur l'étude de très grandes bases de données, servant de cas emblématique pour illustrer un des deux changements majeurs de la science aujourd'hui : la production et l'exploitation de données de masse. Plus précisément, le Chapitre 3 est consacré à une plateforme de bases de données biologiques. La recherche en biologie génère de très grandes quantités de données. La gestion de ces données est devenue une question complexe au fil du temps, car elle nécessite de très grandes quantités d'espace de stockage et une énorme capacité de calcul. C'est pour cette raison qu'il existe un besoin croissant d'e-infrastructures qui rendent accessibles, analysent, intègrent et résument les données disponibles, fournissant ainsi une ressource d'une valeur inestimable pour la communauté scientifique (Bolser et al., 2012 ; Gong et al., 2011). Les principaux domaines de recherche concernés par ces bases de données sont la génomique et ses différents sous-domaines et la discipline qui exploite et gère ces bases de données est la bio-informatique, qui est un domaine interdisciplinaire par essence. Il existe de nombreuses bases de données proposées par des institutions commerciales et publiques.

Cependant, lorsque nous parlons de bases de données volumineuses, il n’y a que quelques acteurs pertinents. Dans ce chapitre nous étudions EBI, « European Bioinformatics Institute ». Cette plateforme de bases de données est utilisée par les chercheurs du monde entier car elle contient toute l’information générée par la science sur un très grand nombre de gènes, protéines et autres composants. Concernant l’usage de cette plateforme, on compte 33 millions de connexions d’adresses IP uniques par mois, 38 millions de requêtes par jour et 8,7 pétabytes téléchargés en 2016.

Les utilisateurs sur lesquels nous nous concentrons ici sont ceux de l’industrie pharmaceutique. Plus précisément ce sont de grandes entreprises du secteur pharmaceutique. Au cours de la seconde moitié du XXe siècle, la recherche médicale comportait encore une part importante d'expériences randomisées et les données utilisées consistaient en de petits silos de données produits en interne. Au cours de la dernière décennie, l'utilisation de grands bases externes de données pour la recherche en santé, et en particulier pour la recherche pharmaceutique, s'est progressivement développée. L'industrie pharmaceutique s'appuie de plus en plus sur l'utilisation de ces e-infrastructures, de sorte que les bases de données publiques et la bio-informatique constituent aujourd'hui un aspect clé de la découverte de médicaments, contribuant à la fois à la découverte et à la validation de cibles. Dans ce chapitre, nous examinons si ce changement a également eu un impact sur la créativité scientifique des entreprises concernées.

4.2 Résultats

L’analyse de cette étude de cas fournit plusieurs éléments qui soutiennent l’intérêt de notre cadre théorique, qu’il s’agisse de facteurs favorables, de mécanismes ou de résultats créatifs.

Facteurs Favorables	Accès à une plus grande variété de connaissances (par exemples données très diverses comprenant une grande variété de paramètres)
	Accès à des connaissances, provenant d’autres communautés.
	Equipes de recherche plus variées pour pouvoir exploiter les données

Mécanismes	Standards
	Outils de langage : ontologies
	Contrôle de la validité des données
Résultats créatifs	Ouverture de nouvelles voies de recherche
	Plus de composés prometteurs (susceptibles de passer aux essais cliniques)

Une partie de ces résultats méritent d'être soulignés ici. Tout d'abord cette plateforme de bases des données permet l'accès à des connaissances provenant d'autres communautés. Ceci est fait grâce à des outils de langage qui s'appellent des ontologies. Différentes communautés (par exemple communauté de recherche sur le Parkinson et communauté de recherche sur le cancer de l'estomac) ont traditionnellement utilisé des noms différents pour parler des mêmes gènes ou protéines. Ceci fait que, historiquement, il a toujours été difficile de comprendre et donc d'utiliser des connaissances venant d'autres communautés. Cette plateforme productrice d'ontologies permet de surmonter ce problème. Ces outils permettent aussi de faire des liens entre différentes bases de données présentes dans la plateforme. Un autre résultat important est le constat que la grande variété des bases des données présentes dans cette plateforme demande aux utilisateurs d'avoir des équipes de recherche plus variées, avec des compétences et des connaissances très différentes, ce qui est un important facteur favorable à la créativité.

Enfin, concernant les résultats du processus créatif, nous observons l'ouverture de nouvelles voies de recherche. Être confronté à un tel nombre et une telle variété de données conduit à se poser des questions que l'on n'aurait pas pu se poser auparavant.

En plus des résultats créatifs prévus par le cadre théorique, nous avons trouvé un résultat inattendu : l'augmentation de l'occurrence des découvertes accidentelles, aussi appelées découvertes par sérendipité. En effet, avoir accès à un tel nombre d'informations, de nature aussi variée, et avec des outils permettant d'établir des connexions entre elles, augmente fortement la probabilité de faire une combinaison pertinente par accident, ou du moins une connexion qui aura un résultat inattendu.

Ces résultats, avec ceux du Chapitre 2, nous permettent de mieux comprendre les moyens par lesquels les IR facilitent la créativité. Dans le chapitre suivant nous examinons la possibilité de mesurer les résultats créatifs des IR.

5 Chapitre 4 : Mesurer l’impact des IR sur la créativité scientifique

Le Chapitre 4 est consacré au développement d’une mesure d’impact en termes de créativité. Ce chapitre part du constat qu’il n’existe pas de mesure d’impact d’un facteur causal (programme, IR ou autre) sur la créativité scientifique. En outre les mesures classiques d’évaluation de la science se concentrent sur la dimension de la valeur, et les nouvelles mesures visent à évaluer exclusivement la nouveauté, mais on ne trouve pas des mesures de créativité qui incluent les deux critères. L’objectif de ce chapitre est donc de développer une telle mesure d’impact. Pour ce faire nous utilisons des articles scientifiques comme unités de base de l’analyse.

5.1 Développement méthodologique

Nous utilisons des articles scientifiques pour mesurer la création de connaissances ou plus précisément la «production» scientifique. Considérant que notre définition de la créativité est un savoir nouveau et qui a de la valeur, notre méthodologie comporte deux parties. Dans une première partie de l'analyse, nous évaluons l'impact des publications effectuées au synchrotron, car elles représentent une approximation de la valeur perçue par les pairs. Dans une deuxième partie, nous observons le degré de non-occurrence des citations, que nous considérons comme un indicateur de nouveauté. Chacune de ces deux parties est composée de deux étapes, ce qui nous donne une analyse en quatre étapes. L’analyse est effectuée pour chacune des deux lignes de lumière déjà étudiées dans le Chapitre 2.

5.1.1 Impact

La première partie consiste en une analyse d’impact et se fait en deux étapes : dans une première étape, nous comparons les revues dans lesquelles la science réalisée grâce au synchrotron est publiée avec d’autres revues de la discipline. Dans un deuxième temps, nous comparons les articles dus au synchrotron à d’autres articles du même journal. En d'autres termes, nous examinons si les articles dus au synchrotron sont publiés dans de «bonnes» revues et, une fois publiés, s'ils ont une «meilleure» performance que les autres articles de la revue en termes de citations. Cela est effectué systématiquement en séparant les deux lignes de lumière étudiées : rayons X et infrarouge.

Nous considérons le journal dans lequel un article apparaît comme indicatif de sa valeur ou impact. Les revues diffèrent par divers aspects tels que la visibilité, la qualité, le processus d'évaluation par les pairs, etc. Nous faisons l’hypothèse simplificatrice que ces aspects se traduisent et sont résumés pour chaque revue par le facteur d’impact[2]. L'organisation sociale de la science est telle que les chercheurs préfèrent publier et dans des revues à fort impact plutôt que des revues à faible impact. Cela crée une hiérarchie de revues dans laquelle les revues à impact élevé sont capables de publier les « meilleurs » articles. Nous souhaitons donc savoir si les recherches sur les lignes de lumière sont publiées dans de "meilleures" revues, c’est-à-dire des revues à facteurs d’impact plus élevés, ou non.

Comme autre mesure de l’impact, nous examinons le nombre de citations. Nous déterminons ici quelle fraction des articles produits en utilisant le synchrotron reçoit plus de citations que le nombre médian de citations du même journal et de la même année. Cela nous indique si les articles publiés en utilisant le synchrotron sont plus ou moins cités que le reste des articles dans le même journal scientifique.

5.1.2 Nouveauté

Pour la deuxième partie de notre étude de la créativité scientifique, nous étudions le deuxième aspect de la créativité, qui est la nouveauté. Cette analyse, comme précédemment, se fait en deux étapes. Nous continuons à utiliser des articles de revues scientifiques et la logique est toujours de comparer les articles utilisant le synchrotron avec des articles n’utilisant pas le synchrotron. En nous appuyant sur l’idée que la science est un processus combinatoire, nous supposons que les scientifiques combinent des éléments de connaissances déjà acquises. Dans notre approche, ces éléments de connaissance sont identifiés à des citations. En effet, on peut raisonnablement imaginer

---

[2] L’usage du facteur d’impact pour évaluer la valeur d’une revue est criticable dans de nombreuses disciplines, mais il représente une approximation acceptable dans le cas des sciences du vivant.

que les citations représentent la connaissance qui a été utilisée pour produire de la science. Dans un premier temps, nous utilisons un indicateur de la fréquence à laquelle les papiers dus au synchrotron citent des journaux «rares» et, dans un deuxième temps, un indicateur de la fréquence à laquelle ils réalisent des combinaisons rares de journaux cités. Tous les calculs sont effectués par année et par discipline.

Nous suivons l'idée générale selon laquelle ce qui est nouveau ou rare dépend du public. Par exemple, au 16ème siècle, la porcelaine était très rare et nouvelle en France - mais pas tellement en Chine. Il en va de même pour les instruments ou méthodes scientifiques. Par exemple, il y a 20 ou 30 ans, de nombreuses méthodes économétriques fortement utilisées en science économique n'étaient pas encore beaucoup utilisées dans la recherche en sciences de gestion. Nous partons de l'idée qu'un journal a tendance à viser à la fois un public et / ou un sujet de recherche donnés. Ensuite, un article dans un journal donné peut fournir des informations utiles à ses lecteurs et faire progresser leurs recherches. Les références données par un article peuvent être interprétées dans le même esprit. Si un article cite un article d'un journal, ce dernier utilise des informations relatives au public ou au sujet du journal cité.

5.2 Application de la mesure et résultats

Nous appliquons notre mesure d'impact sur la créativité au même cas que le Chapitre 2, celui du synchrotron avec deux lignes de lumière : la ligne de rayons x qui est une technologie mature, et la ligne infrarouge qui est une technologie plus récente.

En commençant par l'étude de la valeur on observe que la recherche faite sur le synchrotron est publiée dans des revues à facteur d'impact plus élevé, ce qui signifie qu'ils sont reconnus comme étant meilleurs que la moyenne du domaine. L'effet est plus fort dans le cas de la ligne de rayons x. Cependant nous observons que ces articles, une fois publiés dans de bonnes revues, ne reçoivent pas systématiquement plus de citations que les autres articles du même journal et de la même année (pas moins de citations non plus). Les recherches sur le synchrotron sont publiées dans de bonnes revues mais ensuite ils sont cités autant que les autres articles de la revue. Cet effet s'améliore avec les années ce qui pourrait indiquer une reconnaissance tardive de la recherche nouvelle.

En ce qui concerne l'étude de la nouveauté, pour le cas de la ligne de rayons x nous pouvons constater que la tendance est de citer des revues qui ne sont pas rarement cités.

C’est le contraire pour la ligne infrarouge, où des citations de revues non usuelles dans le domaine sont souvent introduites. Si on revient au cas de la ligne de rayons x en considérant cette fois l’occurrence de nouvelles combinaisons, nous observons que cette ligne a un effet positif en termes de nouveauté, dans la mesure où des combinaisons de citations rarement faites sont introduites. Nous obtenons le même effet pour la ligne infrarouge, mais avec moins d’intensité. Il est important de noter que ces résultats sont conformes à ce qui avait été trouvé dans le Chapitre 2 : la ligne infrarouge est plus novatrice que la ligne de rayons x. Ceci s’explique donc très vraisemblablement par une technologie plus récente, des utilisateurs plus variés, et un responsable de ligne plus pro-actif.

6 Conclusion

Notre revue de littérature nous a permis de construire un modèle conceptuel et la suite du travail de thèse nous a permis d’en valider la pertinence, tout en suggérant d’intéressants éléments additionnels. Ainsi, alors que la littérature avait mis en avant trois types de facteurs favorables différents et d’importance égale, notre recherche révèle que l’autonomie des chercheurs et le rôle actif des managers dans l’activité scientifique est un facteur déterminant, voire un prérequis, car il permet ensuite à la diversité et à la communication d’agir favorablement pour la créativité. Un autre apport qui n’avait pas été relevé préalablement par la littérature a été d’identifier un nouveau type de résultat créatif, en l’occurrence l’émergence d’une nouvelle communauté scientifique.

En ce qui concerne les apports de la thèse en termes de pratiques managériales, le premier est sans doute d’avoir identifié l’importance de la participation active des responsables à la recherche car celle-ci n'est pas une simple condition favorable à la créativité. Toutes les autres conditions favorables en dépendent. De plus à travers le développement d’indicateurs de mesure d’impact de créativité, nous proposons aussi un outil de suivi des performances créatives qui pourrait être utilisé par d’autres IR, et peut-être même être appliqué à d’autres types d’institutions scientifiques.

Enfin, comme tout travail de recherche, cette thèse présente des limites qui ouvrent autant de perspectives pour de futurs travaux. La première limite a trait à la méthode utilisée dans les chapitres 2 et 3, qui s’est principalement appuyée sur des études de cas. Les résultats n'ont donc pas une portée générale. Une première perspective serait d'étendre l'étude à d'autres IR et, pour le cas de la plateforme de bases de données, d'étendre l'étude à d'autres types d'utilisateurs tels que les utilisateurs académiques, ou encore des utilisateurs de petites sociétés pharmaceutiques ou de biotechnologies.

En ce qui concerne l'indicateur de mesure de créativité, avant de chercher à en généraliser l'usage, il serait intéressant d'explorer des mesures alternatives au facteur d'impact des revues, sachant que c'est un indicateur très répandu mais aussi très critiqué voire inutilisable dans un certain nombre de disciplines scientifiques, telles que les sciences humaines et sociales ou les mathématiques. Dans le même ordre d'idées, nous envisageons aussi d'explorer d'autres moyens d'évaluer la nouveauté, à l'aide d'outils d'analyse et de statistique textuelle notamment. Suite à ces approfondissements méthodologiques, il devrait être possible d'utiliser et généraliser notre approche d'évaluation à l'étude d'autres IR ou d'autres organismes de recherche. Elle pourrait plus largement être utilisée pour évaluer les impacts de diverses politiques scientifiques sur la créativité des chercheurs.

Introduction

7 Motivation and emergence of the research question

Science has brought us findings that allow us to live the way we do today. In fact without the existence of science and scientific inquiry we would not have electricity which would mean no computers, internet or smartphones. A world without science is unimaginable as humans have always done science. Science is based on curiosity and the desire to understand how our world works. In fact, we are natural Scientists watch children and you will see that young children play like Scientists work, with investigation. Science is, however, not only about the discovery of electricity or the invention of internet. It is indisputable that science has advanced through history thanks to the accumulation of both, big and small steps in form of new pieces of knowledge. All the knowledge that is produced by science (such as experimental evidence, empirical observation, theories, methods and even new ideas and open questions) serves to advance one step forward towards the solutions posed by today's scientific challenges.

7.1 Some recent examples of scientific discovery: HIV and stem cells

Let us look at one example: human immunodeficiency virus (HIV) has infected more than 78 million people since 1981, it is the cause of AIDS disease and it has result in over 40 million deaths. Finding a cure to this virus is one of the biggest challenges of today's medical research. If science is to find a preventive vaccine for HIV, which, according to Shin (2016) is still not likely to happen soon, several years of research studying how the HIV virus behaves and how it changes the human genome would be needed. These advances vary in size and relevance. For instance, research on the virus started in 1981 and the first treatment for was introduced in 1987 and it consisted on an antiretroviral (Zidovudine) that slowed down the replication of the virus in the human body and reduced the mother to child transmission during pregnancy and breast feeding. It had, however, important side effects and it didn't prevent AIDS disease to develop (it just retarded it). Research continued and several other antiretroviral drugs were developed until in 1997 an antiretroviral therapy, which was a combination of several drugs, became the new standard treatment and caused 47 percent decline in death rates due to important delay on AIDS acquisition. Research continued and today there are more than 40 antiretroviral drugs available to control the virus and in high-income countries the standard treatment assures infected people a life expectancy that is only slightly shorter than non-infected people as death rates have shorten by more than 80%. Most people infected with HIV today in high income countries will probably never develop AIDS and the low viral load in their bodies prevents sexual transmission. The history of HIV research shows several important discoveries despite not having found a cure for the disease yet.

Let us look at another relatively recent example. The discovery of the existence of stem cells was, at the time, a breakthrough discovery. Stem cells possess the innate ability to change into any kind of cell. This means they can turn into, for example, a red blood cell, a white blood cell or a muscle cell. The existence of stem cells was discovered in the 80s. After that a lot of research has been done in order to advance towards the understanding of how these stem cells work. All the discoveries that allowed us to know a bit more about stem cells are relevant. Building on those, the discovery that has been path breaking in the sense that it can by itself bring a solution to a scientific (and social) challenge was the development of a methodology that allow scientists to program those stem cells. Indeed, since 2006 we know that any cell of the body can be reprogrammed and turned into a stem cell. Moreover, once we have a stem cell, we know now how to ask it to transform in any cell we want. This is a breakthrough discovery that brings the solution to long term faced research and societal challenges. We can expect[3], for instance, once clinical trials are finished, to be able to use that technique to replace damaged tissue with new cells and stem cells may be the key, as well, to be able to cure diseases such as Parkinson disease or Alzheimer. All the previous developments, however, were as well big scientific advances.

7.2 Changes in Science and Science Policy

Policy makers, when planning science policy, aim at the occurrence of these kinds of important discoveries. The objective of any policy maker and the objective and desire of the society is for science to advance which needs continuous relevant scientific developments. Sometimes to do this we will need from very large projects and some other times from everyday research that advances towards a better understanding of the

---

[3] https://www.nature.com/stemcells/index.html

problems that we are facing. How can we assure this? How to assure path breaking discoveries to occur?

Today, science is relying the more and more in big equipment. The scientific advancement has always been to some point set by the advance on instruments and the ability to do good quality research in certain areas by the access to those instruments. Today these instruments are increasingly complex and sophisticated. They often reach a cost that makes them inaccessible for single universities or laboratories (Stephan 2012). As a consequence, key instruments and other kinds of equipment are the more and more financed by public institutions and offered to the scientific community with an open logic. When provided under this logic this equipment is known ad Research Infrastructure (RI). RIs are facilities, resources and services used by the science community to conduct research. They include large scale research instruments (such as particle accelerators and telescopes), collections, depositories, public repositories (for example insect, mice or grain repositories), libraries, databases, biological archives, networks of computing facilities, research vessels, satellites and aircraft observation facilities, coastal or natural observatories, etc. The European Commission has defined them as places *"to achieve excellence in highly-demanding scientific fields and simultaneously build the European Research Area (ERA)"*. There is an expressed intention of achieving path breaking results and conducting disruptive science [4]. We can, therefore, expect RIs to come up with big scientific advances such as the ones described before.

We have several examples, in the history of science, of discoveries that have been made possible by RIs. One example is the recent observation (2017) of the first candidate exomoon using the Hubble Space Telescope and the Kepler space telescope[5]. To date, astronomers have discovered a few thousands (close to 4000) of exoplanets, which are objects orbiting stars other than the Sun. A hunt for exomoons, which are bodies that orbit these distant planets, has proceeded in parallel. But these natural satellites had lingered at the limits of detection with current techniques. Another example is the discovery of the

---

[4] This is developed in Chapter 1.

[5] https://www.nature.com/articles/d41586-018-06918-9

Higgs boson[6] using the LHC at CERN. Scientists had theorized about sub atomic particles and how they interact with each other. For long time theories had a missing element to understand matter and this element, the Higgs boson, was theorised in the 60s. However, it was not until 2012 that this element was finally observed empirically at CERN.

In order to achieve or enhance these kinds of discovery, however, there is a need for some continuity. As we explained before, the kind of very big discoveries that would make it to the news and win a Nobel prize are great. However there are a lot of smaller, but still big, everyday advances that are necessary for the advancement of science. Breakthrough are preceed by smaller discoveries but they are as well generally followed up and enriched with a flow of additional scientific outputs. RIs are expected to drive us towards all these kinds of discoveries. In order to know whether RIs are going in the right direction towards this desired goal, it is important to look at whether they are enabling not only breakthrough discoveries but also everyday knowledge creation and everyday knowledge advances. Traditionally the ability to create new knowledge have been explained as driven by creativity. Creativity (a concept better developed in Chapter 1), is the ability to produce something new and valuable. The created item may be a physical object (for instance, an invention or a piece of art) or it might be intangible (such as an idea, a musical Master piece or a scientific theory). Our interest, in this work, is creativity in science. We study creativity as the ability to produce knowledge that is new and valuable and we focus on the specific case of Research Infrastructure.

7.3 Emergence of the research question and justification

The general issue, or problematics, behind this thesis is:

How can the new trends in science contribute to discovery?

It has been recognized by scholars in sociology of science that the way scientific knowledge is produced is changing. Gibbons, (2000) explains how a new form of knowledge production started in the mid-20th century. It was context-driven, problem-focused and interdisciplinary. It involved multidisciplinary teams that worked together for

---

[6] https://atlas.cern/updates/atlas-feature/higgs-boson

short periods of time on specific problems in the real world. Gibbons and his colleagues use the label "mode 2" to describe this paradigm. But the changes that science is experiencing go further than that. Actually, a large part of science is more and more done at a large scale and it depends increasingly on large and costly equipment. Scientific advance has always been influenced by the advance on instruments (Stephan 2012). Today these instruments often reach a cost that makes them inaccessible for single organizations (public or private). Because of this Governments have to provide for this equipment. This is often done by means of joint projects between institutions and even between countries. They tend to follow a logic of openness and are accessible to all researchers who require it. Research Infrastructures (RIs) offer these resources that are done at a very large scale, with a logic of openness and that aim to achieve breakthrough research. These resources often consist in big instruments, but data and the associated resources are also becoming more and more important in science and are believed by many to become central for science in the coming years. With recent instruments for measurement (such as sequencers, satellites or large telescopes) there is as well an exponential increase in the data available in many fields of research (André, 2014). This big mass of data poses technical challenges for its use and provision and is provided by RI due to the cost and size of the task. This mass of data does not come only with challenges, it comes also with promising opportunities derived from its use. Because of this, research in several fields has become the more and more data driven and the analysis of the data has become a very important part of the science production process.

We observe therefore that the science is moving towards an increasing use of large amounts of data and large-scale advanced instrumentation, and that RIs play a major role in providing the scientists with these crucial resources. Surprisingly enough, the impact of these new trends (i.e. instruments, data and RIs use) on the discovery process has not been very studied yet. Because the creation of knowledge (which is at the bases of discovery) depends strongly on creativity, we reach the following research question:

Do Research Infrastructures drive scientific creativity?

There are several reasons that make us think RIs can have a positive effect on scientific creativity.

If look at the website of ESFRI, which stands for “European Strategy Forum on Research Infrastructures” and is a strategic instrument to develop the scientific integration of Europe and to strengthen its international outreach, we see that the objective is to make science advance at big steps. ESFRI, by assuring a rationale of excellence and open access to high quality Research Infrastructures aims to support and benchmark the quality of the activities of European scientists, and to attract the best researchers from around the world. ESFRI operates at the forefront of European and global science policy and contributes to its development translating political objectives into concrete advice for RI in Europe. This, together with some big discoveries that have recently used RIs make us believe, intuitively, that RIs are likely to be a proper place for the study of scientific creativity. Later, on Chapter 1, we will see that the literature review points into the same direction.

7.4 Brief summary of future literature review

Let us have a quick look at what literature says about scientific creativity. Scientific activity has been described as the combination of knowledge in order to produce new knowledge. The process of combination has been studied by multiple authors (i.e. Klahr and Simon, 1999; Merton, 1975; Simon, 1977; Simon et al., 1981) who describe science as a combinatory activity that integrates different perspectives, methods and concepts.

Creativity consists on the production of knowledge which is new and valuable (Amabile, 1988). New knowledge, in order to be creative, must be the result of novel, unexpected and relevant combinations. It needs as well to be considered as valuable by the peers, that is the scientific community. Teresa Amabile finds that individuals working in organizations are creative when these organizations allow them to have intrinsic motivation. She explains that the managers of an organization have a crucial role to play in order to assure this intrinsic motivation and to make it compatible with the objectives of the organization. The most relevant management leverages are to allow for certain autonomy and to ensure effective communications. There is a robust body of literature that tries to identify and explain the organizational conditions helping scientific creativity from the point of view of Sociology of sciences and focusing on the creativity at the level of the organization and not the individual (i.e. Hage and Mote, 2010; Heinze et al., 2009;

Hollingsworth, 2002; Zuckerman, 1967). The most relevant factors identified are communication and diversity.

Pelz and Andrews (1966) find that scientists are most creative when they do not work alone. When interacting actively with each other they are more creative. Creativity is also correlated with high frequency of intra-organizational communication (Heinze et al., 2009). When scientists in the same organisation communicate, they nourish each other's ideas and viewpoints and they are more likely to come up with an original idea. In addition to an effective intra-organizational communication, it is also important to allow communication across the boundaries of the organization. When communicating with scientists working in other institutions the gap between each other's experiences is bigger and there is, therefore, more room for learning and exchanging (Hollingsworth 2000, 2004).

Indeed, what makes communication crucial is the exchange of experiences, concepts and viewpoints. Hollingsworth (2000, 2004) introduces as well the idea that this communication is more beneficial when it operates not only across organizational boundaries, but also across discipline boundaries. The more the amount of disciplines that participate in a research project, or share a laboratory or institution, the greater the chances for the scientists of being creative as there will be a higher variety of elements interacting. This diversity allows for a greater set of knowledge and ideas to be available in order to solve a problem (Hollingsworth, 2004). Similarly, Heinze (2009) shows the importance of the access to a relatively large variety of technical skills. RIs are described by the European Commission[7] on its website and on ERA and ESFRI[8] documents as *"facilities, resources and services used by the science community to conduct research and foster innovation. By pooling effort and developing RIs, European countries can achieve excellence in highly-demanding scientific fields and simultaneously build the European Research Area (ERA) and Innovation Union. They include: major scientific equipment, resources such as collections, archives or scientific data, e-infrastructures such as data and computing systems, and communication networks. RIs can be single-*

---

[7] http://ec.europa.eu/research/infrastructures

[8] ESFRI stands for "European Strategy Forum on Research Infrastructures" (https://www.esfri.eu/objectives-vision)

sited (a single resource at a single location), distributed (a network of distributed resources), or virtual (the service is provided electronically)". RIs come in a variety of forms such as large-scale research instruments (particle accelerators or telescopes) but also collections, depositories, public repositories (insect, mice or grain), databases, biological archives, networks of computing facilities, satellites[9].

Let us study the definition in order to compile the most relevant aspects. As we can see in the definition, RIs are "us*ed by the science community*". RIs follow a logic of openness, the research conducted at RIs is not only performed by an internal team, but it is open to the scientific community. This makes of RI very particular places and suggests that they may enable the encounter of different individuals, which might allow communication and therefore creativity. Additionally, RIs are often, although not always, multidisciplinary and allow as well for multidisciplinary encounters. These means putting a variety of disciplines together, which is another of the factors traditionally identified as favouring creativity. It is because of this reason that we thing that the study of our research question, concerning the role of RIs as facilitators of scientific creativity, is promising.

The research question can be divided into two, more specific, sub questions:

How can research Infrastructure contribute to scientific creativity?

Is it possible to measure this impact?

These questions will be faced by using qualitative case study methodology and will be studied from the point of view of Science Policy and Innovation Studies.

8 Relevance of the topic, contribution and field

Creativity is supposed to be at the roots of knowledge creation, invention and, consequently, innovation. For this reason, this thesis aims to contribute to the literature in the field of Science Policy and Innovation Studies, often referred to as "innovation studies" or "science and technology studies". It is a multidisciplinary field which is

---

[9] For a detailed list of European RI and a European RI map:

http://ec.europa.eu/research/infrastructures/pdf/ri_landscape_2017.pdf#view=fit&pagemode=none

nourished mainly by four sizeable scientific disciplines: policy, sociology, management and economics (Martin, 2012). Because describing the scope of the field is a complex task, we opt for using the one proposed by the journal that is considered as the more relevant in the field, namely, Research Policy. The field is devoted to *"analysing, understanding and effectively responding to the economic, policy, management, organizational, environmental and other challenges posed by innovation, technology, R&D and science. This includes a number of related activities concerned with the creation of knowledge (through research), the diffusion and acquisition of knowledge (e.g. through organizational learning), and its exploitation in the form of new or improved products, processes or services"*[10]. This definition of 'science policy and innovation studies' is broad but it illustrates an essential element which is that the subject (innovation, technology, R&D and science) is studied using a range of social science disciplines (management science, organisational studies, sociology, economics and economic history, policy studies, etc.) which aim is to study empirically and theoretically the interaction between innovation, technology or research, on the one hand, and social, organizational, economic and political processes, on the other.

8.1 The study of science

At the emergence of the field most of the contributions came from the areas of economics and sociology and the focus was the study of economic activity (growth, transformation and cycles) in an alternative way to the neo-classical tradition as the intellectual foundations of the field are "evolutionary economics". Innovation and science took a central role in this view of the economy. Soon research started focusing on the study of the actors of science and innovation, namely universities and the firm. This focus started in the late 60s with a growing contribution from economists and economic historians, sociologists, from the fields of organisational studies, management and to a smaller extent political science. These research activities were at the beginning done in relative isolation, but they started to interact with each sometimes in teams of researchers, such as those at SPRU and Manchester, who were less constrained by disciplinary boundaries than those working in a single-discipline.

---

[10] https://www.journals.elsevier.com/research-policy/

One of the focus of the field was the study of National Innovation Systems. Innovation is considered by the most knowledgeable authors of the field as the motor of economic dynamics. Schumpeter (1934) argues that radical innovations are also linked to the change of technological paradigms and often have their origins in science and knowledge creation at universities. Freeman (1987) introduces the concept of National Innovation systems (NIS) as a network of interacting public and private institutions who initiate, import, modify and use new technologies. These NIS are the focus of analysis to understand the appearance of radical and incremental innovations and finally technological change and economic growth. Nelson (1993) and Lundvall (1992) will continue studying NIS paying particular attention at the role of science and the link between science and technology. Pavitt (1998) shows how academic research provides with fundamental laws of nature and general knowledge but also with solutions to specific technological problems, methodological tools, instruments and human capital. All in all, when industries interact strongly with Academia technological development is more likely to be found and ultimately leads to economic growth. Patel and Pavitt (1991) show this correlation between academic productivity and economic performance. They propose, however, that rather than a causal-effect relationship there is an interdependence: both elements feed each other. Scientific advance comes with economic growth which provides society with the resources to finance science. Research has shown strong historical evidence of this bi-directional interaction between science and technology which is, for example, provided and studied in-depth by Rosenberg (Landau et al., 1986; Rosenberg, 2009). Because science has properties of public goods leading to market failures, it is often publicly funded. The funding of science does not only require providing laboratories and researchers with grants for resources. The resources needed are sometimes large and therefore very expensive. In this case they are directly provided by governments and are sometimes even built at a supra institutional and supra national level (e.g. E.U.). This is typically the case of Large Research Infrastructures, which are the focus of this thesis.

8.2 The study of creativity

Literature on management focuses on how the firm can adapt to an evolving environment and keep up with technological advance. The role of the researchers in management in the field has traditionally been the study of innovation, reflecting growing knowledge about the nature of the innovation process in its various forms. It is in this context that the idea of the study of creativity first appeared in the field. Creativity is needed in order to keep having new ideas and innovating. For instance, Woodman et al. (1993) developed a theoretical framework for understanding organisational creativity. When it comes to the study of creativity in science, the field has only taken some interest recently. Most of the knowledge we have today has been produced by researchers in other fields, mainly the disciplines of psychology and sociology of science.

8.3 The scope of this thesis

This thesis examines the role of Large research Infrastructure on scientific creativity. As explained earlier and because creativity consists in the generation of original and valuable ideas which are at the origins of innovation, this thesis belongs to the field of Science Policy and Innovation Studies. The scope of the field is large and this thesis is situated between two of the core topics of the field, namely Science Policy and Creativity (see Figure 1). The interest on the topic lies firstly in its impact on innovation and therefore on economic growth but it is as well of high relevance for the issues related to the management of innovation and creativity at the organisational level. More specifically our work focuses on a very specific kind of science policy: the development and evaluation of Large Research Infrastructure. This topic is rather understudied, especially when it comes to the focus on creativity. The originality of the thesis lies here but the novelty of the topic also comes with little literature to build up on.

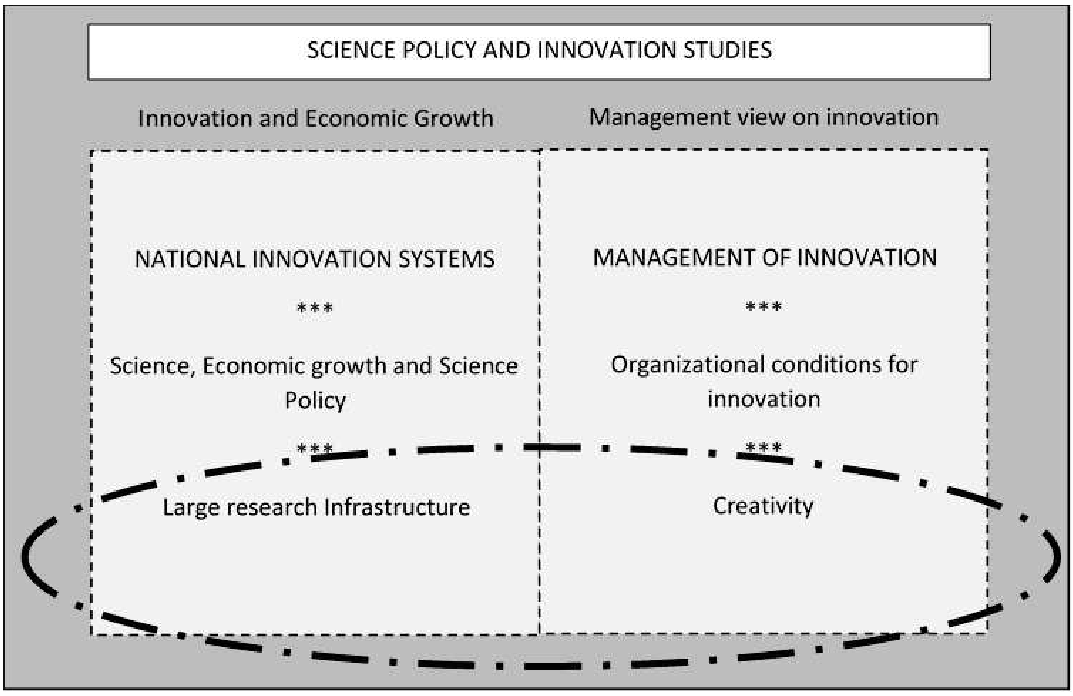


Figure 1 Scope of the thesis: Large Research Infrastructure and creativity

9 Structure of the thesis

In order to answer the research question, this thesis is going to be structured in four chapters as follows:

Chapter 1 (Context and theoretical background) consists in a literature review in the areas of public science policy, research infrastructure and scientific creativity. The disciplines that are mobilized to establish this theoretical framework are the economics of science, the management of innovation and the sociology of science. We introducing the conducting line of this thesis work which is the study of a process of science creation. This process has the particularity of being creative, contrary to the focus on individuals that have characterized the study of scientific creativity in history. Additionally to this, the ideas of value and novelty will be discussed as well as they are crucial for the study of creativity and one of the most important building blocks of our theoretical framework. In addition, the chapter will introduce the epistemological foundations of the thesis as well as the methodology.

Chapter 2 (Large Instruments as facilitators of users' creative process: The role of organizational factors, collaborations and communities in the case of a European Synchrotron) consists of a qualitative case study analysis where we try to identify the favourable conditions for creative research that are offered at the RIs. More particularly we look at the case of a large instrument, a synchrotron, which is a particular type of particle accelerator. The range of scientific fields that use this generic technology is very broad: life sciences, drug manufacturing and research, materials science, computer chip design, chemistry, medicine, physics and geology. We want to know whether and how the synchrtron has an influence in the creative process of its users. We find that this multidisciplinary facet, among some other factors will be found to be crucial for creativity.

Chapter 3 (Large Bio Medical Databases as drivers of creativity: An analysis of the case of the Pharmaceutical Industry) consists in a second qualitative case study analysis where we analyse the case of another kind RI, a biological database platform. This type of database platforms collects information on several biological compounds, the most relevant being proteins and genes. These databases are free to access and used by both researchers from public universities and researchers from private companies. The chapter focuses on the case of users from the Pharmaceutical Industry and how their creativity is triggered by the use of these databases. We find that this platform allows scientists to access to knowledge produced by other communities, which was close to impossible before due to the use of different standards and different languages. This means that they have access to a wider quantity but also a wider variety of knowledge which allows them to be more creative.

Chapter 4 (Measuring RI's impact on scientific creativity. The case of the synchrotron) purpose is to discuss ways to identify and measure scientific creativity and to find an appropriate way to measure the impact of RIs on scientific creativity. Chapter 2 and 3 tend to support the idea that RIs reunite several of the organizational conditions for scientific creativity. In order to see whether these conditions conduct to a creative output empirical quantitative analyses of the output are required. However the literature does not provide us with a suitable methodology to measure creativity and therefore we try to develop one. In our methodological development we consider articles - that is, peer-reviewed journal articles - as the main results of the research. We compare the bibliometric statistics of the articles written in the framework of the synchrotron with those of the articles not realized at the synchrotron. Because the concepts of novelty and value have been one of the main driving threads of this thesis, of this our methodology consists of two parts, one where value is considered and another one where novelty is considered.

Figure 2 sums up all these aspects and shows how the thesis is built in order to answer the research question

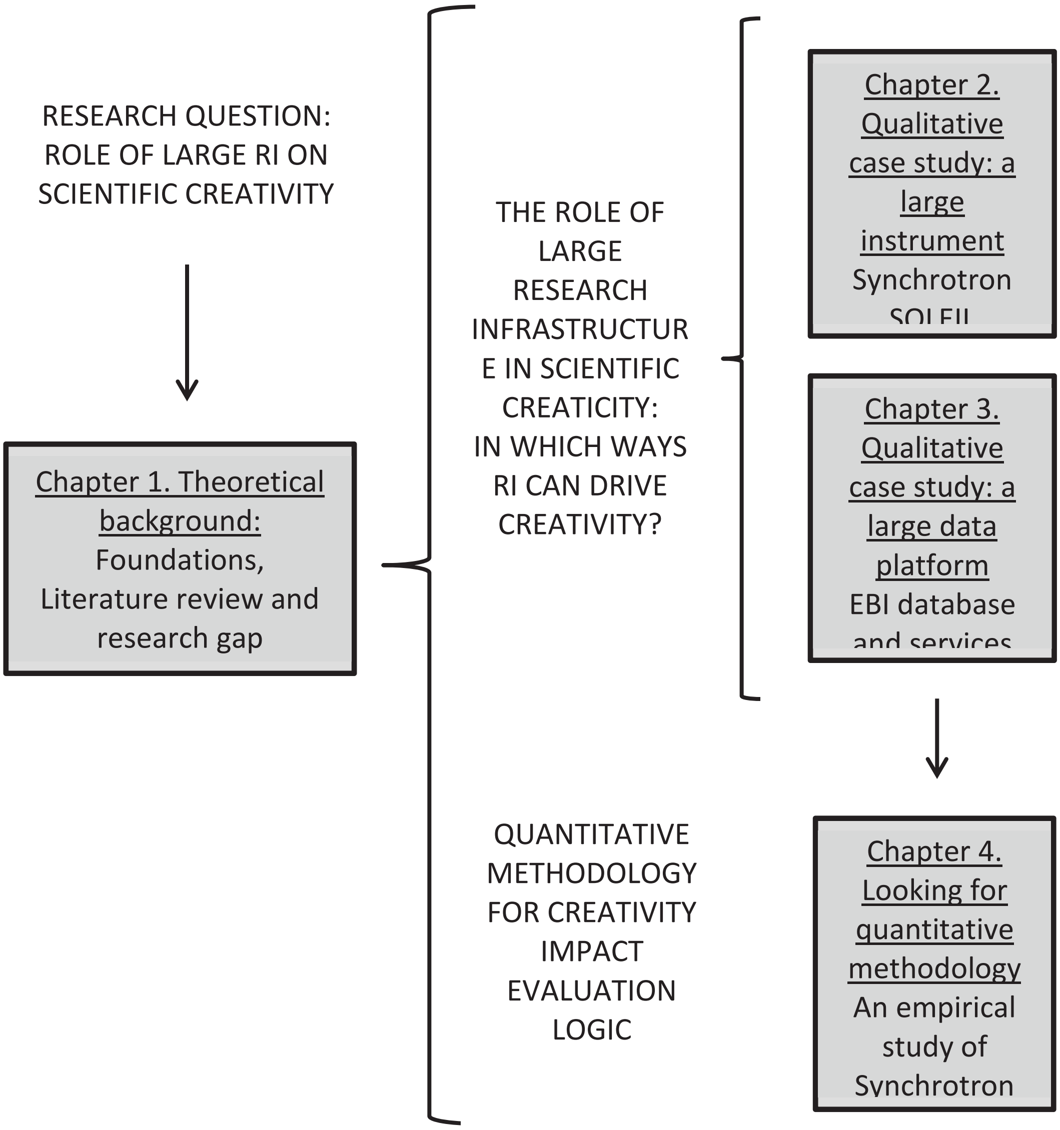


Figure 2 Structure of the thesis

Chapter 1: Context and Theoretical Background

1 Introduction

The way science is done has evolved through the centuries. It is difficult to set a starting point on when humanity started doing science since peoples have been observing and documenting the world for millennia. It is common, though, to situate the starting point of science during the 6th to 5th century BCE as the collection of knowledge started to be done in a systemized and organized way and it included a set of rules and codes to be followed. Let us look at the Pre-Socratics, at Ancient Greek. They did not create knowledge based on the observation of nature, but they did create theories which aimed to explained why things are the way they are and believed that nature is ruled by discoverable laws. The objective was to separate myth from truth and the main method was rational debate. It was at this period that mathematics started developing as something that had sense by itself instead of a tool to describe the world (i.e. Pythagoras). Later, observations, and not thought by itself, became the bases of knowledge creation. Aristotle, for instance, believed in the empirical observation of the natural world as the basis for the creation of general laws that explain it. This order of operations is at the heart of modern scientific practices.

If we focus on modern science, and more particularly on natural and fundamental sciences, the evolution has been based, not on how we define science and how we define knowledge. Modern science has gone through a few paradigm changes. This evolution of modern science has been based mostly on changes on the process of scientific inquiry. This process has often evolved together with the technology available to observe the world. Many think that science is experimenting right now a change of paradigm that started by the middle of the 20th century, and it has been accelerated during the last couple of decades. This change was first defined as “big science” but the concept is too narrow to include all the changes that science is experiencing today. Some of the most commonly mentioned characteristics of this new way of performing science are the following. Firstly, there is the increase of large-scale long-term projects that aim to solve the biggest scientific and societal challenges. They often need costly technological resources that are built and developed only in order to face those challenges. It is the case of the research performed with the LHC[11] collider at the CERN[12]. Sometimes these large-scale projects are not performed in a single institution using very big equipment, but they require collaboration projects at a very large scale, which include several institutions and dozens of scientists working together in the same project. Despite this growth of big projects, smaller projects still exist, and they constitute most of the science that is performed today. These smaller projects, however, often need more and more big equipment. Indeed, there is a growing dependency of science on large-scale equipment, which is costly and therefore publicly financed and open to all the scientific community wishing to use them. Other big need of today's science is data. There has been an exponential growth in the existing data for the last few decades and with this growth the potential uses of the data have increased as well. Data have always been important, in science, however today the role that these data have taken in the research process is increasing at a high speed (André, 2014; Hey et al., 2009).

Because of this growth on large-scale science, the use of large instruments and the need for a big mass of data, public institutions have been investing in financing those projects, instruments and databases. One of the objects used in science policy in order to provide science with the resources that today's challenges demand is Research Infrastructure (RI from now). RIs are publicly funded in Europe. Since the early 2000s, the development and the coordination of large RIs have been increasingly recognized by the European Commission as an essential pillar of the building of the European Research Area (ERA). With the ERA, the European Commission pretends to create a competitive and dynamic economy which is based on the creation of knowledge and it is to this goal that RIs have become the priority of European Science Policy. RIs are expected to help science provide society with breakthrough scientific discoveries.

Discoveries and the creation of scientific knowledge are strongly dependent from creativity. Creativity is, indeed, thought to be at the bases of knowledge creation, generally speaking. The study of creativity has, traditionally, been pursued by psychologist and the most common examples used to identify cases of creativity come

---

[11] The Large Hadron Collider (LHC) is the world's largest and most powerful particle collider

[12] CERN (derived from French "Conseil européen pour la recherche nucléaire") is a European research organization that operates the LHC.

from science. Creativity has traditionally been defined as crucial for understanding the creation of new knowledge. The process that Darwin or Einstein went through in order to achieve their theories is often used as examples of how creativity is done. Several disciplines have looked at scientific creativity and from a multiplicity of points. Psychology was one of the first disciplines to study creativity when trying to understand which cognitive traits characterized creative people. Sociology on the other hand has focused on the study of social structures and processes of scientific activity that impact creativity. The study of all this literature, together with the study of the empirical context, will allow us to understand to what extent these new ways of doing science may affect creativity and therefore, knowledge creation.

More specifically, in this thesis we aim to understand whether the new ways of organizing and performing science via the use of RIs and, and more generally, via the use of large instruments and databases, allow science to keep being creative or even to be more creative than before. To do so we will look at two different empirical contexts, a large RI providing instrumentation and an RI providing large databases, in order to understand how these RIs work and which elements of RIs can lead us to expect that they have a positive impact on the creativity of their users. We will also study the literature on scientific creativity in order to establish a theoretical framework appropriate to each RI case. This first chapter constructs the overall theoretical framework of the thesis and it founds an appropriate methodological approach, after having established the interest of the research field (Research Infrastructure as a relevant and interesting place for studying creativity). To do so, the chapter starts (Section 2) by presenting the empirical context of the thesis, that is, the growing importance of Research Infrastructure in science, as well as its descriptive elements. Then (Section 3) it provides a selective literature review on the topic of scientific creativity, which shows how the topic has been studied by different bodies of literature and helps us to identify the most relevant concepts for our analysis. This section is intended to set up the conducting line and the building blocks of our conceptual framework, as well as the research gaps that our work aims to fill. Finally (section 4) the chapter presents the epistemological and methodological approaches of the thesis.

2 Empirical context: Trends in Science and the landscape of Large Research Infrastructures

Scientific activity has been recognised by policy makers as a fundamental driver of innovation and how both drive economic growth. However, not only science is at the origins of innovation, innovation can be crucial for the advancement of science through the development of tools, methods and instruments (Krammer, 2015; Rosenberg, 1982). Research Infrastructures lie among these elements and they have a growing importance in science with several disciplines being dependent on them. There are as well more and more large-scale projects that depend on RI. Indeed, capital-intensive research projects are developing at rapid pace in number, size and cost. In some research domains, the imperatives of the science itself requires the creation of large infrastructures and for some research domains there is simply no other way to conduct the needed experiments, observations or computations than the use of RIs.[13] Furthermore, society is confronted with global-scale challenges that demand innovative, science-based solutions in areas such as health, energy, climate change-fighting and food security. All components of the research landscape are being solicited to address such global scale challenges, sometimes with efforts that are on the same vast scale as the challenges themselves[14]. The public investment in RIs is growing in order to answer to this call as well as to satisfy the demand of users. It is because of this growing importance of RIs for science and because of its particularities that in the past few years, there has been a growing body of literature, especially impact assessment reports, that provides with methods and tools to evaluate RIs and studies the role of RI in science (Donovan, 2011). This literature is, however, still limited and creativity is an aspect that remains disregarded in impact evaluation of RI. This reason leads us to empirically study how RIs can facilitate scientific creativity. To do so we will start by giving some descriptive elements of RIs to show their uniqueness as well as the properties that suggest that they are likely to be a proper and highlighting place for the study of creativity.

---

[13] OECD Science Forum Report on Establishing Large International Research Infrastructures: Issues and Options http://www.oecd.org/science/inno/47057832.pdf Consulted Dec 11, 2018

[14] OECD Science Forum Report on Internation Distributed Research Infrastructure.: Issues and Options http://www.oecd.org/sti/inno/international-distributed-research-infrastructures.pdf Consulted Dec 11, 2018

2.1 The future of science: a growing dependency on big equipment and data

Science depends the more and more on large and costly equipment such as RIs. The scientific advancement has always been to some point set by the advance on instruments and the ability to do good quality research in certain areas by the access to those instruments (Stephan 2012). Today these instruments often reach a cost that makes them inaccessible for single universities or laboratories. Because of this reason key instruments and other kind of equipment are the more and more done under the form of RIs, meaning at a large scale, by means of joint projects or financed by (inter)national funds. They also follow the more and more a logic of openness and are accessible to all researchers that require it. Additionally, RIs but also laboratories are the more and more asked to participate in large collaborations and endorse into ambitious collaboration research projects. All of this gives RIs a growing role, which is an additional reason for exploring how these kinds of objects impact scientific creativity.

Data and the associated RIs are also becoming the more and more important in science and are believed by many to become central for science in the coming years. Science has always produced data. Very early the observation of the surrounding world led to the production of data. Counting, measuring and keeping have always been important tasks in scientific activity. As an example, in the field of astronomy, the successive inventions of different tools such as the astrolabe, the quadrant, the telescope, the spectroscope have advanced towards more precise, richer data, but also more numerous and more voluminous. With recent instruments for measurement (such as satellites and large telescopes) the increase of the amount of data has grown exponentially (André, 2014). It is the same in other fields of research which have seen in recent decades their practices disrupted by the development of instruments able to produce mass data: genomics and its sequencers, climate and environmental sciences and multiple sensors at land, sea, air and space are just some examples. It is true also in social sciences, with the growing data found in the web and social networks. Together with big opportunities this mass production of data poses some technical challenges. As André (2014) explains, the rapid growth of data challenges the technological ability to store and maintain this data as well as to offer it to the scientific community. In the case of voluminous, reference data sets such as the entire human genome, there is a need for cutting edge technology, expert human resources and thus very large amounts of funding. Here again, RIs and more precisely e-infrastructures tend to play a crucial role to provide the researchers with open access to high quality databases which are very costly to maintain and require from very large storage capacity.

The society is moving towards a data-driven economy and this change could be considered as a new industrial revolution. This is true particularly for science, where the development of data collections and processing tools started a few decades ago. This is being the more and more accelerated with the development of storage and access infrastructure under an openness principle. Several scholars agree on the fact that this logic will be generalized to all the domains of great challenges. The development of these data infrastructures and its openness is believed to have the potential of meeting several societal changes in the areas of health, environment, energy and innovation. They are also the predecessor of similar societal changes in domains different from research and they are hoped to create wealth and employment[15]. The provision of data from research, as well as the associated processing and visualization tools, will participate obviously to the development of this economy based on the data.

2.2 Descriptive elements of RI

Because RI is a complex and multi-face object, it is difficult to reach a comprehensive definition. We will adopt the definition given by the European Commission[16] on its website and on ERA and ESFRI[17] documents. The term "research infrastructures" refers to *"facilities, resources and services used by the science community to conduct research and foster innovation. By pooling effort and developing RIs, European countries can*

---

[15] European Commission, 2014. Towards an Economy of data

http://ec.europa.eu/transparency/regdoc/rep/1/2014/FR/1-2014-442-FR-F1-1.Pdf, consulted November 28, 2018

[16] http://ec.europa.eu/research/infrastructures

[17] ESFRI stands for "European Strategy Forum on Research Infrastructures" and is a strategic instrument to develop the scientific integration of Europe and to strengthen its international outreach. The competitive and open access to high quality Research Infrastructures aims to support and benchmark the quality of the activities of European scientists, and to attract the best researchers from around the world. ESFRI operates at the forefront of European and global science policy and contributes to its development translating political objectives into concrete advice for RI in Europe (https://www.esfri.eu/objectives-vision)

achieve *excellence* *in highly-demanding scientific fields and simultaneously build the European Research Area (ERA) and Innovation Union. They include: major scientific* *equipment,* *resources such as collections, archives or scientific data,* *e-infrastructures such as data* *and computing systems, and communication networks. RIs can be* *single-sited* *(a single resource at a single location), distributed (a network of distributed resources), or* *virtual* *(the service is provided electronically)".* RIs include large scale research instruments (such as particle accelerators and telescopes), collections, depositories, public repositories (for example insect, mice or grain repositories), libraries, databases, biological archives, networks of computing facilities, research vessels, satellites and aircraft observation facilities, coastal or natural observatories, etc[18]. Let us study this definition in order to compile the most relevant aspects.

- From this definition we understand, firstly that there is a variety of objects that we can consider as research infrastructure since they are described as "*facilities, resources and services*". This variety comes with a complexity in its study. If every single RI has a different nature and therefore different characteristics, they probably need to be studied individually. This is one of the reasons why we focus on the study of complex and dynamic processes and we use qualitative methodology.
- As we can see in the definition, RIs are "us*ed by the science community*". RIs follow a logic of openness, the research conducted at RIs is not only performed by an internal team, but it is open to the scientific community. This makes of RI very particular places and suggests that they can be places that enable the encounter of different individuals which, as we will see later in this chapter, can enable creativity.
- The idea that RIs might be suitable for the study of creativity is also suggested by the definition of RI by itself, as they are meant to contribute to "*achieve excellence in highly-demanding scientific fields*". Excellence, in science, refers to the production of knowledge of good quality that will help face societal and scientific challenges. This requires creativity and therefore the seek for excellence is a reason to expect RIs to be a suitable place for the study of creativity

---

[18] For a detailed list of European RI and a European RI map:

http://ec.europa.eu/research/infrastructures/pdf/ri_landscape_2017.pdf#view=fit&pagemode=none

- Finally, when describing the different characteristics that RI might have and the different kinds of RIs we can find it mentions big equipment and e-infrastructures such as data. Because of the increasing relevance of these two kinds of object in science, we will study particularly these two by looking at the cases of one single sited equipment and one e-infrastructure which is virtual and consists on the provision of data.

Research Infrastructure is a complex and multi faced object not only because of the large variety in its typology but also because each RI type presents a large variety of roles and objectives[19]. RIs, however, have many common features. They all are organizations that provide the researchers with services or access to a technology, but also conduct research by themselves. Indeed, RIs have scientists working and performing research there, usually this research consists in long term projects aimed to answer fundamental problems of today's science. They also cooperate with other RIs for technological development and try to stay up to date on the last available techniques and technologies. However, RIs main role is to provide users with services going from the maintenance of a database or an archive to the maintenance and technical assistance in the use of instruments. Because of this double role RIs can, on the one hand, accumulate technological expertise, and on the other hand, share this expertise with external users. Because of this we believe that RIs have the potential to foster creativity. Another characteristic is their openness, RIs are open to all[20] external users, this gives them a collective dimension which, as we will observe later on this chapter, is as well often associated to creativity.

In other words, large-scale research infrastructures are defined as those facilities with many or all of the following features: large research capacity, trans-national relevance, requiring sizeable investment and, generally, having high operating costs. They may be unique or rare, and have a set of peculiarities that distinguishes them from other objects in the research landscape. RIs provide with the latest cutting-edge technology and have a consequential impact on science and research at both the global and European level. RIs

---

[19] OECD The impacts of Large Research Infrastructures on Economic Innovation and on Society, http://www.oecd.org/sti/inno/CERN-case-studies.pdf

[20] If they produce public research and with a selection of projects through excellence when the time or space are limited.

rely strongly on a multiplicity of types of collaborations and often bring together several individuals and institutions. A final crucial element of European RIs is their free access. When performing public research (e.g. research leading to an academic publication), laboratories and other institutions do not need to pay for the use of the RI. In the case of large instruments, access is based on research quality. In the case of databases, one can talk about entirely free and open access for any potential user.

Considering the size of their investment in RIs, policy makers naturally ask for impact studies to assess and monitor the activity of RIs. Research Infrastructures require to be evaluated in a way that is adapted to their peculiarities. As mentioned above, there is a diversity of types of resources combined with a wide variety of missions, a variety of the disciplines that they are used for, as well as of locations and access methods. Moreover, the permanent dimension of RIs necessarily brings different impacts than those brought by other R&D policies. The criteria used for the classification of the impacts is, however, very similar to the general ones used for the evaluation of other S&T policies. The focus is not put only on the science that is performed at that RI but on several categories of impacts. For instance Zuijdam et al. (2011) will distinguish between Scientific, Economic, Human Capital and Societal impacts. To this classification several authors add the dimensions of direct and indirect effects and the geographic dimension commented above (Ilbeigi, 2017). Building on top of the BETA method mentioned above, the BETA-EvaRIO[21] method was developed for the specific case of impact Evaluation of RIs. This methodology classifies impacts into four categories of effects: direct, indirect, performance and capacity effects, all assessed through the consideration of several metrics and indicators. When it comes to the impact of RI on science itself, the focus is on capacity effects, i.e. increasing knowledge in science and technology, management and organisation as well as reputation and networking capacity. The former includes indicators based on number of publications, number of patents as well as number of PhD thesis done using the RI. The latter includes information on co-publications as well as co-patenting and other kinds of collaboration. In all cases there is an accent on the gain in excellence due to RI use, by using information such as impact factor (IF) and number of citations (Bach and Wolff, 2017).

---

[21] http://evario.u-strasbg.fr/beta-method

Creativity, and especially its underlying mechanisms, is an aspect that remains disregarded in impact evaluation of RIs, as it does when it comes to impact evaluation of science policy in general or the allocation of resources in science. The reason why this happens is twofold. Published research is meant to be new and valuable. Therefore, measuring impact through publication counts and impact factor is expected to be equivalent to measuring creativity. This is however not the case. As we will see in section 3 below, creativity is desired but often not properly encouraged. This happens in several domains of society and is particularly true in the case of science. Next section will develop this aspect that will remain central all along this thesis.

3 A selective literature review on the notion of creativity

In order to better understand how RIs can facilitate the creativity of their users, this section explores the notion of scientific creativity in general, as characterised by important scholars of the field. What is scientific creativity? How is it produced? By who? When does it occur (under which conditions)? A comprehensive review of literature about all these questions would be well beyond the scope of the thesis. Our objective is to establish the theoretical framework and to present the areas of inquiry of this thesis.

We first define and discuss the concept of scientific creativity as well as the main elements of its definition, novelty and value. Then, we look at different aspects of the scientific creative process, such as the underlying mechanisms, the drivers and the (more or less) creative outputs. The creative actors and the conditions under which creativity occurs are also topics of interest. In order to provide insights on these issues, we analyse the most relevant findings from different scientific domains We focus particularly on three authors with great contributions to the topic and whose findings have been the basis upon which the study of creativity in science is built: Arthur Koestler, who introduces the idea of bisociation, Dean Keith Simonton who discusses the main factors behind individual creativity and Teresa Amabile who highlights the key role of environment and organisational conditions, rather than the cognitive characteristics of creative individuals. This is followed by some complementary and relevant contributions of management literature to the notion of collective creativity in the case of innovation processes. The section ends with a positioning of how this thesis considers creativity.

3.1 Creativity in science

This section studies the different bodies of literature from which we pull the building blocks of our theoretical framework. To do so we start by the study of creativity in science as such, which elements are involved, which mechanisms drive it and which environments favour it. Because the literature studying creativity in science has some limitations, we study in a second part of this section, the literature on management of creativity. We finish the section by positioning the thesis and assembling the building blocks of the thesis.

3.1.1 Scientific creativity as the production of new and valuable knowledge

Similarly to other domains of knowledge, creativity in science has been described as consisting in the production of knowledge and capabilities that are **new** and **useful** (Hollingsworth, 2002) or **valuable** (Amabile, 1988). This definition is widely used by the literature on scientific creativity. Its main terms deserve some further reflexions. We choose Amabile's definition and consider creativity in science as *"the production of knowledge and capabilities that are new and valuable"*. Knowledge needs to be new, in order to be creative, but it also needs to have some value which means that it must be considered as useful by someone. These two notions, of novelty and value, are very important for the understanding of creativity and will be present all along the thesis.

3.1.1.1 The production process of science

Science, defined as the production of knowledge and scientific activity, has been described by many scholars as a process where existing knowledge is combined in order to create new knowledge (Koestler, 1964; Simonton, 2004; Stephan, 2012). More specifically we can consider that this new knowledge is the outcome of a process of scientific inquiry, itself considered as a problem-solving activity built up by combining different elements such as concepts, perspectives or methods, in order to find answers to research questions (Simonton, 2004). This knowledge is to be transformed into a scientific publication (journal, book or other) and validated by the scientific community by means of a peer review process where the originality and relevance of the knowledge produced is evaluated (Spier, 2002; Voight and Hoogenboom, 2012). Scientific activity produces as well new capabilities such as skills, know-how and methodological development. Indeed, the production process of science is not completely linear. And in the middle of the process while using methods and techniques some new methods and techniques are developed. These, and not only the formulation of a theoretical law, are considered as a scientific output as well (Heinze et al., 2009).

Additionally the literature that studies science presents an ambiguity regarding the study of creativity, as often the words productivity and creativity are used with the same meaning and authors differ regarding their preference for one or another (Csikszentmihalyi, 2014, pp15-19) The origin of this ambiguity lies on the definitions of science by itself. What we observe with these definitions is that science, to be considered as such, requires to have at least some degree of novelty. It requires to have as well at least some degree of value. To be published, new knowledge needs to be relevant, which means that it has some worth, merit or importance to at least someone. This raises some question as the terms used to define science in general (originality and relevance) are equivalent to those used to define creative science (novelty and value). Can we consider that all science is creative? The literature review below provides us with some insights, more particularly the part dedicated to the work of Simonton (2004) where he distinguishes between big Creativity and small creativity.

Another question that has been raised by the literature regarding the definition of scientific creativity is whether the terms used to define it, that is to say, novelty and value, are compatible with each other. An important tension or even conflicting forces have been found between these two dimensions of creativity. The next subsection is dedicated to the study of that tension.

3.1.1.2 Novelty versus value: a tension between the constitutive dimensions of creativity

Despite being highly beneficial for the society, being creative in science is also tricky as the two criteria used to define creativity pull research into opposite directions. As explained by Csikszentmihalyi (2014) and Csikszentmihalyi et al (1995), there is the need to convince the guardians of the domain that the idea is creative. However, the generation of ideas that have a high degree of novelty entails going through a process in which these ideas are put at risk to be rejected because they tend to be considered as bizarre, inappropriate, unlikely or risky (Staw 1995). Creative behaviours have been defined as risky ones by several scholars (Carver & White, 1994; Keltner et al., 2003; Lee et al., 2004; Mainemelis, 2010). Novel research has a potential for a higher impact, but it also faces a higher uncertainty. Indeed, the potential impact is higher but so is the probability of a low impact (Heinze et al., 2009). Moreover, the impact of novel research might arrive with some delay, as might do the recognition of this impact by the community. Important scientific discoveries are often not well appreciated at once by the community.

It is well known that in science recognition often arrives with a certain delay. Delayed recognition is caused by different phenomena and it consists in a longer time to integrate the findings of radically novel research than incremental one (Garfield, 1977). One of the phenomena that lead to a delayed recognition is caused by the prematurity of some novel research. Research is considered premature when, because of the novelty, there is very little for other scientists to build up on. *"A discovery is premature if its implications cannot be connected by a series of simple logical steps to generally accepted knowledge"* (Stend, 1972). Prematurity, among other factors, results in delayed recognition. Another reason for delayed recognition consists in the resistance that the incumbent scientific paradigms show when new paradigms emerge (Kuhn, 1962, Merton 1973).

Science, to be published, is judged by peers and to do so they usually have three criteria (Heinze et al., 2009; Wang et al., 2017): plausibility, validity and originality. Because of this, creativity in science suffers from the known problem of a tension between the criteria of plausibility / validity and the criteria of originality. This often results into scientists choosing conformity instead of dissent. By choosing originality and novelty at the expense of plausibility and validity, they would defy some established rules, and thus take a risk. For this reason, scientists may decide not to assume that risk, choose to perform less novel research and continue exploiting research paths that were already open. This issue raised the concern of several researchers, as well as authorities, who think that science might become less and less creative, so that the advancement of the knowledge frontier is endangered (Heinze, 2009).

3.1.1.3 Conclusion

As discussed above, the definition of creativity points up two key dimensions of creativity, namely novelty and value, which go to opposite directions but have to be both present for a scientific production to qualify as a creative one. But some questions remain open at this stage of reflexion. We need to explain the creative process, i.e. the mechanisms lying behind the production of new and valuable knowledge. We also need to explain who is behind this production of knowledge. Can everyone be creative? Is it the property of some unique, gifted individuals? And if yes, which ones? Or is it the property of groups, teams or entire organizations? Finally, we do not know what makes this creativity occur. The next subsections of section 4 look at the literature in order to see how these questions have been treated. Trying to concentrate on the most relevant research on the field, we identified three interesting aspects studied by literature when it comes to scientific creativity. First, one of the main mechanisms that allows new and valuable ideas to emerge seems to be the **combination of knowledge**. Here we focus on the work of Arthur Koestler, whose idea of bisociation is recognised as one of the most important contributions to the study of creativity. Second, we explore the elements behind this mechanism of bisociation. There are two bodies of literature, one that appeals to chance and another one that appeals to the idea of geniuses, namely individuals who have a talent for science. We use Simonton's synthetic view, who adds two additional factors: zeitgeist and logic. Finally, Teresa Amabile shows the relevance of everyday creativity that can be achieved by everyone. Her conception of creativity does not depend on the individual’s intellectual ability but rather on his motivation. And motivation depends on the organizational conditions under which the researcher works. This view has been very influential, and it is present in all recent works on scientific creativity.

3.1.2 Combining knowledge, the work of Arthur Koestler and the mechanism of bisociation

The first question we asked ourselves while studying the definition of scientific creativity is: "which mechanism is behind creativity?". Today it is widely accepted to consider the production of science as a process by which scientists combine different pieces of knowledge in order to create new knowledge (Romer, 1994; Simonton, 2004; Weitzman, 1998). How this process of combination is done will determine how creative the output is. In order to produce creative science, it is often necessary to combine multiple pieces of knowledge that have not been combined before or that are not very often combined. Although he was talking about innovation, Schumpeter (1934) already introduced the combinatorial perspective at the centre of novelty as he defined innovation as a "new combination". Nelson and Winter (1982) argue as well that new combinations of concepts or of materials that already exist are the bases of novelty in art and science. When focusing on creativity, the idea of recombination was first developed by Arthur Koestler through the concept of bisociation.

3.1.2.1 Arthur Koestler and the mechanism of bisociation

Arthur Koestler is one of the first to theorize creativity and to try to develop a general theory that explains it. To do so he introduces the concept of bisociation, which he understands as being at the centre of the creative act. **Bisociation** consists in putting together two frames of reference that had been, up to the moment, never united. The concept of bisociation defines a way of combination, which is one of the main components of the production of knowledge. Koestler insists on the idea of unexpected combinations and he focuses more on the novelty side of creativity. He considers as well creativity as inherent to individuals and provides with several examples of known inventors and scientists.

Koestler bases his analysis on the study of big scientific discoveries and he defines the existence of a common structure to all the acts of creation, regardless of the field where the acts happen, where bisociation is at the centre. He uses three main cases, which are humour, scientific inquiry and art. Koestler's fundamental idea is that any creative act is a bisociation (which is more than mere association) of two (or several) supposedly incompatible frames of thought. Koestler also calls those frames of thought matrices, and defines them as “any ability, habit, or skill, any pattern of ordered behaviour governed by a 'code' of fixed rules”. These matrices include routines, rules and codes that determine the behaviour of individuals as well as the conceptual frame of thinking. Koestler argues that the diverse forms of human creativity all correspond to variations of this model of bisociation. In jokes the audience is led to expect a certain outcome compatible with a particular matrix (e.g. the narrative storyline); there is, however, a punch line, that replaces the original matrix with an alternative and produces the comic effect. In scientific inquiry, the two matrices are often fused into a new larger synthesis. The recognition that two previously disconnected matrices are compatible generates the experience of “Eureka”. Finally, in arts the two matrices are held in juxtaposition to one another. Observing art is a process of experiencing this juxtaposition, with both matrices sustained (Koestler, 1964 p.45).

Koestler uses several examples of the history of science and technology in order to illustrate the idea of bisociation of two frames of thought. Darwin's contribution, for instance, has its originality lying in picking up some disconnected frames of thought and integrating them in a new one. The main frame was the evolutionist's belief that the various species in the animal and vegetable kingdom had not been independently created, but had their origin as descending varieties of previous species. This idea, however, gave no explanation of the reasons that caused the common ancestor to turn gradually into different kinds of animals. The second frame used by Darwin was the domestic breeding achieved by the selective mating of favourable variations. He found the last element in the frame of Malthus's “An Essay on the Principle of Population” in the idea of competitive struggle. From this new combination (bisociation) of three different frames of thought, Darwin started constructing his theory, which was completed by an extensive accumulation of experiences and fieldwork afterward (Koestler, 1964 p.140).

Another example is the Gutenberg's invention of the printing press. He combined two frames of thought, the first one being the seal, where a text was “carved” and could be use several times, by applying pressure, in order to print a symbol on a paper. Yet, this was not enough as the process of production of the seals was still laborious and the printing slow. He took part in the wine harvest and saw the power of the press. He thought about it and realized that the same steady pressure could be applied by a seal or coin preferably made of lead, which is easy to cast on paper, and that owing to the pressure, the lead would leave a trace on the paper. This was his “eureka!” moment (Koestler, 1964 p.123)

As we have seen in the mentioned examples, the act of bisociation is very often accompanied by some field experience. Experience makes it possible to create links and connections that will result into the foundations of a new frame of thought with its own rules. Putting together different frames of thought does not result into a random combination of concepts, codes and rules, but rather into building a completely new one that makes sense by itself to the point of not being able to consider them separately anymore (Cohendet, 2016). The framework for the understanding of scientific creativity proposed by Koestler is still used today and it was the basis on which several other works on creativity have been constructed.

3.1.2.2 Building up on the idea of rare combinations

This combinatory aspect of science has been recognized by several other authors (i.e. Klahr and Simon, 1999; Merton, 1975; Simon, 1977; Simon et al., 1981) who describe science as a combinatory activity that integrates different perspectives, methods and concepts. New knowledge, in order to be creative, has to be the result of novel, unexpected and relevant combinations. Recent works focus on the idea of novelty of these combinations. Neumann (2007) explains that scientists create knowledge by bringing together previously unlinked ideas that allows them to generate a new concept. Others, rather than looking at absolute novelty, prefer to focus on very uncommon or rare combinations. Uzzi et al (2013) study the long-term impact of articles including very rare combinations (of knowledge, ideas, concepts, etc.). Lee et al (2015) study the impact of team composition on the novelty of scientific results. They find that the wider the variety of skills present in the team, the more creative the knowledge created will be. Building on the aforementioned work, Wang et al. (2017) continue viewing science as a combinatorial process of previous knowledge and they measure novelty in science by examining whether a published paper makes first-time-ever combinations of referenced journals. They found that highly novel papers, defined as those using new combinations, deliver results that have a higher impact in the long run.

The question that rises up to this point is: "how do these valuable new combinations appear"? Does creativity come from some cognitive characteristics of scientists? Is it due to a favourable environment? Is it a matter of chance? We can find useful elements of answer in the works by Dean Keith Simonton, who explains the role of different elements in scientific discovery. These elements are: chance, logic, genius and zeitgeist. He thus focus on **how** scientists find the frames of thought that, put together, will construct a whole new frame holding a whole new body of knowledge.

3.1.3 From genius to chance: the works of Dean Keith Simonton

Up to now we have seen that scientific creativity is studied mainly as inherent to individuals. One may ask, who triggers the occurrence of those new combinations? Are individuals all capable of doing them, or only a few of them? Many scholars in psychology have studied the conditions that trigger the rare combinations leading to the creation of new knowledge. In line with this discipline, a large body of this literature focus on the psychological peculiarities of very creative individuals. Using big scientific breakthroughs as an analytical point of departure, these authors consider that some individuals have a natural talent for discovery and they analyse the psychology of those individuals. Others consider creativity as a matter of chance or good luck. Dean Keith Simonton contribution focuses on both aspects and proposes a framework that additionally to these two, adds two additional determinants of scientific creativity (i.e. chance, logic and *zeitgeis* in addition to genius*)* as well as their respective influence on "small creativity" versus "big Creativity".

3.1.3.1 Dean Keith Simonton: chance, logic, zeitgeist and genius

In his book about scientific creativity, Simonton aims to analyse scientific creativity from multiple points of view and summarize research done, up to the moment, on the topic. The first relevant contribution of his book is to consider two different kinds of creativity, which he calls small and big creativity, or **small c** and **big C**. Big creativity would occur in the form of big scientific discoveries such as those that are awarded with the Nobel prize. Small creativity, on the other hand, would include every-day problem solving and smaller scientific advances. Big creativity would consist on "eureka" moments whereas small creativity would consist on "aha" moments. Although he discusses the sometimes-collective aspect of creativity and provides with some interesting figures about creativity within scientific communities and across disciplines, most of his work focus on how individual creativity works and he does that by looking at the cognitive particularities of the scientist himself but also by studying the context to which the scientist is exposed to. For Simonton (2004) creativity can be viewed from four main perspectives which are: chance, logic, zeitgeist and genius. He dedicates a chapter to each of these aspects.

Simonton starts by stating that **chance** is the primary basis of scientific creativity. The first argument he uses to defend this position is the fact that the distribution of occurrence of scientific discoveries has the typical features of a completely randomized probabilistic process. He reaches the conclusion that scientific discovery and therefore creativity approximates to a random combinatorial process (Simonton, 2004. p.41). Whether he looked at the scientific output at the peaks on the careers of scientists, at the output of groups of scientists or at the emergence of new relevant paths within scientific communities, the randomized distribution seemed to be always present. Although one could consider that chance should not be considered a cause of scientific creativity because it is not a variable or a factor, Simonton argues that this prevalent role of chance is compatible with some deterministic factors. He argues as well that chance hides all the complexity that we do not know and therefore creativity in science should be studied through other factors as well.

The second aspect of creativity discussed in the book is **logic**. Simonton argues that logic can only be considered as having a secondary role in scientific creativity. Creativity is supposed to lead to the unexpected and it is therefore the opposite of where logic leads. Using logic is helpful to solve routine problems and it is at the origins of some "aha!" situations rather the "eureka!" ones. He argues as well that because the most challenging scientific problems are often illogical or inconsistent in the sense that "*they cannot be noticed without noticing the trick or the catch*" (Simonton, 2004 p.164), these problems are strongly resistant to a purely logical analysis. According to Simonton logic is only useful in small problems and it is therefore a cause of small creativity but not of big Creativity. This does not mean that logic is not at all present in big C as big scientific breakthroughs build up on smaller scientific advances where the process has been sustained by logical thinking. Logic, however, can be as well a constraint for scientific creativity. Logic establishes the paths that are most likely to lead to a solution, it imposes norms and ways of thinking and therefore it prevents divergent thinking. In other words, logic would impose some ideas and given concepts would enter the combinatory process rather than others. This might lead to skipping the unlikely combination that would result in an unexpected creative result. Logic would, very often, drive the researcher to think within the borders of an established theory or paradigm. If we think about the frames of thought of Koestler and the examples that he gave, logic would not allow bisociation to occur. Within logic, two different scientific frames of thought are not supposed to coexist. Let us look at the case of Darwin and his theory of Evolution, if he had followed logic, he would not have used the theory of Malthus, which aimed to understand human population, together with the ideas about evolution of species. Logic is, therefore, useful for small c but it can be detrimental to big C.

The third aspect studied by Simonton is ***Zeitgeist*,** which is strongly related to logic as it consists in the body of rules that set what is logical and what is illogical. The disciplinary zeitgeist consists of the phenomena, facts, variables, definitions, theories, laws, questions, goals and criteria that constitute a scientific domain in a specific moment of time. These ideas constitute the basis on which discovery is supposed to be built up on. Each discovery, once validated by the discipline, will be integrated in the disciplinary zeitgeist that is in continuous change. Because, within Simonton's framework, big discoveries occur almost by mere chance, there is little that the disciplinary zeitgeist can do in order to favour scientific creativity. The only thing that can be done is to encourage new trials and to have some openness and flexibility in its rules in order for new discoveries to be integrated faster. Simonton highlights here the previously defined tension between the different criteria needed for creativity. The disciplinary zeitgeist has the role of assuring the validity (or value) of new knowledge but by doing so it can impose some restrictions to novelty. Indeed, he highlights how the different norms and rules are often proven to be erroneous and counterproductive for the development of new knowledge (Simonton, 2004 p.92). In addition to the disciplinary zeitgeist there is the sociocultural one, which consists of the political, economic, social and cultural circumstances that encroach the scientific activity from the outside. For instance, some sociocultural conditions would encourage individuals to become scientists while others might discourage them. Some sociocultural conditions would encourage risk taking and following one's gut, while others would discourage it and seek for conformity. The sociocultural context might as well determine the access to instruments, infrastructure and simply financial support. For Simonton cultural zeitgeist is relevant for the speed and frequency of discovery, but it cannot have a direct impact on the discovery by itself.

The last of the aspects studied is **genius** and it consists on the personal ability of a scientist to be creative. Although he considers chance as the bases of scientific discovery, Simonton shows that a set of personal characteristics have as well a high impact on scientific creativity. He finds, indeed, that the level of creativity of a scientist is correlated with certain cognitive traits such as capacity to generate multiple and diverse associations of ideas, defocused attention and receptiveness to novelty, variety, complexity and ambiguity (Simonton, 2004 p.103-120). These traits are more needed for big C than for small c and their impact also varied from one discipline to another. In highly paradigmatic disciplines they are not highly needed while in more revolutionary disciplines they are crucial. There is, however, an important trade-off between these abilities and the ability to be more conservative and follow the establish rules. As creativity is defined not only in terms of novelty (i.e. novel associations), but also in terms of value, one has to stay on a certain way near the known references in order to be accepted by the peers. Education and training are considered important as well by Simonton as they contribute to the previously mentioned intellectual qualities. Finally, he shows that creative scientists work simultaneously in a variety of scientific projects and are involved on a variety of professional tasks and activities.

To sum up the findings of Simonton, he shows in first place that scientific discovery depends heavily on chance. He finds as well that there is a continuous trade-off between novelty and value at different levels of creativity. Logic is necessary to validate knowledge and to keep rigor in science but it constraints novelty. The same happens with cultural and disciplinary zeitgeist. Finally, he defines the personal attributes that contribute to creativity (i.e. capacity to generate multiple and diverse associations of ideas, defocused attention and receptiveness to novelty, variety, complexity and ambiguity), but these have to be present only up to a certain level as science needs to be validated by disciplinary rules and proven valid. These ideas, although integrated into a framework by Simonton, were already present in the literature on psychology (for the notion of genius) and sociology of science (for the notion of chance). These bodies of literature offer some insightful ideas for the understanding of creativity.

3.1.3.2 Traditional focus on talent: creative science as an activity of geniuses

Simonton is not the first one to have looked at the idea of genius. Many early studies on creativity were based on the idea of some individuals having a natural talent for creativity. Indeed many think that some individuals are naturally gifted to be creative (Gruber and Barrett, 1974). It is the literature in the discipline of psychology that has contributed more to this idea. It is, for example, the case of Guilford (1950), one of the pioneers who have studied the intellectual abilities favouring creativity. More specifically he finds that having a fluid and flexible thought allows individuals to move easily from a group of ideas to another, which favours problem analysis. He shows as well that divergent thinking has an important role for creativity because it allows individuals to seek and use alternatives. Divergent thinking consists on being able to generate multiple answers when trying to solve a problem, contrary to convergent thinking, which aims to find a single solution to a problem[22].

Another relevant focus of psychology literature is the potential relationship between creativity and intelligence. This body of literature highlights the importance of the relationship. Researchers are interested not only in whether creativity and intelligence are related, but also in how and why. Empirical evidence shows that there is a positive (although moderate) relationship between creativity and intelligence (by using IQ). There are multiple theories explaining this relationship. Several studies consider intelligence as a necessary yet no sufficient condition for creativity (Barron, 2012; Guilford, 1967; Wallach and Kogan, 1965). In a similar line there is another influential theory considering that creativity is not intrinsically related to intelligence. What it shows is that individuals are required to meet a minimum level intelligence in order to gain a certain level of education and skills, which then will offer the opportunity to be creative (Hayes, 1989; Karwowski et al., 2016; Shi et al., 2017). Exhibits of creativity are, therefore, moderated by intelligence instead of having a direct cause-effect relationship[23]. Finally several authors consider that creativity depends mainly on curiosity (Hagtvedt et al., 2019; Li et al., 2018) and on critical thinking (Wechsler et al., 2018).

Although in this thesis we do not consider creativity as being an ability of only certain individuals, the research in psychology provides with a lot of insight for the study of creativity. Having a fluid mind which can easily move from a group of ideas to another shows the relevance of being able to take into consideration several ideas. In the same line, divergent thinking and curiosity can be understood as having an open mind towards several alternatives. Finally, research looking for the link between intelligence and creativity shows how intelligence simply allows for the accumulation of knowledge that is required to be creative. Indeed, we consider this idea insightful. Nevertheless, this

---

[22] There are many tests of creativity that try to measure divergent thinking. The most popular of them all is a set of tests proposed by Torrance (1966) known as the Torrance Tests of Creative Thinking (TTCT).

[23] . Sternberg and O'Hara (1990) propose a framework of possible relationships between creativity and intelligence: creativity as a subset of intelligence, intelligence as a subset of creativity, creativity and intelligence as overlapping constructs, creativity and intelligence as parts of the same construct and creativity and intelligence as distinct constructs.

exclusively individual vision of creativity is, in our framework, limited. Science today is performed the more and more in group, we consider it as a collective activity as we will see later on in this chapter, for this reason it is important to complete these ideas with other bodies of literature.

3.1.3.3 Scientific discovery as a matter of chance

Another way to study creativity is to focus on the idea of chance. Research shows that most research and innovation attempts fail to achieve relevant results and that it is complex to understand the reasons that bring some projects to fail and others to be successful (Freeman et al., 1982; Rothwell et al., 1974). This uncertainty is a central issue in research and innovation activities (Arrow, 1962; Nelson and Winter, 1982; Rosenberg, 1992). Additionally, when research manages to reach some relevant results, these are often different from what was expected. Sometimes the scientific results even give answers to different problems than the ones aimed to be solved in first place. The term serendipity has been used to refer to this kind of accidental discoveries (Merton and Barber, 2004; Murayama et al., 2015). It usually refers to finding the solution to a problem when searching for another one. The concept can be extended to include also the finding of the solution to a searched problem, but in a way that was not planned, e.g. through some kind of methodological "error" (Yaqub 2016). What is common to these phenomena is that at some point there is a connection between two elements and an observed effect. Although chance is at the basis of the phenomena, the capacity of observing this effect is very relevant for serendipitous discoveries to happen (Merton and Barber, 2004). Interestingly enough, Yaqub (2016) considers that chance is a subordinate to researchers' skills, knowledge, curiosity and ability to share and communicate information.

What is interesting about this view of scientific discoveries as a matter of chance is that it appeals to the idea of uncertainty. As we have seen previously in this chapter, novelty in science often comes with a high degree of uncertainty. Uncertainty is a major issue which makes scientific activity risky and poses some problems to creative science to be recognized when it is too novel and it differs too much from the current ways of understanding problems.

3.1.4 The impact of the organizational environment: the works of Teresa Amabile

Teresa Amabile was one of the first scholars in psychology to highlight the importance of work conditions for creativity. Contrary to previous literature, she thinks that anyone can be creative. The genius aspect is not the most relevant one and every single individual can be creative. She thinks that chance plays a small role and that the context is crucial for determining the occurrence of creativity. In line with her major contribution, several other authors have analysed as well the influence of social environment and work conditions, which makes it today one of the main area of study when doing research in creativity.

3.1.4.1 Teresa Amabile: the role of intrinsic motivation and work conditions

Before Amabile's work, most of the research in creativity done by scholars had aimed to identify the cognitive attributes of creative people. As mentioned earlier, creativity was considered as a talent that only a few people had, and the studies focused on the ways of thinking likely to favour creativity. According to her, by contrast, ordinary people could be creative, and the main point was to identify the organizational factors that may influence creativity. She seeks to understand how to favour everyday creativity in organizations: which factors facilitate it and which factors constrains it. More specifically she highlights the importance of motivation and how organizations are able to raise this motivation. Her work has had a big impact on management research. She uses experimental methods that provided a lot of quantitative and qualitative data. She studies mainly scientists in R&D laboratories in organizations as well as employees in innovation units.

Teresa Amabile defines creativity as the ability to produce new and useful ideas, in any field. To be considered creative, an idea must be different from what has been done before, appropriate to the objectives pursued and have value (Amabile, 1988). The author focuses on the final creative product or result, contrary to previous research, which was mainly on the individual cognitive process that allowed creativity. She identifies creative results by questioning experts about their field, then she tries to associate levels and kinds of motivation to the individuals and the moments related to those creative results. In order to analyse what motivations drive the actors, she uses interviews and surveys. She also asks researchers to keep a daily journal in which they have to answer several questions related to how they feel while doing their task and what motivates them the most during the day.

The most relevant findings coming from her works consist in explaining the role of intrinsic and extrinsic motivation in creativity. **Intrinsic motivation** occurs when the task or process is interesting by itself for the individual and it is at the origins of creative results. It consists, more specifically, in the motivation caused by doing something that is considered interesting, satisfying and enjoyable, or by considering the job as challenging by itself. **Extrinsic motivation**, on the other hand, is considered as undermining for creativity and it consists in the motivation that comes from other factors than the enjoyment of the task itself. Some examples in an organizational context are performance rewards or career development linked to achievements, which often puts co-workers into competition. When put into competition or aiming to get some rewards, creativity is undercut as individuals seek for straightforward solutions that will meet the expectations of the managers in the fastest possible way. Extrinsic motivation has also a negative effect on intrinsic motivation by making individuals feel constrained and unable to advance guided by curiosity. These results are found to be valid in organizational contexts (mainly scientific laboratories of private organizations) but are also tested in experiments done with children and young adults (Amabile, 1988).

Intrinsic motivation appears to be very relevant in the early stages of the creation of an idea. In firms, as the idea turns into an innovation or a product there are other factors that, together with intrinsic motivation, participate in enhancing creativity. They consist in domain expertise and creative-thinking skills. **Domain expertise** is composed by knowledge and technical skills allowing individuals to identify and integrate relevant information. **Creative-thinking skills** consist in the cognitive abilities likely to be linked to creativity, such as perseverance, ability to take risk, critical thinking and ability to see problems from multiple perspectives.

When these three characteristics appear together (motivation, creative thinking skills and expertise) creative results are the most likely to appear. But what is the role of the organization in this context? The organization can influence motivation through **work environment**. We can find, for instance, all the extrinsic motivators that have been shown to undermine intrinsic motivation, as well as other factors in the work environment that may harm intrinsic motivation and creativity, such as low risk attitudes from the management, strict norms on how tasks should be done and/or lack of autonomy for the workers. Other factors may stimulate creativity, such as work teams with a diversity of skills, policy in favour of collaboration, freedom in carrying out the work, supervisors who encourage the development of new ideas, and norms of active sharing of ideas across the organization. Amabile recognizes, however, that extrinsic motivation is needed even if it is in conflict with intrinsic motivation. Firms have precise objectives and several restrictions that impede the establishment of complete freedom for workers to do their jobs based on the sole intrinsic motivation. For this reason, what she proposes is to promote a fluent communication that allows the managers and the employees to find common areas of understanding and common goals where extrinsic and intrinsic motivations meet (Amabile, 1997).

Finally, more recent works of Amabile introduce the idea of **inner work life effect**. She finds that individuals have a higher intrinsic motivation (and are, therefore, more creative), when they have positive emotions at their workplace and about their work, their organization and themselves. When this happens, they are also more productive, in a better mood and they are nicer to each other, which increases creativity of the others as well. If this happens there is a positive feedback loop. She studies as well what leads to those positive experiences and she finds out that it is mostly **making progress in relevant meaningful work**, which means work that the individual considers relevant (Amabile, 2011). Here again the role of organization's managers to favour positive emotions at the workplace is crucial. They must allow individuals to make progress in their work, which often means reducing the number of side-tasks, allowing them to be focused and giving them the necessary resources. It is important as well to provide them with enough **autonomy** in order to let them focus on finding solutions to problems. Finally, **communication** is again a key factor in order for everyone to understand what the other is expecting and to develop a trust relationship.

To sum up, the most relevant contribution of Teresa Amabile is to point out that all individuals have the potential to be creative and that this creativity depends highly on the organizational conditions that foster intrinsic motivation and hence creativity. She finds that intrinsic motivation, which is the one that comes from the pleasure of the task by itself, is crucial for creativity, contrary to extrinsic motivation that demines creativity. This intrinsic motivation, which can be enough by itself in the early phases of the development of an idea, needs to be accompanied later by creative-thinking skills and an expertise on the subject. A very useful impact of her work consists in explaining that the managers of an organization have a crucial role to play in order to assure this intrinsic motivation and to make it compatible with the objectives of the organization. The most relevant management leverages are to allow for certain autonomy and to ensure effective communications. The same factors can have a positive effect on what have been called the “inner work life”, which in the end affects positively intrinsic motivation.

Interestingly enough, the works of Teresa Amabile, although done in the field of psychology, have had a substantial influence on Management science as it highlights the importance of organizational conditions for creativity. By focusing on these conditions and on motivation issues, she introduces the notion of the process of creation rather than on a set of requirements (cf section 4.2 for developments by Management science). In the area of sociology of science also, several other authors have studied the organisational conditions beneficial to creativity. In the following section, we present the corresponding literature in the specific case of creativity in science.

3.1.4.2 Organizational conditions for scientific creativity

After Amabile, there has been a growing interest for continuing the study of the organizational conditions that favour creativity. For this reason there is a robust body of literature that tries to identify and explain the organizational conditions helping scientific creativity from the point of view of Sociology of sciences (i.e. Hage and Mote, 2010; Heinze et al., 2009; Hollingsworth, 2002; Zuckerman, 1967). The most relevant factors that have been identified can be divided into two different groups. On the one hand we have communication and diversity, on the other hand we have the attitude of management and the autonomy of the researcher.

Communication and diversity

Pelz and Andrews (1966) find that scientists are most creative when they do not work alone. Indeed, scientists who interact actively with each other, and involve the whole team when setting up their research goals, are often more productive. Research productivity is correlated with high frequency of **intra-organizational communication** (Heinze et al., 2009). When scientists in the same organisation communicate, they nourish each other’s ideas and viewpoints and they are more likely to come up with an original idea. Communication needs to be effective, which means that the individuals need to be open to share their ideas and open to integrate other's ideas. This happens for instance when working together to solve a problem, communication is therefore materialized in the form of teamwork. When communication happens across teams it is even more likely to generate new ideas because there will be a greater gap among frames of thought, concepts and viewpoints. Mobility of researchers and teamwork between departments has, in this line, been recognized influencing positively creativity (Hage, 2006).

In addition to an effective intra-organizational communication, it is also important to allow **communication across the boundaries of the organization**. When communicating with scientists working in other institutions, not only there is a greater gap in terms of concepts and viewpoints, but there are also additional gaps in terms of know-how, ways of working, routines or methods. If it is done efficiently, this kind of communication can be extremely beneficial for creativity (Heinze 2009). The ways in which this cross-organization communication is the most efficient are often collaboration, participation in conferences and workshops and mobility. When developing joint projects with other institutions, communication between scientists of both institutions is facilitated by the existence of a common objective and the development of trust through repeated interactions. When participating in conferences and workshops, scientists are challenged to defend their ideas by peers, and this initiates discussions with a strong exchange of ideas. Finally when moving from one institution to another, scientists will bring with themselves know-how and ways of working that they will exchange with the colleagues from the new institution.

What we observe up to now is that what makes communication crucial is the exchange of experiences, concepts and viewpoints. We have seen as well that the greater the gap between the individuals, the greater the chances of being creative. Hollingsworth (2000, 2004) introduces the idea that this communication is more beneficial when it operates not only across organizational boundaries, but also across **discipline boundaries**. He argues that research breakthroughs are typical for research organizations where scientists communicate both inside and across the boundaries of the discipline and of the research topic. The more the amount of disciplines that participate in a research project, or share a laboratory or institution, the greater the chances for the scientists of being creative. This result is highly connected to the notion of diversity. It has been found that scientific

diversity is also very important for creativity. For example, because the Rockefeller University was organized around laboratories rather than scientific disciplines and fields, it had scientists with a diversity of backgrounds sharing a similar exchange. This diversity, on the condition that it was interconnected by means of an effective internal communication policy, allowed for a greater set of knowledge and ideas to be available in order to solve a problem (Hollingsworth, 2004). Similarly, Heinze (2009) shows the importance of the access to a relatively large variety of technical skills. However, it has to be noted that the scientific diversity of an organizational environment alone may not foster creativity unless it is also linked to organizational arrangements that support multidisciplinary contact. Diversity and effective communication must therefore, come together.

The role of managers and the need for autonomy of researchers

Since organizational conditions matter, it is important for managers to endorse a policy that ensures the presence of some kind of variety in the research process, as well as the communications that will allow the organization to benefit from that variety. But the literature reveals that it is not enough that management promotes variety and communication. Allowing scientists to be autonomous and having managers involved in current research activities is also beneficial for creativity.

Management studies show how individuals are more likely to be creative when they are allowed to follow the objectives that they consider better and to use the methods that they consider the most appropriated. This is typically the case in science. When researchers have **autonomy** to make their own decisions regarding research agendas and about how to accomplish these agendas, they are more likely to come up with creative results (Pelz and Andrews, 1966). Freedom to define and pursue individual scientific interests within or beyond a broadly defined thematic area is central to understanding why scientists and their groups are highly creative (Heinze et al., 2009). Additionally, Pelz and Andrews (1976) find that effective leaders are **involved in ongoing research**. Active participation in the praxis of scientific work is important for leaders to understand the problems of the group, to motivate group members and to organize a coherent research program. This finding is also reflected in the literature review by Mumford et al. (2002), who suggest that leadership in creative environments requires predominantly technical and scientific expertise.

The reasons that have been found to explain this relationship between autonomy, investment of managers and creativity are the following. Creative thinking is defined by being new and somehow unexpected. Therefore, when establishing strict rules on how an activity should be performed, novel ideas cannot easily emerge. The same blocking appears concerning the objectives: if the goals are dictated by leaders, there is very little space for new ideas and especially for the pursuit of intuition. And the latter has been found to be strongly correlated to creativity (Simonton, 2004). Finally, as highlighted before, psychologists have found that creativity is highly dependent on motivation, therefore goals need to be decided by the people performing the activity (Amabile, 1996).

3.2 Creativity as a collective process: contributions of management literature

Up to now we have seen how most of the research on scientific creativity has been done by considering creativity as something that happens at an individual level. The works of Amabile, as well as the ones done after her in order to identify the organizational conditions that favour creativity, recognize the crucial role of the group. However, they consider creativity as lying on the individual, rather than the group. This comes from considering the individual scientists as the unit of production. We consider the focus on individual creativity as a limitation of current works on scientific creativity, especially when Big Science and/or large research infrastructures are at stake. This is more generally a limitation since collective aspect is becoming crucial as science is increasingly done in collaboration, as is shown by several studies on the growth of co-authorship (E. Y. Li et al., 2013; Wagner and Leydesdorff, 2005). Indeed, science is the more and more performed by teams. Scientific papers signed by a single author are becoming rare in almost every discipline and the individual cannot longer be considered as the creative unit. It is because of this reason that we study, in this section, the management view on creativity. In management, although the focus is usually set on innovation and not on science, the idea of a collective creation is always present. Three important aspects of creativity, in the literature in management science, can be emphasised in this respect: the notion of a dynamic process, the role of learning and the role of communities.

3.2.1 The creative process: perspectives from the study of innovation

Although an important part of the study of scientific creativity focuses on the individual, several studies in the field of Sociology of science consider the production of knowledge as a collective activity. Callon (1988) for instance, considers the production of knowledge as a result of a collective activity that lies on a network of interacting humans and non-humans "actors[24]" (such as instruments or the evolution of the discipline itself). (Latour, 1999) follows the same idea when defining the scientific inquiry as a collective activity that puts together a multiplicity of objects and actors who cooperate, negotiate and organise everyday work. Science production is indeed a complex activity where the elements that enter this process are humans as well as instruments or data or whatever you want. Simply say that you are using it as a general approach, as a background, for example that you will take into account the interconnections between very different and very different elements, not only between human actors (for example)

Literature on management focus on how the firm can adapt to an evolving environment and keep up with technological advance. It also focuses on the innovation process at the level of the firm. Innovation process is traditionally viewed in the following way. It starts before the emergence of an idea, with the gestation of this idea. Then there is the emergence of the idea. It continues with the elaboration of innovation by building on top of this idea and it finishes by the acceptance and spread of that innovation (for instance by transforming it into a product). This process can be interrupted at any moment and being re-taken at some of the previous steps (Garud et al., 2013). This process is, however, more complex than this. For this reason, several scholars have proposed models of the productive process of innovation that allow for certain accuracy while keeping.

Let us look, for example, at the Chain-linked model of innovation from Kline & Rosenberg 1986. It is a first attempt to describe complexities in the innovation process and it represents a non-linear process for innovation within one entity (such as a team or an organization). There is a central path which has one main direction of events but is constantly giving feedback to the precedent ones. In the chain-linked model, new knowledge is not at the origins of innovation. The innovation process begins, instead,

---

[24] In the sense of Actor Network Theory initiated by Callon and Latour, the term "actor" refers to an "acting agent" and includes objects in addition to human beings.

with the identification of a potential market (such as an unfilled need). This drives analytic design, then detailed design and test, then redesign and production, and finally distribution and marketing. This process has as well complex feedback loops between all the stages. The process is fed with the use of knowledge from outside the organization, external knowledge, as well as by research. There are also important feedback loops with the organization's and the world's stored base of knowledge.

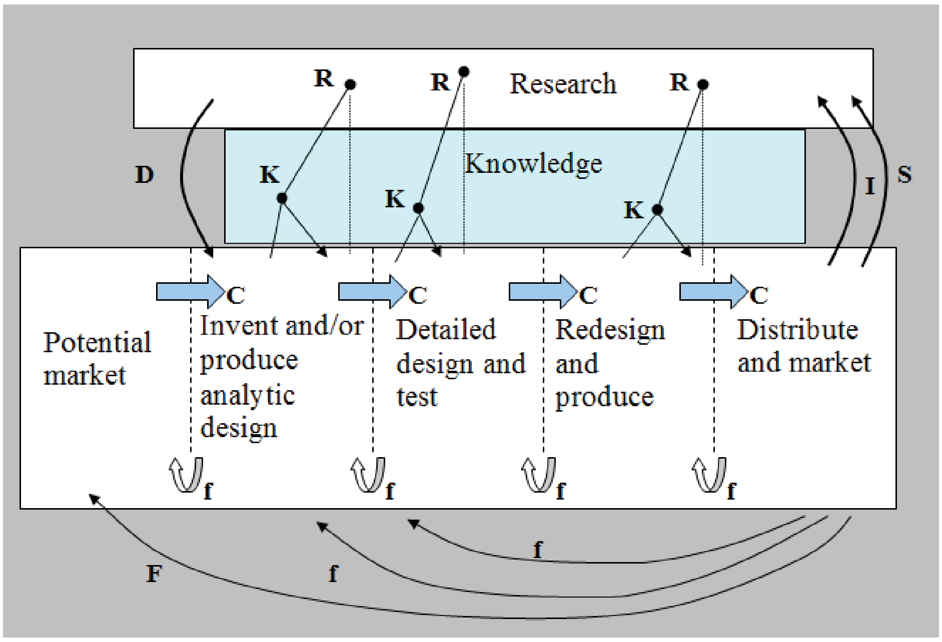


Figure 3: (Source: KLINE S.J., ROSENBERG N. (1986), "An overview of innovation", In Landau R, Rosenberg N. (Eds), "The Positive Sum Strategy", Academy of Engineering Press, p 275.)

We think that, although science works in a different manner, there are several interesting aspects of this process that can be applied to science. Indeed, we can think about a central path that includes different stages and which is non-linear but has a general trend or direction of actions. In science we look for a research gap, we think about how to fill it in, we design the research with the collection of data, methods and tools that will be used. Then we do this research and we end up with a final outcome which can be a result in the form of a published paper. There is also external knowledge and feedbacks between the different stages.

3.2.2 The role of organizational learning

Management research has often focused on innovation, rather than on creativity. Creativity has been recognized, however, as the pre-condition for innovation. Innovation requires being able to combine properly exploitation and exploration activities. Exploitation consists mainly in selling on the market the products developed by the organization. Exploration is the ability to create new products or develop new methods and technology, hence opening new options (new markets) and being able to cope with the changing environment (March, 1991; He and Wong, 2004). Exploration demands a capacity to be creative and therefore management research got to the point of studying creativity.

One of the factors that has been described as important for being able to innovate is the use of external knowledge. Cohen and Levinthal (1990) develop the concept of absorptive capacity, which consists in the firm ability to recognize the value of new information, assimilate it, and apply it to commercial ends. This idea has been largely used to explain how, without absorptive capacity, firms cannot profit from the knowledge that is produced at universities and the technological developments produced by other firms. Indeed, the knowledge that is produced by other firms and universities (patents or scientific articles for instance) is not easy to be identified and understood. This is because it is complex, specific and also because there is often a cognitive distance between the firm's knowledge and the external knowledge that can be exploited (Nooteboom et al., 2007). Traditionally, literature proposes that, in order to maintain its absorptive capacity, the firm should continuously develop knowledge creating activities (R&D or innovation) to remain competitive (Minbaeva et al., 2003; Todorova and Durisin, 2007).

Not only the use of external knowledge is important, learning, within an organization, is crucial as well. In a context where management research concentrated on the study of innovation, James March analysed the relevance of creativity in order to allow firms to produce innovations. He considered creativity as being strongly dependent on learning. He worked on the concept of organizational learning, which he described as routine-based, history-dependent, and target-oriented. Organizations learn by encoding inferences from history into routines that guide behaviour. Organizations can learn from their own experience, but they can also learn from the experience of others. When learning from other's experiences, new conceptual frameworks or paradigms for interpreting those experiences are developed, together with the organization's own experience (March and Levitt, 1988). These new frameworks allow the organization to be creative. March shows in several of his works that creativity is crucial for the firm to survive in a changing environment, and more specifically that firms need a continuous flow of new ideas to be sought beyond the limits of the organization itself (March, 1995). He maintains that a part of the organization's resources, particularly the surplus from good periods (i.e. organizational slack), should be dedicated to the free production of ideas by individuals and groups without any formal control by the organization, without monitoring or defining goals.

The importance of collective learning is central in management literature when studying creativity. The most relevant model for learning within organizations is probably the one proposed by Takeuchi and Nonaka (1995). They developed the well-known S.E.C.I. model of organizational learning, which describes the different types of learning created in organizations based on a principle of knowledge transmission and conversion. The possible conversions are: from tacit to tacit knowledge (Socialization), from tacit to explicit (Externalization), from explicit to explicit (Combination) and from explicit to tacit (Internalization). By means of socialization, tacit knowledge, that is to say, the kind of knowledge that cannot be written down and formalized and that can be transmitted only by practice, is shared. Here there is a need for face to face interactions and shared experiences. Socialization typically occurs in a traditional apprenticeship, where apprentices learn the tacit knowledge needed in their job through hands-on experience, rather than from written manuals or textbooks. Externalization consists in turning the tacit knowledge into explicit knowledge by writing and articulating the knowledge. For example, concepts, metaphors, images, and in the end written documents can support this kind of conversion. When tacit knowledge is made explicit it can be shared, transformed and used by others who do not necessarily shared the practice, and it becomes the basis of new knowledge. Combining explicit knowledge is another form of learning as it consists in integrating different bodies of existing knowledge. Explicit knowledge is collected from inside or outside the organisation and then combined, edited or processed to form new knowledge. The new explicit knowledge is then disseminated among the members of the organization. Finally, there is the learning from explicit to tacit knowledge: explicit knowledge becomes part of an individual's knowledge and it will constitute the assets of the organization. Internalization is also a process of continuous individual and collective reflection, leading to the ability to see connections and recognize patterns, hence the capacity to make sense between fields, ideas, and concepts.

All of these kinds of learning are crucial for the integration and development of new knowledge and ideas. Learning has been shown to happen often at a collective level and more specifically to happen at groups where there is trust and a common identity, which are communities. Indeed, literature on communities show how important they are for knowledge leaning within and across organizations, as they play a crucial role for knowledge transfer. It is for this reason that there is an extensive body of management literature dedicated to the study of communities.

3.2.3 The role of communities

Communities are groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise on this area by interacting on an ongoing basis and they play a crucial role in learning, as learning is made easier with cohesive social ties. The role of communities for organizational creativity was first analysed in studies of work practice in firms. They showed that creativity increased within communities having cohesive social ties (Lave and Wenger, 1991, Orr, 1996). These studies also showed that communities engaged in a common practice provide the context for collective learning and problem solving as well as knowledge creation and integration in the firm (Brown and Duguid, 1991; Grant, 1996, Hargadon and Bechky, 2006, Kogut and Zander, 1996). Communities provide their members with shared identities and a common language, which is instrumental for efficient communication and thus coordination (Kogut and Zander, 1996). It is important to note that communities are not always delimited by the boundaries of an organisation (O'Mahony and Lakhani, 2011). Therefore, management needs to consider the existence of members who belong to the community but are external to the organisation, and to ensure that communication with them remains fluent and easy. Not only the communities facilitate the **creation** of new knowledge, they also facilitate the **adoption** of this new knowledge and they constitute a proper structure for learning (Cohendet et al., 2013; Harvey et al., 2015).

A second relevant perspective on communities and creativity is the communication **between** different communities. Communities are described as key when participating in interdisciplinary networks that gather together knowledge from different actors and from different disciplines. Organisations need to provide efficient communication within a community but also with **multiple** external communities. When the boundaries of the firm are blurred, organizational creativity becomes the capability to leverage community efforts and integrate knowledge across communities whose members are mostly outside the organization (Harvey et al., 2015; Powell et al., 1996). If these different communities collaborate, the output will be more creative and the more they collaborate the more creative the output will be (Batallas and Yassine, 2006).

3.3 Towards a conceptual framework

Here we try to position the thesis in the literature reviewed above, before establishing its conducting line and the main conceptual building blocks.

3.3.1 Synthesis of the most relevant points of the literature

Psychologists have been studying creativity for several decades now. The focus was often put on the analysis of creative individuals, as in the case of the research done by Koestler (1964) and Simonton (2004) for example. Later on, Amabile (1997) argued, from the perspective of psychology as well, that every individual could be creative. Thus rather than studying the cognitive traits of creative people it was important to examine the environment that led people to have the motivation to be creative. In all these relevant contributions to the study of creativity, scientific activity has been one (if not the) focus of study. Creativity in science has been defined as the production of knowledge that is new and valuable.

These two criteria, which are supposed to be present together in order to obtain creativity, often push science in opposite directions. When validating scientific knowledge scientists must prove the originality and plausibility of their work. But these two criteria of validation are in continuous tension. Highly original work has often very little to build up on and it is therefore difficult to recognize its value. This tension has been recognized and explain by multiple works (i.e. Garfield, 1977; Heinze et al., 2009; Hollingsworth, 2002; Merton, 1975; Simonton, 2004; Stent, 1973).

Despite this tension some authors have conceptualized and studied scientific creativity. They have studied the organizational conditions under which scientific creativity occurs. Amabile (1997) first talked about intrinsic motivation as the key concept for understanding creativity. Intrinsic motivation, however, is an independent variable that individuals come equipped with. Intrinsic motivation is strongly affected by the organizational conditions in which the work is done. Organizational conditions and structures that have been found to have a positive relationship with creativity are the following: diversity of team background, communication within the discipline and across disciplines, autonomy of researchers and management involvement on scientific activity (Heinze et al., 2009; Hollingsworth, 2002; Zuckerman, 1967).

Although there is a general agreement on this definition of scientific creativity, it is not enough to get an appropriate vision of the creative processes in science today, that is to say, a vision that takes into account the collective dimension of scientific production when it is based on RIs or highly sophisticated instrumentation platforms, large databases or e-infrastructures, together with an increasing use of R&D collaborations. In order to complement the literature dedicated to scientific creativity in this respect (collective dimension), we analysed some important contributions of management science, especially those related to innovation management

3.3.2 Positioning the thesis along the dimensions of creativity

The literature review at the outset of Section 3 brings to the foreground several dimensions of creativity, that help us to position the works of scholars of the field. More precisely, we have identified four main dimensions of creativity: the novelty versus value components, the big C versus small c distinction, the focus on creative output versus creative process, and finally, the individual versus collective scope to which creativity may apply. The following lines are dedicated to discuss these dimensions further and to position our work with respect to each of them.

3.3.2.1 Novelty and Value

The most important dimension of creativity relates to the weight of the two attributes that define the term of creativity by itself, namely novelty and value. Novelty and value, although they might seem like two different concepts are two forces that would pull into opposite directions and that are in continuous tension. Although recognized, this tension makes it difficult to observe and theorize creativity while giving the same weight and relevance to both attributes. This difficulty is made visible when looking at how creativity is empirically studied: works about scientific creativity have often overlooked one of them and put emphasis on the other. Some authors would exclusively (although in an implicit manner) focus on value (i.e. Amabile, 1983; March, 1991) and others would exclusively talk about the novelty of the creation (Carayol et al., 2018; Lee et al., 2015; Wang et al., 2017).

This thesis will try to avoid this trap by continuously looking at both aspects. All along the thesis chapters, we will consider that the co-existence of novelty and value are the two pillars that constitute the essence of creativity. They will be systematically analysed since they represent the necessary elements for creativity to be qualified as such.

3.3.2.2 Big and small creativity

The literature distinguishes between big and small creativity. Big creativity would be when the scientific production has a big degree of novelty and newness. It is typically the case of breakthrough discoveries that receive the Nobel prize. Small creativity, on the other hand, would include every-day problem solving and smaller scientific advances. Simonton (2004) explains this difference and calls it big C and small c. This distinction is rarely used in articles about scientific creativity but when studying creativity in science some authors implicitly focus on the big C, since they look at major scientific discoveries (i.e. Hage and Mote, 2010; Heinze et al., 2009; Hollingsworth, 2002); while others consider also the small c since they look at all the scientific articles produced by any member of a scientific community (i.e. Lee et al., 2015; Wagner et al., 2016; Wang et al., 2017), thus studying all the scientific advancement, big or small.

As the attributes that compose creativity (value and novelty) are not necessary binary, we could talk about different degrees of value and different degrees of originality/newness. We could even consider a continuum. Here in this thesis we will consider creativity as an attribute that scientific work can have, and we will observe the whole spectrum, from small c to big C and including all the options that there are in between. The reason to do so is that we think we cannot content ourselves with a dichotomy between the creativity present on everyday tasks and the creativity that has as a result a breakthrough discovery. There are several middle points such as relevant results in a specific field that, without being disruptive, help science advance significantly. Chapters 2 and 3 will look at creativity in general, regardless whether it is big or small creativity and Chapter 4 will develop a quantitative measure for this creativity.

3.3.2.3 Creative process and creative output

What we have seen in Section 3.1 is that the study of scientific creativity often focuses on the creative results. Arthur Koestler, for instance, explains the concept of bisociation by studying how some breakthrough discoveries by famous scientists emerged. He studies for instance the cases of Albert Einstein and Charles Darwin. Early studies would identify creative individuals by studying successful scientists and then try to understand their ways of thinking and approaching problems. It is the case of Teresa Amabile as well, when she relies on the idea of value to identify creative results as those that have been recognized as such by the peers. Once the creative results are identified she aims to find to which states of mind that creation was related, and later on which organizational conditions have favoured it. More recent works identify creative science by identifying laboratories with a large number of Nobel laureates or nominees to other prizes and then study the organizational conditions that favoured the occurrence of that creativity. Although the organizational conditions play a role in the creative process, the latter is not studied. This thesis aims to study the whole creative process, including its outputs and its organizational determinant, but also its mechanisms, in order to analyse whether and in which ways research infrastructures may impact it.

3.3.2.4 Individual and collective creativity

We can consider, particularly when observing a creative product, creativity as lying on the individual or as lying on the group. Traditionally most of the research in scientific creativity has focused on individual creativity. For instance Guilford (1950) identifies intellectual abilities that favour creativity such as flexible and analytical thought or divergent thinking. Koestler (1964) discusses the individual ability to put together different frames of through in order to create a completely new one. Amabile (1988 and 1983) is one of the first to consider creativity not as a trait of exceptional people but as an attribute to the final product. She, however, keeps focusing on the individual as she considers that every individual can produce a creative output if he has the motivation to do so. She, like some more recent works (i.e. Hage and Mote, 2010; Heinze et al., 2009; Hollingsworth, 2002) will study the organizational conditions that lead individuals to produce creative results. When doing so the collective aspect appears to describe which kind of context allows for individual creativity to emerge (communication with colleagues, existence of cohesive ties, etc.). It is the research in Management sciences (i.e. Cohen and Levinthal, 1990; Nonaka, 2000; Nooteboom et al., 2007; Takeuchi and Nonaka, 1995) that has considered the creative product as the result of a collective action and has considered the collective (i.e. the organization) as the agent that produces a creative output or that is able to endorse into creative process.

3.3.3 Building blocks of the collective science creation process

In order to understand creativity in science we have identified the most relevant elements that will build up our theoretical framework. We consider creativity as made up of novelty and value together. Because of this we, all along this thesis work, consider this both dimensions and do not consider that there is creativity when one of them is missing. Another important building block of this thesis is to consider creativity as an attribute of science that has continuity. Building up on Simonton (2004) we think that there is small, everyday creativity as well as the big creativity that brings breakthrough discoveries. Maybe there are interactions between small c and big C in the creative process, the former depending on the latter, or vice-versa. Maybe there are intermediate levels of creativity that shed light on some creative mechanisms. Hence, we prefer to consider the whole spectrum between small c and big C.

Still more importantly, we consider creativity in science as a **process** where knowledge is continuously combined. As in innovation studies, we conceive this process at a collective and not at an individual level. In the figure below, we present a simple conceptual model of a collective process of science creation, which we have designed to support our analysis about the creativity made possible by RI use.

When studying the innovation process, we can identify several of the elements studied for the understanding of scientific creativity: the process is moved by selecting and combining a variety of concepts, ideas and viewpoints. If we adapt this traditional process to the case of science, we could assimilate the gestation of an idea to the combinatory process before the emergence of the research question. The emergence of the idea corresponds to fixing a research question as in science the research question corresponds to the idea, that will be developed and tested later. The elaboration would consist on this development and testing of the idea which would include the observation and collection of data and its analysis. Finally, we would have the creative output. This figure represents this process.

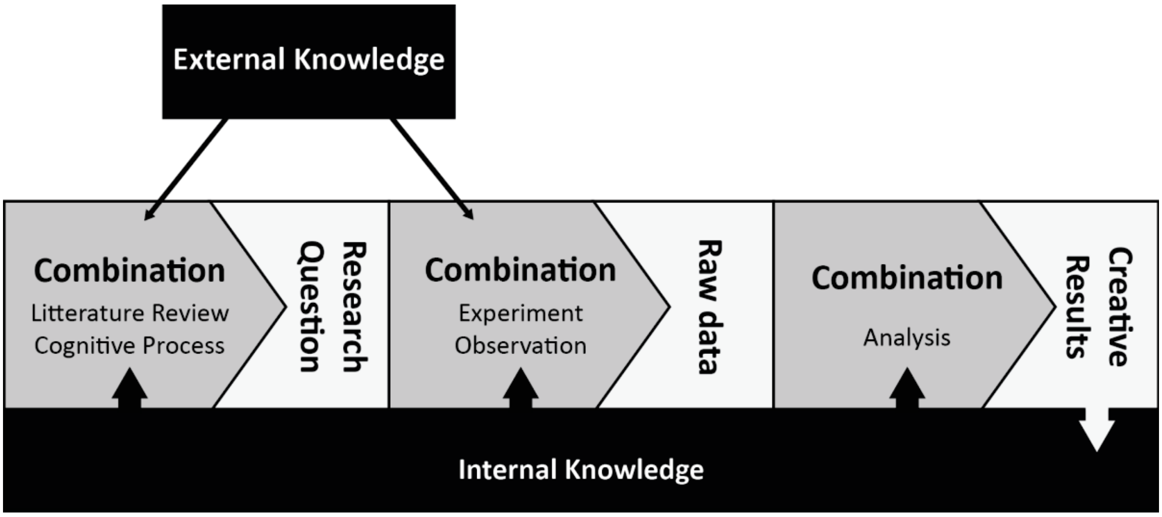


Figure 4: Collective knowledge-creation process

The focus on a process means that we study creativity in a dynamic way and not by looking at a creative object or result. This output is characterized by a continuous recombination activity. During the process of science creation different elements (formal knowledge, tacit knowledge, methods, data, ideas, concepts and viewpoints) are combined. This is expressed in the figure above. At the beginning of the combinatory process researchers are working on building up an idea, a step leading to what we call research question in science. To do so they use their own knowledge and external knowledge, such as literature, and they combine them. When the research question is set the combinatory process consist in combining techniques (experimentation, observation) in order to produce data. The data set will be analysed by using internal knowledge in the form of formal knowledge, techniques and methods. This process will result into a creative output in the form of creative scientific results. In this sequence of combinatory processes, knowledge (both internal and external) may be present all along the central path.

Literature has shown us which conditions make this combinatory process creative. We know that the knowledge used in order to be combined needs to be varied. The pieces of knowledge that are combined need to be different from each other. This variety appears often in the form of multidisciplinarity, which provides the creation process with formal and informal knowledge, methods and techniques that come from different disciplines. Communities allow for this multidisciplinarity by facilitating communication, especially of the informal type, within the community and across it.

This thesis, through its three core chapters (Chapters 2, 3 and 4), will study the role of RIs in this process, as well as develop a technique that aims to assess the creativity of research done by using RIs. We will look at two cases, a European Synchrotron and a Biological database, and use qualitative case study methodology which is pertinent for the complexity of the studied topic.

4 Epistemological and methodological foundations

The objective of this section is to detail our research strategy; more precisely we want to present the empirical and methodological mechanisms assembled for this graduate work which aims at studying how RI can favour creativity. We consider that the quality of a research work lies on both the logic of the whole research process and the coherence of all the elements that constitute it (Zarlowski, 2007). Previous section presents the research question of the thesis by constructing the theoretical framework and identifying the research gap. In this section we will start by presenting and justifying the choice of epistemological paradigm as well as the reasoning model. This will bring us to understand the link between the research question and the methodology. Finally we detail the empirical methods, tools and approaches that we use. We justify the choice by showing how it respects the coherence mentioned above as this coherence between the research questions and how the question is answered is crucial for the validity of knowledge. More details concerning the methodology (such as choice of cases) will be developed in each chapter.

4.1 Epistemological foundations

This section is devoted to explain and justify the scientific approach of this thesis. Up to now the chapter has been dealing with the topic and aim of this PhD thesis, that is to say "what do we want to know?". This section will focus on its epistemological position, that is to say "how do we want to develop knowledge?" and "why this way?". To answer these questions, we should firstly specify that this thesis lies in the field of "science policy and innovation studies" which has also been labelled as "innovation studies" or "science and technology studies". It is a multidisciplinary field which is nourished mainly by four sizeable scientific disciplines: policy, sociology, management and economics (Martin, 2012). From the multiple definitions of the field, as we can see in Chapter 1, we use in our case the one used by the journal that is considered as the more relevant in the field, namely, Research Policy. The field is devoted to *"analysing, understanding and effectively responding to the economic, policy, management, organizational, environmental and other challenges posed by innovation, technology, R&D and science. This includes a number of related activities concerned with the creation of knowledge (through research), the diffusion and acquisition of knowledge (e.g. through organizational learning), and its exploitation in the form of new or improved products, processes or services"*[25]. The link to this field will condition the epistemological choices.

Epistemology is the scientific field, a sub discipline of philosophy that studies the theories of knowledge generation. The term describes as well how the researcher approaches the construction of new knowledge (Gavard-Perret et al., 2012) and the choice of a paradigm. This choice of a paradigm defines how the research is designed. The concept of paradigm, introduced by Thomas Kuhn (1970), has been widely used, sometimes in a vague way, to describe the idea of a certain grid of analysis possibly co-ordinated with a set of practices to which the members of a given social group adhere (Soler, 2000). In this section, which deals with the few dominant epistemological paradigms, we will sketch the main lines that differentiate them in order to be able to choose a satisfactorily epistemological framework.

4.2 Epistemological paradigms

Different epistemological paradigms will offer different pathways on how to pursue research and assure value and validity of the results. It is the community, namely the peers, who will determine that validity and value according to a certain number of rules of the field. It is, therefore, crucial, to choose the epistemological approach that is more coherent with the studied research topic and the field where research is performed. This

---

[25] https://www.journals.elsevier.com/research-policy/

sub-section is devoted to that choice. Before describing the different epistemological paradigms, we will look at the two major streams of thought in the history of Science. The first, rationalism, considers the analytic reasoning as the fundament of all scientific knowledge. The construction of valid knowledge is done only by deduction of the generic abstract to the particular concrete. The second, empiricism, considers that sensitive experience and experimentation are the only source of valuable knowledge. The hypothesis check and the generation of valid knowledge can only be done by induction, going from several particular cases to conclude a general law.

From these ideas two big epistemological paradigms emerged in science: positivism and constructivism. According to Besnier (2016) positivism strives to establish the laws linking phenomena without being interested in the search for their causes. Subsequently, the term "positivism" qualified the attitude focused exclusively on the study and description of facts. This paradigm excludes from science anything that cannot be reduced to sensory experiences or to submit to statements based on a logical analysis. In constructivism, on the other hand, the knowledge about phenomena comes from the result of a construction made by the individual. Several other epistemological paradigms are considered today but literature does not provide with a unified view on the cartography of epistemological paradigms. For instance Gavard-Perret et al (2012) distinguish between five epistemological paradigms (scientific realist, critical realist, pragmatic constructivist, interpretivist and constructivist). Van de Ven (2007) talks about philosophies of science and distinguishes between four of them (positivism, relativism, pragmatism and realism. Giordano (2003) describes only three (positivism, interpretivist and constructivism). We adopt here the typology of Gavard-Perret (2012) and choose critical realism as our epistemological positioning.

4.2.1 Critical realism

This graduate work aims the role of research infrastructure in science by looking at the collective process of knowledge creation. This process is complex, with multiple actors interacting with each other and multiple factors at stake. We consider that this process have exists and that it has many observable patterns. The ontological hypotheses of critical realism consider, in line with the positivist approach, that there is a reality independent of the researcher. In our case, we consider the collective knowledge-creation process an existing reality. The fact that the use of RI might change this process is a reality as well, regardless of the researcher that studies it. We have several ways to observe all the aspects of these interactions. However, because this reality is complex, and social, our ability to observe it is not unlimited. Additionally this topic that we study is rather original and therefore there is little previous knowledge to build up on. Therefore we are aware that our knowledge of the reality is imperfect. Because of this we believe that critical realism is the appropriated approach for our research.

Critical realism presents itself as the post-positivist alternative to both, positivism and constructivism (Archer et al., 1998), interpretivism being considered as a moderate form of constructivism (Gavard-Perret and al. 2012). From an ontological point of view, critical realism postulates, as do positivists, that laws exist independently of the facts of man or his ability to perceive them. Nevertheless, like Scientific realism, Critical realism recognizes that reality is not easily reducible to our perceptions and experiences of it. In other words, the nature of reality is not apprehended, characterized and measured without difficulty (research methods are fallible, and the cognitive abilities of the researcher limited). In addition, a distinctive aspect of critical realism is the idea of stratifying reality into three nested domains: (1) the real, where the generating mechanisms occur i.e. the rules that govern the occurrence of events, (2) the actual, which includes events occurring when generating mechanisms are implemented, regardless of whether they are or not observed by man, and (3) the empirical, which represents the human perception of the actual. Figure 5 represents this.

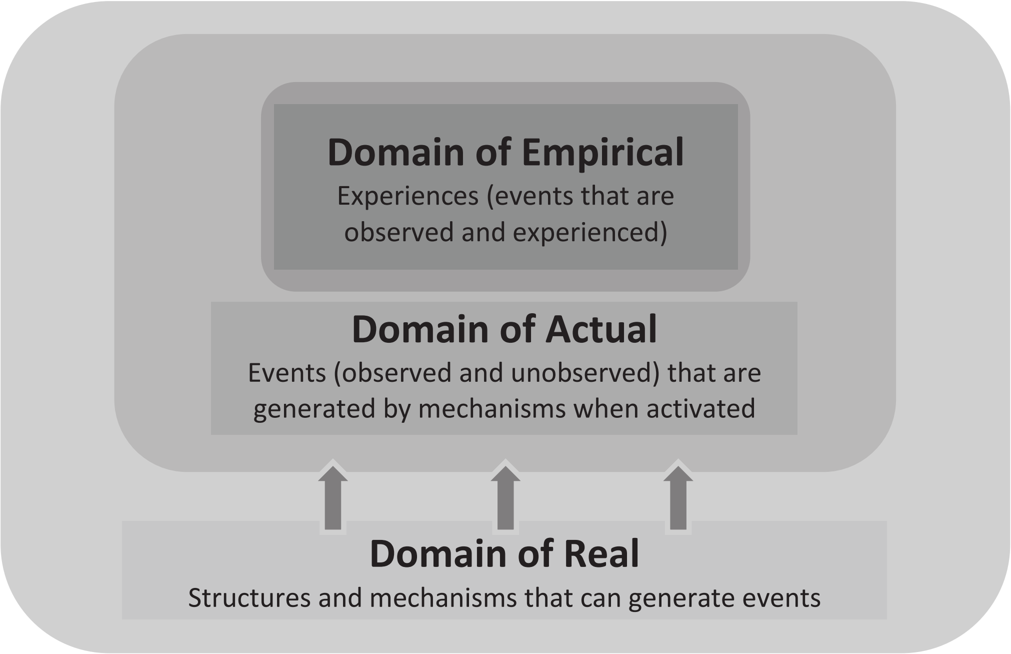


Figure 5 Critical realism, extracted from Bhaskar 1979

As far as the epistemic dimension is concerned, this paradigm posits that what is knowable is the empirical, which reflects the human perception of the real updated resulting from the activation of the generating mechanisms residing in the real deep. The goal of the cognitive process in critical realism is to identify these generating mechanisms and their mode of activation reflecting how they manifest themselves in the actual. Critical realism recognizes that our knowledge of this deep reality is subject to all sorts of influences including historical influences. It thus postulates an epistemic relativism admitting that deep knowledge of the real is always local and historical (Mingers et al 2013). However, this epistemic relativity should not make to lean the researcher towards an ethical relativity. Critical realism recognizes the need for objectivity and knowledge of the real as an ideal to achieve but it admits that it will never be achieved. This is what gives it its critical character because it questions the inference used to derive an ontological knowledge of the world. Also, critical realism adopts a vision of reality as an open system supposing multiple explanations across several generating mechanisms (Bhaskar 1998). To identify the generating mechanisms representing this reality and explain their modes of activation, it is necessary to consider different internal and external conditions (Collier 1994), including contextual social, organizational factors, environmental and technological issues, which may play a causal role in the occurrence of the observed phenomenon.

Figure 6 represents how research is done within this paradigm. The observer (the scientists) observes the empirical and by contrasting this observation to its own knowledge (theories and reasoning) tries to reach a rich description of the events that she is observing. The objective is to make the empirical as close at is possible to the actual by the use of reasoning and by contrasting the empirical to the previously acquired knowledge. Once this actual is induced, the researcher aims to be able to explain the underlying mechanisms that can generate events, or in other words, the researcher aims to explain the Real. Despite this will to give explanations that are as close as possible to the real, the researcher knows that these explanations can be only imperfect.

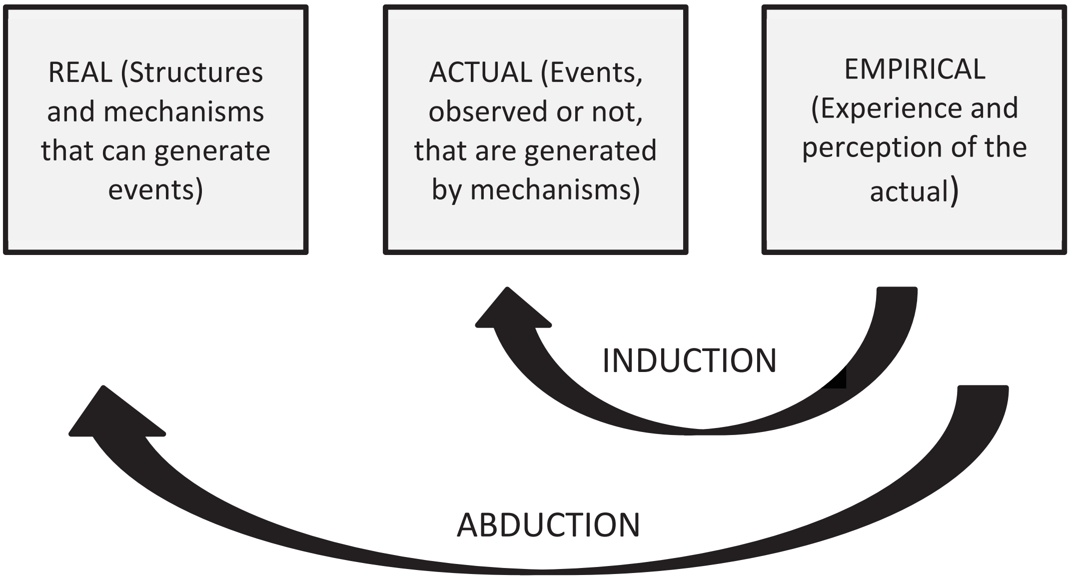


Figure 6 Research within a critical realism paradigm

4.3 The reasoning model

Inductive and deductive reasoning are the traditional reasoning approaches used to construct a valid argument. Inductive reasoning moves from specific instances into a generalized conclusion. It often used data and observation of the world in order to construct theory about that world. Deductive reasoning moves from theory that is built by abstraction and reasoning towards an hypothesis and later on it uses observations of the natural world to confirm those hypothesis (Collis and Hussey, 2003). Finally, a third option is the abductive reasoning. This is the reasoning chosen for this thesis as it is the reasoning that is most often suggested within a critical realism paradigm.

Abductive reasoning is often used in management sciences when the goal of the research is not to develop a general rule but rather to propose new conceptualization in form of a new model, a new theory or new hypotheses. It is the chosen reasoning model when the observation of the studied reality (in form of data or other) is done in parallel to the reasoning and building of theory (Gioia et al., 2013). Within an abductive reasoning there is, at the beginning, the construction of a theoretical framework which will be used for the study of the object of interest. By observing the object, this theoretical framework will be adapted to include all the mechanisms that had not been expected or taken into account ahead because the existing knowledge didn't allow for imagining them. Within an abductive logic, therefore, the theoretical framework evolves as the research advances. In this sense abduction compares ongoing empirical evidence to support the development of conjectures. In this thesis we follow that logic and we confront the field while developing a new conceptual framework.

In this graduate work we proceed as follows:

- We start by looking at the literature and simultaneously exploring the fieldwork. Being familiar with the fieldwork and in parallel with the literature in economics of science, management of innovation and sociology of science, both activities can be done easily in parallel. The exploration of the fieldwork consists on looking at documents such as policy documents, journalistic articles and annual reports of the studied RIs.
- We start early to build the theoretical framework. While building this theoretical framework there are, often, returns to both, the fieldwork and the literature. At this stage of the study, the contact with the fieldwork is still exploratory. Although we are already familiar with it, we do not know its adequation to the current theoretical questionings. In parallel to the theoretical framework a research question emerges.
- Finally, we confront our theoretical framework to the fieldwork. Our objective is to verify how adequate this framework is but also to complete it and concretize it. It is for this reason that the fieldwork is continuously feeding back the earlier steps of the process.

The following figure shows this process:

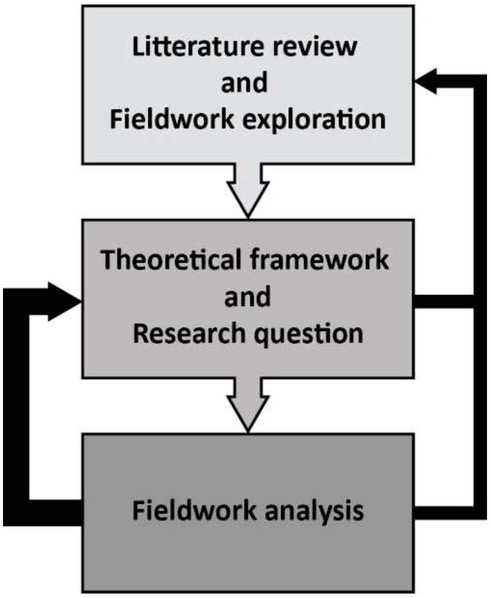


Figure 7: Abductive reasoning model

4.4 Research methodology

In this thesis, the use of different bodies of literature as well as various methodologies comes mainly from a willingness to approach the reality as closely as possible. This thesis implements a mixed-method approach, mobilizing first qualitative and then quantitative processes. Since this reality is unexplored empirically and theoretically we use a qualitative approach in chapters 2 and 3, with two different case studies of two different RIs. In both cases the methodology used is qualitative case study methodology. Chapter 2 explores the case of a synchrotron, which is a big scientific instrument. Chapter 3 focuses on the case of a biomedical database platform, which is a digital data infrastructure. In both cases we depart with a theoretical framework that aims to explain the role of these RIs on scientific creativity. We confront that theoretical framework to the fieldwork and then we complete the theoretical framework with the missing elements. As for Chapter 4, it consists in the development of a methodology and its application on a quantitative study, which aims to better clarify some of the conclusions reached in the previous chapters. The quantitative analysis allows us to better back up some of the knowledge constructed during the previous chapters and to describe more accurately the studied phenomena. This mixing of methods has been done in order to extend as much as possible our knowledge on the topics we are exploring and to be able to answer our research questions as accurately as possible.

4.4.1 Qualitative methodology

There are three elements that explain our choice of a qualitative methodology for chapters 2 and 3.

First, our epistemological positioning, critical realism, determines this choice. Although critical realism positioning can be based on both, qualitative and quantitative methodologies, Gavard-Perret et al (2012) believe that qualitative methods are particularly adapted to this paradigm. Other authors go in the same direction. Charriere Petit and Durieux (2007) believe that although doing exploration does not require a choice of a qualitative tool, it is more efficient as it allows for understanding complex dynamics and processes. Qualitative methods allow as well for the elaboration of conjectures and the perception and understanding of the mechanisms generating the deep reality and their modes of activation.

Secondly the literature review showed how the study of scientific creativity in the way we intend to do it are rather original and therefore there are little elements to build up on. Large research infrastructures are understudied, despite the particularities that make them unique places and the growing role they play in science. Additionally, the way we study scientific creativity is new. Indeed, the study of novelty and value together and the consideration of creativity as a collective process are not really explored yet by the literature, specially when looking at the specific case of research infrastructure.

The third and final argument consists in the complexity of the object that is studied. This thesis studies a collective process with a multiplicity of actors, mechanisms and other elements involved. Qualitative methodology has been pointed out as optimal when studying these kinds of complexities because they produce rich data with dense descriptions (Miles and Huberman, 2003 pp:26-28) and allow for the observation of multiple dimensions and a more thorough analysis.

4.4.2 The case study

The qualitative approach looks, to us, as the logical choice for the first part of this research work. Yin (2003) distinguishes five main methods: experimentation, investigation, archive analysis, history and case study. The choice of the research method is based on three criteria (Yin, 2003) which are the type of question asked by the researcher, the level of control that the researcher can exercise over the events he seeks to study; the degree of focus on contemporary events. Following the works of Yin, we have chosen to rely on case studies for chapters 2 and 3 of this thesis. The first reason is that this method is adapted for research aimed at understanding the "how" of a phenomenon. The second reason is that the case study methodology is recommended for research in which the researcher has little control over observed phenomena and behaviours as it is the case for us in this research. Finally, our research focuses on contemporary events and therefore the historical analysis is not adapted.

Additionally to the above mentioned reasons for the choice of case study research, case studies have, over time, become the most widely used qualitative methods in management science, proving this method is an efficient research strategy (Eisenhardt, 1989; Rispal, 2002; Yin, 2003). The reason of its success is that it explores a contemporary phenomenon, the case, in depth and in its context. This makes of it a very valuable tool for management research, especially when the boundaries between the phenomenon and the context are not clear (Yin 2003). The case study can be mobilized to improve, test or generate a theory and it is not affiliated with any firmly defined paradigm or reasoning although it is more often used in critical realism, constructivism and interpretivism (Eisenhardt, 1989).

This thesis has two chapters (chapters 2 and 3) relying on case study research. The specific strategies for each of the cases will be developed in the chapters themselves.

4.4.3 Methodological development and quantitative analysis

The objective of Chapter 4 is to develop a tool for the quantitative evaluation of scientific creativity. Today there is an absence of an accepted approach to the measurement of scientific creativity. Because of this, and basing the choice of unities of measurement on an exhaustive literature review, we develop a tool that aims to become a tool for the assessment and monitoring of research activity. More specifically we aim to use this tool to assess and monitor research activity that has been performed by using research infrastructure. The units of measurement are inspired by the recent research in scientific novelty. We, however, provide the originality of assessing novelty and value together. Our tool allows as well for comparison and therefore, rather than studying whether creativity is steady or growing, we assess whether creativity of a certain subset of research (such as the research performed at a specific RI) is more or less creative than the rest of the research in the field. The developed methodology is applied to one of the cases studied qualitatively in this thesis, the case of the Soleil synchrotron. The details on the design of the methodology are developed in Chapter 4.

5 Concluding remarks

The objective of this thesis is to answer the questions: **Does research infrastructure drive scientific creativity? How can research Infrastructure contribute to scientific creativity? Is it possible to measure this impact?**

These questions are relevant for many reasons. In first place, scientific advance requires creativity in order to face big societal challenges. We need science to keep making discoveries and creating new knowledge to understand the world and this requires creativity. This need for creativity is particularly true in the case of Large RIs because doing creative science is the very reason they have been built for. Indeed, RIs aim to push the knowledge frontier and contribute to big discoveries and to solve some of the big societal and scientific challenges. Additionally, RIs are becoming the more and more important in the research project. Science is experiencing several changes and one of them consists, precisely, in the increasing dependency in big instruments or big databases.

Although this topic is understudied, some researchers have studied creativity in science. Several scholars have looked, for example, at the organisational conditions under which creativity takes place in science. From this body of literature, we know that variety in its multiple forms has a positive impact on creativity. More specifically this variety consists of variety of knowledge involved in research, variety of methods, variety of backgrounds or multidisciplinary research. Other factors that have been linked to creativity in science are autonomy of researchers and the ability to communicate actively. Management literature has studied creativity and, although it rarely discusses the specific case of scientific creativity, there are several insights that we learn from that body of literature such as the important role of communities for creativity, the role of organizational learning and how creative processes work

Building on this research we will look at the case of research Infrastructure and at whether these favourable conditions and mechanisms for creativity are found at RIs. To do so we will look at two case studies (Chapters 2 and 3) of very different RIs, a big instrument and a biological database platform. We will try to analyse as well, qualitatively and quantitatively, the kinds of creative outputs that are produced at RIs. We will do so in a qualitative manner in Chapters 2 and 3, then we will develop a quantitative methodology and apply it in Chapter 4.

Chapter 2: Large Instruments as facilitators of users' creative process: The role of organizational factors, collaborations and communities in the case of a European Synchrotron

1 Introduction

Technological resources are an important and necessary input to do science and they consist mostly of scientific instruments, as well as the know-how related to those instruments. In several scientific fields the ability to perform science is strongly determined by the equipment available. It is crucial to keep up on the latest technological novel instruments allow for novel experiments and observations as well as for conducting known experiments in shorter time (Rosenberg, 1992). One can understand that in some fields of biology a microscope is essential and without it, it is not possible to do the needed observations and experiments. Stephan (2012) argues that for scientists the access to certain instruments can be the most important point in their decision of taking a job in one faculty or another because the pace of discovery is often determined by the access to these instruments. This chapter focuses on the case of large instruments, which are a specific kind of Large Research Infrastructures (RIs). They consist of sizeable instruments that cannot be built by single institutions and are therefore constructed at a national and very often supranational level. The costs and usage of such RIs have to be shared by multiple institutions[26].

The reason to focus on the case of big instruments is the growing dependency of today's science in the use of these big instruments. Scientific advance has often been considered as a precursor of technological advance. According to the "science pushed" linear model of innovation, science is supposed to provide the society with fundamental laws that allow for the understanding of the natural world. These laws and general knowledge can then be applied in order to solve specific technological problems. We know, however, that there is a bilateral interdependency between technology and science. The reverse causal relationship is also true, so that technology often sets the path of science (Rosenberg, 1984; Stephan, 2012). Let us look at the case of astronomy; recently the changes that have set the path in astronomical research are the development of very large telescopes. They are large scale facilities which contain the most advanced optical, radio or atmospheric measuring instruments. They are operated in collaboration between several institutions and several countries. The technology they contain is very complex,

---

[26] OECD Global science Forum. Report on Establishing Large International Research Infrastructures: Issues and Options https://www.oecd.org/sti/sci-tech/47057832.pdf

requires several years of planning and leaves them with a very long life-spam. Because of their high construction and operating costs, there are only a few of them in the world, and often only one equipped with the latest technology and specialized in a particular kind of technology. Yet their development is key to the advancement of our knowledge of the universe. This is not the sole feature that makes very large telescopes unique places. Research in these telescopes is often conducted by non-local teams which use the telescope remotely or go in-site in order to do observations. The access to these telescopes is often crucial for teams to be able to perform good quality science in certain domains.

This pattern is repeated in several other domains of knowledge. Large Instruments are being developed in several domains of science and the dependence of science on these Large Instruments seems to be growing. As the knowledge frontier is pushed further and the scientific challenges become more complex, the needed instruments become more complex as well. As science becomes more and more dependent on technology, authorities focus on assuring the access to these technologies as well as their development. The main reason is that these instruments are expected to allow science to advance and to do more path breaking discoveries. Actually, some of the objectives exposed by the European Commission related to Research Infrastructures (of which, large instruments are one type) are: "*responding to the rapidly evolving Science frontiers*", "*pushing the frontiers of knowledge*", "*solving our grand societal challenges*" or "*performing cutting-edge research*"[27]. Since big steps in science are often associated with large RIs while requiring a lot of creativity at the same time, it is rather unselfconscious to consider the possible effects of RIs on the creative process of their users. Do big instruments provide users with favourable conditions for scientific creativity? Do they help them to elaborate creative outcomes? In the case of a positive answer, what are the mechanisms that enable such a creativity and how do they play in the science production process? Surprisingly enough, this kind of issues has not yet been examined in the literature in innovation and science management, at least to our knowledge. In this chapter our aim is to contribute to fill in this gap through an in-depth empirical analysis in the case of a large synchrotron radiation source.

---

[27] European Strategy Forum on Research Infrastructures https://www.esfri.eu/objectives-vision

We have several reasons to think that large instruments are likely to favour creativity as they have several properties which make them a suitable place for creative science to occur. They are a place where multiple actors joint forces to pursue common objectives[28]. For instance, in many cases the size, scope and cost of research infrastructures require multi-lateral or even global agreement. They are also involved in international R&D networks, including networks of instruments where they cooperate in technological development and follow up of the main global scientific challenges. Additionally, to this, they are very often open to users coming from different organizations. In Europe, in line with free access policy and shared use and management, they represent a place of encounter where scientists from different countries and backgrounds meet.

We focus on the case of a synchrotron. A synchrotron is a particle accelerator, the main property of which is to create a very powerful source of radiation (e.g. a highly bright white light). Originally developed for the use of X-rays to explore inert or living matter, it was then declined along the whole spectrum going from infrared to hard X-ray. Electrons are accelerated in boosters until they reach the speed of light, then the highly energetic electrons are transferred and continue to move in a large storage ring. The whole technology depends on a physical phenomenon: when a moving electron changes direction, it emits energy. When the electron is moving fast enough, the emitted energy is at X-ray wavelength. A synchrotron machine accelerates electrons to extremely high energy and then makes them change direction periodically. The result is that photons are emitted in several beamlines all around the storage ring. All the beamlines correspond to different kinds of laboratories or experimental stations. The machine operates day and night, with periodic shutdowns for maintenance[29]. The range of scientific fields that use the synchrotron technology is very large: fundamental life science, drug fabrication and research, material science, computer chips design, chemistry, medicine, physics or

---

[28] European Commission on Research Infrastructure: https://ec.europa.eu/research/infrastructures/index.cfm?pg=policy

[29] European Synchrotron Radiation Facility (ESRF) http://www.esrf.eu/about/synchrotron-science/synchrotron

geological science use synchrotrons frequently, which doesn't mean that all need to use the same beamlines[30].

Synchrotrons in Europe are open equipment, the objective of which is to help researchers to perform high level scientific research. Demands for beamtime[31] are formalised in proposals and selected through periodical panel reviews. So the synchrotrons offer the possibility to ant kind of groups and organizations to make use of its advanced technology. Both small teams and large well-known research centres have access to the cutting-edge technology, regardless if they have significant or few resources. To make this effective there is no price to pay for those who produce publicly available science, normally university departments and other research units of Public Research Organizations (PROs). The foreign or distant users of a synchrotron, i.e. the researchers coming from outside the region or the country where it is located, can apply for European grants covering the cost of trip and accommodation.

Studying creativity at the synchrotron is particularly interesting for several reasons. First, despite its extended use in several domains of science, its creativity has, to the best of our knowledge, never been studied before. Its study is, nevertheless, very interesting given the characteristics of synchrotrons. In addition to being places to perform science, there is a big technological component to the activity. In synchrotrons, as it is the case for several other instruments, the latest technology is in continuous interaction with the latest scientific development. There is, as well, a collective dimension of synchrotrons that make them an interesting place for the study of creativity. Synchrotrons are used by multiple people with multiple origins (country, institution, etc). Additionally, they are multidisciplinary facilities as they are used by many different fields of research. This importance of the collective dimension leads us to think they are proper objects for the study of creativity. There are studies on team creativity however there are no studies that focus on creativity at a large instrument.

---

[30] European Synchrotron Radiation Facility (ESRF) http://www.esrf.eu/about/synchrotron-science/synchrotron

[31] It refers to time using the beamline

The chapter is structured as follows. We start by introducing the theoretical framework which consists on the study of scientific creativity at three different stages: the favourable factors for creativity, the mechanisms for creativity and the creative outcomes of science. After that we will present the empirical analysis which first explains the choice of methodology, the data and the method of analysis and then shows and analyses the results. Afterwards there is the discussion of these results and the conclusion.

2 Theoretical Framework

Except for Avadikyan & Mueller (2017)[32] we could not find any academic work regarding the creativity of large RIs. These authors investigate to what extent creativity and innovation management at the level of synchrotron organisation contributes continuously to the development of its technological platform. They highlight the role of different communities in the process, including RI-user's interactions. Our focus is somewhat different. We explore the impact of a synchrotron on the creativity of its scientific users. Although the role of instrumentation in general is not denied by scholars in history of science, who show that new instruments often lead to breakthrough discovery, the underlying mechanisms and impact on researcher's creativity remain poorly understood. To our knowledge there is almost no literature on that specific topic. This is even truer in the case of large shared instruments such as synchrotrons, even though their specific properties may well create a favourable context for users' creativity. There, is, however, an important body of research on creativity and on creativity in science that could help us to understand what we can expect concerning synchrotron impact on the scientific creative process. This literature is presented in the next subsection, by distinguishing between the three broad conceptual building blocks elaborated in Chapter 1 (cf. section XXX): the organizational favourable factors, the creative mechanisms and the scientific outcomes.

---

[32] AVADIKYAN Arman, MÜLLER Moritz (2017) "Management of creativity in a large-scale research facility" in The Global Management of Creativity (Eds) WAGNER Marcus, VALLS-PASOLA Jaume, BURGER-HELMCHEN Thierry, pp. 140-158, Routledge, NY.

2.1 Favourable factors for creativity at the synchrotron

In this sub-section we remind briefly some of the organizational factors propitious to scientific creativity as they were identified in the different streams of literature reviewed in Chapter 1 (i.e. psychology, sociology of science and management). We explore the possibility to apply them to the case of synchrotron users' creativity, focusing on those likely to have the most explanation power. A lot of scholars from multiple disciplines have focused on the organisational factors favouring scientific creativity. From all the factors highlighted by the literature there are three general groups that we consider as particularly relevant to the case of Large Instruments and more specifically to the synchrotron: (i) researchers' autonomy and leadership, (ii) scientific diversity, and (iii) communities. Their potential role is developed in the next subsections.

2.1.1 Scientific autonomy and leadership

Because Synchrotrons are facilities used by multiple individuals coming from multiple different places, we can expect the employees of the synchrotron to play a very special role as they will be in contact with this multiplicity. We can expect the Synchrotron (or more particularly the people working there) to be knowledge workers, as they will be at the interface between multiple different individuals which should give them an important role in creativity (Burt, 2004). Synchrotrons have several beamlines, which are a sort of scientific laboratory with specific technologies to each of them. Each beamline has an employee who is responsible for its management.

In Synchrotrons, beamline managers are responsible for the decisions made concerning technological developments at the level of the beamline as well as responsible for the acceptance (or the criteria of acceptance) of projects at the beamline. This autonomy, together with their scientific activity, is suggested by the literature to be positive for creativity. Indeed, as studied in Chapter 1, a factor that has been found to be positively correlated with the existence of scientific creativity is the independence of scientists when setting goals and the need to have as managers and leaders of scientific institutions; individuals that have been actively working in research themselves. Furthermore, it is positive for scientific creativity that the leaders of research groups are still active in scientific production. Let us start by the independence of scientists. Literature shows that when researchers have autonomy to make decisions such as which research paths to follow and by means of which strategies, , they are more likely to be creative (Pelz and Andrews, 1966, p146). Freedom choose individual scientific interests within or beyond a broadly defined thematic area is a central factor to explain which scientists and groups are highly creative because intuition and good knowledge on the subject are crucial and no one is better placed than the scientists themselves to have this (Heinze et al., 2009). Literature claims that science is more creative when performed in a bottom up logic, rather than having leaders deciding research agendas for the rest of the group (Amabile, 1996).

Similarly, Heinze et al (2009) find that in order to assure creativity of the group, leaders must be, themselves, involved in doing research. Active participation in the praxis of research is important for leaders as it helps them understand the problems of the group, how to motivate group members and how to organize a coherent research program. This finding is also reflected by Mumford et al. (2002) who suggest that leadership in creative environments requires predominantly technical and scientific expertise.. What literature explains us is that scientists themselves are the best judges when it comes to selecting their research questions or agenda, and how to perform research. For this reason, it is also important that leaders are scientists themselves and if possible are involved in scientific production. Because of the nature of the instrument, at the synchrotron we cannot talk about team leaders as such. The main role of beamline managers is to assist users. However, these employees of the synchrotron are scientists as well, meaning they are active in performing research. We can, therefore, expect synchrotrons to be a good place for creativity. The reason is that, although scientists at the synchrotron are not team leaders, they are often **involved in research** as shown, for example, in the EvaRIO report and the subsequent work by (Avadikyan and Müller, 2017).

2.1.2 Scientific diversity

Synchrotrons are, by nature, multidisciplinary instruments. There is a large variety of disciplines that uses them which brings to the same place people with multiple different competences and skills. Additionally, the Synchrotron by itself offers variety in terms of methods and technologies that can be used there. This diversity can play as a favourable element for creativity as literature has shown that creativity is enhanced by diversity.

Diversity or variety are concepts that have been linked to creativity for a long time. Management research has largely studied how the introduction of diversity in companies (diversity of employees profile and cultural background, diversity of work methods) can foster creativity (Chamorro-Premuzic, 2017; McLeod et al., 1996; Paulus and Nijstad, 2003; Rodan and Galunic, 2004). Lee et al (2015) show this positive impact of diversity in scientific creativity. This diversity consists on diversity of backgrounds involved in the research, diversity of methods used and diversity of disciplines involved. This kind of diversity is often found at synchrotrons. In a similar line Hollingsworth (2002) argues that great creativity in science is typical for research organizations where scientists communicate across the boundary of the discipline and the research topic. It is also important that leaders endorse a strategy that aims to integrating scientific diversity. Because synchrotrons have several disciplines coexisting in one single facility one could expect that there is (a minimum) exchange among these disciplines and this would favour creativity. For example, because the Rockefeller University was organized around laboratories instead of being organized around scientific disciplines and fields, it had a greater capacity to adapt quickly to research strategies and to allow effective communication across cognitive boundaries (Hollingsworth, 2004, pp 34-35). Teamwork between departments has also been recognized influencing positively creativity (Hage and Mote, 2010). Similarly, Heinze et al. (2009) show the importance of access to a relatively large variety of technical skills in order to achieve creative science. Hage and Mote (2010) argue as well that plural organizational leadership ensures diversity of research strategies and richness in ideas. The three directors of Institute Pasteur (which they studied) operated with diverse recruitment patterns. More specifically, they recruited scientists with a diversity of backgrounds.

Scientific diversity has been claimed by literature as the most important contributor to scientific creativity. This diversity is most often expressed in terms of variety of background in research teams, as well as multidisciplinary in its multiple forms such as multidisciplinary projects or multidisciplinary institutions. Multidisciplinary research refers to the situations in which several different research disciplines are present. Multidisciplinary research has traditionally been linked to creativity as it brings together schemes of thought (with ideas, concepts and viewpoints) and methods. It can consist on multidisciplinary teams, where individuals come from different backgrounds; on research projects with different teams from different disciplines or on laboratories or other institutions that gather teams from multiple disciplines. The closer and more frequent the interactions of the individuals from different countries is, the higher are the chances of creativity to occur(Alves et al., 2007, 2007; Heinze and Bauer, 2007).

The role of communications will be detailed later when discussing about creative mechanisms.

What are the implications of variety in the case of RI user creativity? Are RIs and in particular big instruments such as Synchrotrons likely to offer the knowledge variety needed for creativity?

There are several reasons to believe that we can find variety at a synchrotron. We can in first place expect to find a **multiplicity of scientific and/or technological backgrounds**. The first reason to believe this relies on the fact that we have technological experts on the instrument on the one side, i.e. the beamline scientists, and researchers from multiple disciplines coming to the synchrotron to carry out experiments, i.e. the RI external users, on the other side. This by itself entails the encounter of different scientific and technological backgrounds. A second reason relates to the fact that the synchrotron is an instrument that is used simultaneously by numerous users from diverse scientific disciplines. This entails a potential for unexpected inter-users encounters and exchanges Because of all these encounters of diverse backgrounds we can also expect, in a second place, a few **multidisciplinary projects**, which do associate complementary competences coming from more than a single scientific discipline. Last but not least, and strongly related to the previous points, we expect to observe the encounter of a **multiplicity of communities**[33]. Individuals with different backgrounds belong often to separate communities of scientists, but the concept of community is more demanding than the notion of scientific background or disciplinary knowledge. Communities have rules, methods, ways to communicate and perspectives. The notion of communities, their role and the way they may interact - or not - at a synchrotron require additional analysis, which is developed in the next subsection.

---

[33] The concept of communities will become very relevant later on this Chapter and it is extensively developed in next section.

2.1.3 The role of communities

This section aims to study the literature that investigates learning, creativity and innovation in the context of knowledge communities, focusing on the kinds of communities and dimensions likely to be relevant in the synchrotron case. Several types of communities are likely to be present at a synchrotron. At the very least, we should observe a kind of "technology based" community composed of engineers and scientists who develop and maintain the instrument. We should also observe several user communities, belonging to broader and more conventional academic, disciplinary-based communities. We can also consider different scopes, or concentric circles, for these communities (e.g. from specialists of a specific detector to the whole range of particle accelerators, from users of a specific beamline - or belonging to a narrow research field - to the whole set of synchrotron users, etc.).

As discussed in the first chapter of this thesis, communities have been found to be crucial in several aspects that drive creativity. We will present two lines of ideas. First, the "raison d'être" of some communities is precisely to learn and produce new knowledge or knowhow. Hence they are intrinsically creative. Second, when they meet the members of external organizations, companies or other communities, they become a source of innovation for the latter. In turn these creative communities are themselves fed by contacts with external entities, thus the co-location of diverse communities may lead (under propitious condition) to the co-creation of new knowledge and capabilities.

In the following lines, we explore different concepts of communities[34] in order to characterize the creative communities present at a synchrotron. Then we go beyond intra-community creation, pointing up the possible existence of trans-community knowledge sourcing as well as inter-community co-creation.

---

[34] An extensive review about the well known and rich management literature on communities would fall well beyond the scope of the present chapter. Here the aim is modestly to complement the notions presented in Chapter 1, so as to provide useful conceptual elements for analysing the creative process of synchrotron users.

2.1.3.1 A brief review on knowing communities

The term "knowing communities" refers to those communities which aim towards the creation of knowledge (Cohendet and Diani, 2003; Harvey et al., 2015). Broadly defined as groups of people concerned by the same topic and building collective knowledge about it by interacting with each other, they encompass different kinds of creative groups that of interest to us, among which the Communities of practice (CoPs) and the Epistemic Communities of the literature in management as well as the Invisible Colleges and other scientific communities identified by sociology of science.

Communities of Practices

CoPs, are defined as groups of individuals that develop the same practice or, in other words, the same kind of activity. They are characterized by sharing a set of practices that include exchange of tacit knowledge by communicating and exchanging information frequently. What is particular about this concept is that communities of practice do not need to share a common institutional affiliation and yet there is a feeling of identity and belonging. Brown and Duguid (1991) define CoPs as groups of individuals who gather together to propose a solution to a problem (often a technological problem but not always) when a formal group, for instance a firm, fails to do so. There is an emphasis on the fact that the lack of flexibility of formal institutions prevents them to addressing new problems, and therefore there is a need for these informal groups or communities. Organisations should embrace these communities as they are efficient on solving new problems and as they represent a good place for informal learning. But the distinctive property of CoPs is more on the former advantage than on the second, that is, more on finding solutions to new problems than simply transmitting knowledge. Because of this, CoPs are central in the study of innovation and creativity. By means of case study research, literature has shown that CoPs trigger innovation (Lesser and Storck, 2001; McLure Wasko and Faraj, 2000; Wenger et al., 2002).

The way scientists work has several elements of a CoPs. Scientists work together pursuing the same goals, which are often linked to solving the global scientific challenges. The way scientists organize themselves is typically not constrained by the limits of a single organization and scientific communities are informal in the sense that all members do not share the same affiliation. The concept of CoPs seems especially relevant in the case of synchrotron scientists and engineers, as they devote themselves to developing and maintaining the instrumentation, as well as the associated methods and practices. Also, they exchange knowhow and technology with the staff of other synchrotrons all around the world. Unlike conventional CoPs however, they aim at knowledge creation, namely the engineering science around their instrument. This is typically a characteristic of another type of communities, i.e. Epistemic Communities, which are presented below.

Epistemic Communities

Whether they are composed of synchrotron staff or synchrotron users, scientific communities also have important characteristics of Epistemic Communities. The way scientists organize themselves is based on rules that, although they are not always written, strongly shape the interactions within the community. Because of this, and also because their essential goal is to create and transfer knowledge, scientific communities are often considered as Epistemic Communities. Used to describe the communities who aim at knowledge production, the phrase "Epistemic Communities" refers to *"knowledge-based network of specialists who share beliefs in cause-and-effect relations, validity tests, and underlying principled values and pursued common policy goals"*(Haas, 1992, p187), or to *"Small groups of agents working on a commonly acknowledged subset of knowledge issues and who at the very least accept a commonly understood procedural authority as essential to the success of their collective [knowledge] activities"* (Cowan et al., 2000 page 234). In other words, Epistemic Communities are groups of people who share a common field of studies and whose common goal is to advance the knowledge of this field of studies. A particularity of these communities is the existence of a set of norms helping to give validity to the produced knowledge. Latour et al (1979) studied in "Laboratory Life" how science is organized. They identified the existence of several social norms as well as procedural norms in the way scientists organize themselves. These norms concern the way scientific work is conducted, descriptions of the complex relationship between the routine lab practices, the publication of papers, scientific prestige, relation to other labs, research finances and other elements of laboratory life.

Both epistemic communities and CoPs are terms that have been widely used and for which several definitions have been given. Because of this, the idea of "Knowing communities" have been developed to cover them all. It concerns all groups of people who share a set of problems or a passion about a topic and who aim to increase their knowledge and interact on an ongoing basis (Harvey et al., 2015). In this context of several concepts and definitions to describe knowing communities, where do communities of scientists lay? Communities of scientists, for us, are knowing communities that are both epistemic and of practice. The concept of invisible college below represents it quite accurately.

Scientific communities and Invisible Colleges

The term "Invisible College" was first used in the 17th century when some scientists in Europe referred to themselves as an "Invisible College" because they did not belong to any formal institution, while meeting frequently to share information about their common scientific interests and to monitor the advance on their field. In modern literature it has been used again to describe groups of scientists who specialize in a specific subfield of studies and who use informal channels of communication and collaboration (Crane, 1969; Cronin, 1982; De Solla Price and Beaver, 1966).

Recently, Zuccala (2006) proposes a definition, including all the different aspects that the literature about Invisible Colleges has considered important: *"An invisible college is a set of interacting scholars or scientists who share similar research interests concerning a subject specialty, who often produce publications relevant to this subject and who communicate both formally and informally with one another to work towards important goals in the subject, even though they may belong to geographically distant research affiliates*."

Figure 8 illustrates this definition. Invisible Colleges correspond to that space of interaction between the discipline itself, the social interactions between scholars and, which is of interest for the purpose of this chapter, the physical and technological resources. This idea of Invisible College includes all the aspects of both CoPs and Epistemic Communities: there are formal and informal interactions, disciplinary norms and rules as well as practices concerning the use of resources. The stress on the use of physical resources makes the concept of Invisible Colleges particularly relevant to the case of synchrotron users. Such communities of scientists are, therefore, invisible colleges, but also epistemic communities and communities of practice.

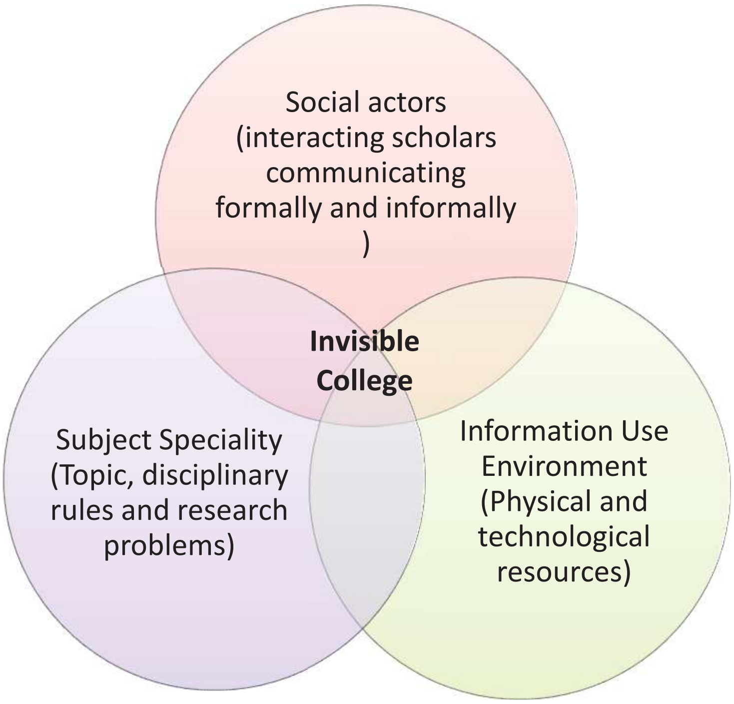


Figure 8 Invisible colleges (Adapted from Zuccala 2006: Modelling the invisible college)

Interstitial communities of "research technologists"

The case of instrumentation calls for a dedicated concept of community. In their instructive chapter "*A fresh look at instrumentation*", (Bernward and Terry, 2001) revisit the notion of Research-Technology, which they define as *"instances where research activities are oriented primarily toward technologies [instruments and methods] which facilitate both the production of scientific knowledge and the production of other goods*". Connected to the increasing sophistication and large-scale of technical systems, the growing cognitive specialisation of scientific activities and, finally, to the scientization of engineering, Research-Technology is developed by particular communities. The latter are referred to as **interstitial communities**, because they are characterized by fluidity, hybrid carriers and more generally by a trans-community positioning. Research-technology practitioners work with universities, companies, states and other organizations, but they remain separate from them. They work on **generic devices** that are open-ended, flexible and can be used in several applications (including industrial and military ones). The devices are designed to solve problems requiring precision detection and measurement in general. They may even "offer answers to questions that have hardly been raised". In accordance with this high degree of flexibility in use, the practitioners also develop standardized languages in order to diffuse the device as broadly as possible.

The parallel with synchrotron scientists and engineers is appealing: in synchrotron we have a community operating a generic (set of) instrumentation allowing for (outstandingly) precise detection and measurement with multiple applications. The idea of interstitial community points out an interfacing role and an autonomous capability including both experimentation and theory. Due to this trans-community positioning, we expect that synchrotron scientists will have a positive impact on the creativity of their users. More specifically we expect that they will understand users' preoccupations and help them to create valuable knowledge, eventually they will provide them with complementary competences. They may even enter in a co-creation process. These trans- or inter-community learning processes are explored in the next subsection.

2.1.3.2 Crossing of communities: knowledge sourcing and co-creation

Authors studying the concepts of communities above have insisted on them being the place for the generation of new ideas. But they highlight also additional explanations concerning the sourcing and generation of these new ideas.

Sharing knowledge within a given community is not enough. The ability to offer solutions to current problems depends also on the ability of the community to develop absorptive capacity[35] and thus to use knowledge which is external to the community (Cohendet et al., 2013). Indeed, pooling only the knowledge produced at the intra-community level and finding the way for individuals to learn and enrich it may not guarantee a high degree of creativity, especially in the case of a rather old and mature community. Individuals need to learn as well from the knowledge produced outside the community, and they must be able to combine it in proper and creative ways to solve the most complex problems that they face.

The role of communities for external knowledge sourcing is widely acknowledged in the literature. Well known investigations analyse specifically how some users' community, especially in the case of lead users, are crucial for internal innovation efforts (Hienerth et

---

[35] Concept better developed in Chapter 1

al., 2014; von Hippel, 1976). All of this makes us think that connexions with external communities could have a relevant role to play in Research Infrastructure and particularly in synchrotrons. Actually, part of the research conducted within the EvaRIO project tend to support the idea that the openness to external communities plays a important role in the continuous creation of internal technological knowledge and innovation, making synchrotron a highly creative place. Expanding this idea further, Avadikyan and Müller (2017) show how the synchrotron becomes a place where many communities meet, and how this crossing of communities triggers creativity within the synchrotron organization itself. Hence, one can easily hypothesize that the reverse influence, from synchrotron to user, is also relevant. More specifically, working with a synchrotron team or meeting another user community at the synchrotron site could exert a positive influence on given users' creativity as well.

To summarize the above discussion, we expect that, due to its broker position between heterogeneous groups, a synchrotron team of "research technologists" (quoting Joerges and Shinn, 2001) can draw from different sources of knowledge and its scientists are expected to be very creative. The point is, this creative potential may well lead to feedback loops toward other communities, specifically the creativity of synchrotron users, which is the core issue of this chapter. More generally, the users as well can profit from the heterogeneity of knowledge and competences of the other communities present at the synchrotron (here other communities mean those of synchrotron teams but also of other users' groups), if they hold the required openness and absorptive capacities. Moreover, bi-directional exchanges of knowledge pave the way for interactive learning and co-creation processes. Hence the crossing of multiple and diverse communities is an additional topic of interest. Nevertheless, the underlying creative and combinatory mechanisms have to be analysed in more detail. This is the object of Section 2.2, in an attempt to review relevant concepts for analysing creative mechanisms at synchrotrons.

What are the communities at stake at a synchrotron? What could be their impact on the users' scientific creativity? What are the underlying mechanisms? These are open questions into which we aim to give some insights via the SOLEIL case study. Hence, one of the key points to be examined empirically will be the role of the synchrotron SOLEIL as a hub of communities where actors from various communities meet and interact. Below we sum up the other concepts and main hypotheses which emerged from the literature overview up to this point, and which will feed an important part of the field study in Section 3, concerning the organizational factors presumably present at the synchrotron site and propitious to users' creativity.

2.1.4 Summary of factors and objectives for the case study

The following table is a wrap up of what we will be looking for in the empirical analysis of the favourable factors for creativity at the synchrotron. We have seen three groups of factors that the literature identifies as favourable for scientific creativity.

The first group of factors concerns the role of managers and it includes two different factors: beamline managers having autonomy to conduct their research and beamline managers being involved in research. Literature shows that when scientists have autonomy to make decisions on their research agenda, they are more creative. It shows as well that when managers who make important decisions are scientists themselves, these decisions will be better for creativity.

The second group of factors that is described by the literature as positive for creativity and is likely to be found in a synchrotron is scientific diversity. This diversity can come in the form of variety of backgrounds of people in terms of discipline, variety of methods that are used and multidisciplinary projects. Research shows that when science is done including ideas concepts and viewpoints from different disciplines it has more chances to be creative. It happens the same when it comes to the methods used, the more varied are the higher the chances to come up with new ideas. Finally, multidisciplinary research, which is the one that aims to solve a research question by putting together different disciplines, has also been traditionally linked to creativity.

The last group of factors consists on the existence of communities. The existence of active communities has been described by several scholars as a crucial factor for creativity. Because communities allow easy and efficient communication, they allow individuals to share their ideas and this exchange of ideas is what favours creativity. Because synchrotrons contain cutting edge technology, we can expect to find interstitial communities which are communities that consist of experts around a particular technology. Because of its multidisciplinary we can expect the synchrotron to bring several communities together and allow scientists to pull knowledge from the other communities (such as formal knowledge, methods and skills). Finally, we can have this learning at a higher level with strong exchanges between communities who would actively share knowledge.

The following table summarizes all of this factors that we will explore at the interviews. Our objective is, rather than confirming their existence, to understand whether these favourable factors are present and by means of which mechanisms they participate to the collective knowledge creation process.

Table 1 Favourable factors for creativity

Scientific autonomy and leadership	**Scientific diversity at synchrotron**	**Creative communities at synchrotron**
Beamline managers' research autonomy	Variety of disciplines and backgrounds	Interstitial communities and invisible colleges
Beamline managers involved in research	Variety of methods	trans-community knowledge sourcing
	Multidisciplinary projects	crossing of diverse communities

As one may notice, all factors are not completely independent from each other. For instance, the diversity of scientific backgrounds depends on the communities they belong to; the autonomy of the synchrotron researchers obviously impacts their agenda and hence their participation or not in trans-community or multidisciplinary projects, as does the involvement of team leaders; the role of interstitial communities and knowledge sourcing would not appear as relevant without scientific diversity, etc.

More importantly, all these favouring factors would remain ineffective in the absence of interactions and communication mechanisms. In other words, creativity means that variety has to be interconnected and combined. Before we proceed to the case study, we have thus to investigate in more detail the **mechanisms** of collective creation at stake: communications, transfers, collaboration and other interactions making it possible to connect different pieces of knowledge in a combinatory fashion.

2.2 The creative mechanisms: combining complementary pieces of knowledge through effective communications and interactions

Building on Chapter 1 presentation about the Koestler's concept of bisociation, and more generally on the knowledge combinatory principle at play behind any creative activity, we focus on the connection mechanisms between the knowledge bases of diverse people, groups and/or communities.

In the literature in psychology about creativity in science, effective interaction with other scientists are supposed to favour the productivity of the individual researcher. In the 70's, Pelz and Andrews (1976) had already shown that research productivity was correlated with a high frequency of intra-organizational communication. This idea has been recently confirmed by several studies in scientific creativity, which systematically found that organizations producing highly creative research allowed for cross discipline interconnexions via a number of mechanisms such as multidisciplinary projects, multidisciplinary departments or simply by enabling communication among heterogeneous groups of the same organization (Heinze et al., 2009; Hollingsworth, 2002; Simonton, 2004; Zuckerman, 1967). Others show how organizing science around research questions and challenges rather than disciplines have a positive impact on creativity (Hollingsworth, 2004).

The same idea is also supported by research focusing exclusively on collaborations. Collaborations can be done between individuals, between teams and between institutions. Due to the growth of co-authorship collaborating between individuals have become the norm. Most of the research looking at collaboration in science would focus on collaboration between individuals and between teams in the form of joint projects and partnerships(Adams et al., 2005; Bozeman and Boardman, 2014). Several authors present collaboration as a positive factor for creativity (Adams et al., 2005; Aragon et al., 2009; De Solla Price and Beaver, 1966; Zuckerman, 1967). In the continuity of this idea of intense interactions as a prerequisite for creativity, mobility is also considered as good for scientific creativity. Scientists changing organization a few times during their career or spending some periods in other organizations can draw from different know-how and different perspectives that they can use, introduce and combine with their knowledge and abilities (Hage and Mote, 2010). Finally, and still following this line, the literature on network brokerage argues that people who are placed at the intersection of heterogeneous social groups have an increased likelihood of drawing upon multiple knowledge sources which leads to the generation of new ideas. For example, managers who occupy brokerage positions are more often than others the source of good ideas (Burt, 2004; Rodan and Galunic, 2004).

Different types of connexions should be distinguished, inasmuch as they lead to more or less radically new creative results. The distinction we propose is twofold. The first criterion concerns the direction of the connexion: a connexion between two knowledge bases can be **unilateral versus reciprocal**, thus making it possible to differentiate between mono-directional transfer versus bilateral exchange of knowledge. The more the connexion is based on a reciprocal, bi-directional basis, the more it is likely to trigger interactive learning, that is, a true co-creation process. Hence, the more it is appropriate to talk about a **new combination** of knowledge, in the sense defined in Chapter 1.

The second criterion concerns the intensity / length of the connexion. Here the idea is to differentiate occasional exchange(s) of information from continuous co-working during a given period of time. The more the connexion is intense and long-lasting, the more it tend to be effective, that is, the more it is likely to lead to a new combination. The underlying intuition is that **effective connexions** are a pre-requisite for creativity to occur.

As a result of the previous discussion we will consider different types of connexions at synchrotron. Mono-directional connexions will be referred to as knowledge transfers, less occasional and less intense connexions will be referred to as communications, while bilateral and intense connexions will be referred to as collaborations.

Effective bilateral connexion, i.e. collaboration is therefore crucial in order to be able to extract value from the co-location of diverse knowledge bases. Nevertheless, knowledge transfer and (less intense) communication have also a role to play, for they may be a first step toward a more intense relationship. We believe that communication is likely to be found at the synchrotron in a couple of forms. First, we can expect **fluid formal and informal communication between users and the synchrotron**. The users of the synchrotron being external to the synchrotron itself, there is a big need for coordination and therefore we can expect formal communication to appear. As for informal communication, it would come from the need for the external users and the scientists at the synchrotron to understand each other in order to actively work together. Additionally, we can expect to find **fluid informal communication between users** since they all share a physical space and to some extent objectives from its use. In all the above cases, we also expect that these communications sometimes turn into effective knowledge transfer. They may even turn into collaborations, that is to say, co-creation processes, provided that actual knowledge recombination and interactive learning actually occurred and led to co-creation.

2.3 Creative outcomes: The synchrotron contribution to users' creative results

By the end of the collective creative process we find the new knowledge which is the outcome of the process. This section aims, by means of literature review, to anticipate which kind of creative outcome could result from the use of synchrotron. As it is the case when studying the favourable factors or the mechanisms for the emergence of scientific creativity when it comes to creative outcome, literature has not focused on the case of research infrastructure. We will, therefore, use the combination of the literature on scientific creativity, research infrastructure and scientific instruments to unmask the kinds of scientific outcome that are likely to appear around the usage of a big instrument.

As it has been shown in Chapter 1, the study of creativity has been approached by researchers in multiple disciplines. Here we aim to understand how Large RI can contribute to creative scientific outcome. Although traditionally the study of scientific outcome has focused on scientific publications in peer review journals, the reality of what is the outcome of the collective science creation process is complex. Indeed, most of the scientific results can be published. However, not all the outcome of the process is published and not all the published outcome takes the same form (Heinze and Bauer, 2007). Scientific outcome is not homogeneous in form and can manifest in a multiplicity of ways, from the observation of a new natural phenomenon to the development of a new theory (Amabile, 1983; Ford, 1996; Woodman, 1993). Recently Heinze et al (2009) propose a list of creative scientific outcomes. This list consists on the following:

1. Formulation of new ideas that open a new cognitive frame or bring theoretical claims to a new level of sophistication
2. Discovery of new empirical phenomena that stimulate new theorizing

3. Development of a new methodology by means of which theoretical problems could be empirically tested.
4. Invention of novel instruments that open new search perspectives and research domains
5. New synthesis of formerly dispersed existing ideas into general theoretical laws.

This classification helps to better understand which kind of outcome we consider. New knowledge is creative if it provides new insights on relevant scientific problems. This means that creative research needs to be original. However, creativity is not only defined in terms of newness. For research to be creative it must be useful as well. Citations have often been considered as a good indicator of creativity because they consist on the recognition by peers of the value of a given scientific production. However, as it has been explained in Chapter 1 about scientific creativity, from the previous criteria the one that is the more often not acknowledged is the criteria of novelty. Usefulness, on the other hand, is often rewarded via publications, grants and citations. Because of this we will be using qualitative analysis to try to determine whether Large research Infrastructure and more precisely Large instruments can offer the favourable conditions to scientific creativity.

2.3.1 Knowledge creation

The first kind of outcome that we could think of would include points 1, 2 and 5 from the previous list. We can, indeed, expect the synchrotron to be a place of knowledge creation in both ways: the generation of new ideas and the discovery of new empirical phenomena. The synchrotron is a tool to perform science and the main expected results of its use is science. Science is mostly measured today in terms of publications and the synchrotron has, in its website, a list of publications that have been done using it. So the production of new knowledge is clearly observed. But to what extent can it be qualified as creative ? There are a few reasons, additional to the ones mentioned in the previous section, to believe that synchrotrons could be generating a creative scientific outcome. As seen in Chapter 1, creativity is defined in terms of novelty and value (or impact). The latter dimension is further analysed in section 2.3.3 about quality of the research outcomes.

2.3.2 Innovation

This section will include points 3 and 4 from the list above. The synchrotron containing instruments that are necessary to produce research, we can expect creativity in terms of technological development. For instance, new methodologies are likely to be developed. The synchrotron being a multidisciplinary tool we can expect for its managers to seek for technological improvements as well as new ways to apply it. This has already been suggested by Avadikyan & Mueller (2017). The relevance of these innovations for synchrotron users is twofold: first it has an impact on the quality of their research (see 2.3.3 below); second, we expect to find some cases of co-development of methods and/or instruments with lead users.

Several reports from the European Commission and the OECD go in this direction as well, they point at research Infrastructure as places that are expected to provide the Scientific landscape with the cutting-edge technology needed for science to advance. Furthermore, RI are expected to collaborate with each other in order to keep the technological developments advancing as a way to make the knowledge frontier advance[36].

2.3.3 Quality and impact

Finally, we have several reasons to expect high quality research at the synchrotron as a possible outcome. Quality of research can be defined in multiple ways but traditionally the focus is put on its impact and more specifically on the recognition or perception of value attributed by peers. More specifically we can observe this by observing where the research is published, the grants that it obtains, etc. One of the characteristics of creative research is to be highly valuable (Simonton, 2004). Another reason to expect quality and impact from research done at the synchrotron is that quality is the first objective expressed by the European Commission for its Policy on Research Infrastructure. “Excellence” and “Quality” are words that appear constantly in the multiple policy documents that we can find on-line. Finally, the reasons that lead us to think that we can find this quality research in the synchrotron is related to the EvaRIO project, that had

---

[36] https://www.oecd.org/sti/sci-tech/47057832.pdf and http://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-infrastructures_en.pdf

results pointing to this direction. More specifically, the EvaRIO project on Research Infrastructure already studied the case of the synchrotron from the point of view of impact evaluation and found that, for a given set of beamline users, the research done at the synchrotron has a relatively higher impact than research done without it by the same group of users (Avadikyan et al., 2014). This EvaRIO result points out the excellence, that is to say the **value dimension** of the research outputs resulting form synchrotron use, which is already an important condition for creativity.

2.4 Conceptual framework and Research gap

As explained in the introduction, literature has not formally defined which the impact of big instruments in scientific creativity is. Creativity is not studied in impact and evaluation analysis and it is, however, crucial for the advancement of science. Literature does give some insight on which the conditions under which scientific creativity are is more likely to happen. With this information in hand we have deduced which of these conditions are likely to be found at the synchrotron and the next section will, by means of qualitative case study, investigate whether these conditions are present there or not. When it comes to the possible creative outcome the same will happen. Crossing literature has allowed us to understand which kinds of creative outcome are likely to happen at the synchrotron, next section allows us to empirically investigate it. Finally, we aim to investigate which is precisely the role that communities have at the synchrotron and whether they have an impact on users' creativity. The next table summarizes the topics that we will be looking at.

Table 2 Creativity: favourable factors, mechanisms and outcomes

Favourable organization factors for creativity at the synchrotron	
Role of leaders	Infrastructure run by scientists
	Synchrotron scientists involved in research
Scientific diversity	Variety of disciplines and backgrounds
	Variety of methods
	Multidisciplinary projects
Communities	Interstitial communities and invisible colleges
	Trans-community knowledge sourcing
	Meeting and crossing of diverse communities
Creative mechanisms at the synchrotron	
Combining complementary competences via effective interactions	Fluid formal and informal communication with the synchrotron scientists
	Fluid informal communication among users
	Collaboration
Creative scientific outcome at the synchrotron	
Kinds of creative outcome found	New knowledge
	Technological development
	Quality of research

In first place we look at the favourable factors for creativity. Among them there is the role of team leaders which is divided in two favourable factors for creativity: infrastructure run by scientists and synchrotron scientists involved in research. In second place there. There is as well the topic of scientific diversity which consists into three favourable factors: variety of disciplines and backgrounds, variety of methods and multidisciplinary projects. We have, as well, the synchrotron as a hub of communities.

After the favourable factors for creativity we focus on the creative mechanisms at the synchrotron which consist on fluid formal and informal communication with the synchrotron scientists, fluid informal communication among users and collaboration.

Finally, we have the results of the research, which consist in creative outcome such as new knowledge, technological development and quality of research.

The following figure shows our theoretical framework. We see the collective knowledge-creation process with a variety of external knowledge going into the process. It does not enter only at the level of combination of formal knowledge but also at the level of Analysis. We think, indeed, that the synchrotron being a technological platform, users can profit extensively from the technological knowledge of its scientists and of other users.

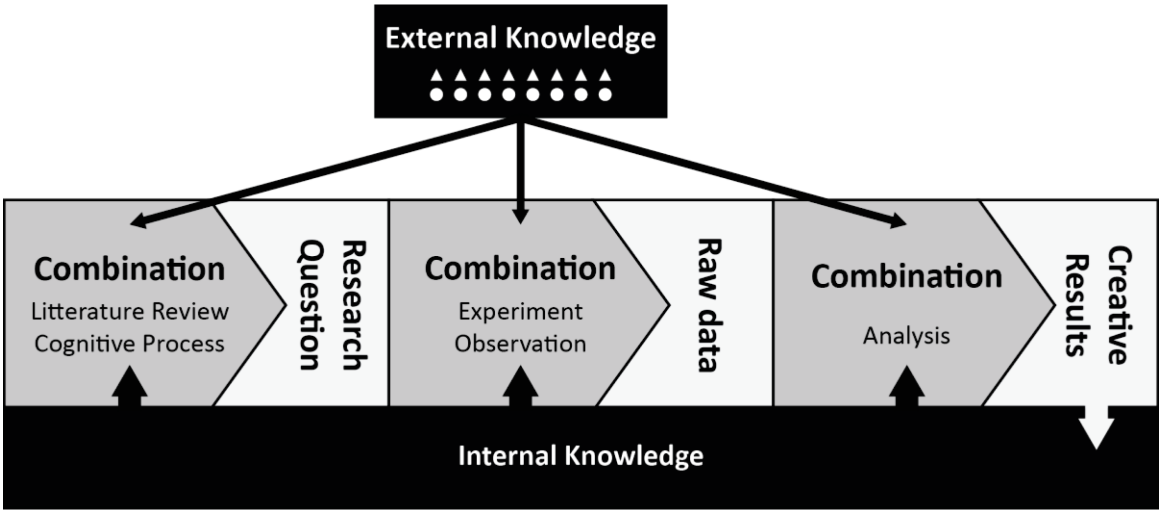


Figure 9 Collective knowledge-creation process

3 Empirical analysis: A qualitative case study about the synchrotron SOLEIL

This section is dedicated to an empirical inquiry concerning the role of a synchrotron in the scientific creative process of its users. To explore this question, we carried on a qualitative case study based on around twenty interviews of RI users and scientists managing or operating the RI. Before explaining our methodology and our research results in great detail, we first present the context of our research and the case itself.

3.1 Context and presentation of the case: The EvaRIO project and the Synchrotron SOLEIL

The empirical study for this chapter has been done in the context of the continuity of the project EvaRIO, a Research project at the University of Strasbourg, integrated in the

Framework Program 7 of the European Commission. EvaRIO (which stands for Evaluation of Research Infrastructures in Open innovation and research systems) is a Coordination and Support Action project funded by the European Commission under the 7th Framework Program (grant N° 262281 - Research Infrastructures INFRA-2010-3.2). EvaRIO is about finding ways for evaluating Research Infrastructure for Biomedical Research (Avadikyan et al., 2014). As largely explained in Chapter 1, RIs are an important aspect of the European Science Policy. Research Infrastructure are expected to be centres of reference were high quality research is performed and where the European Research Area is developed. This means that the objective is, on one hand to provide with significant improvement to technological and scientific questions and on the other hand to do this with a European perspective, rather than a national one. Access to all countries of the European Union and other European partners, mobility of researchers and diffusion of knowledge are the main policy objectives[37].

SOLEIL, the acronym for "Optimized Source of LURE Intermediary Energy Light," is a synchrotron research facility located on the Plateau de Saclay in Saint Aubin (Essonne area near Paris). It is publicly funded by two principal shareholders, the CEA and the CNRS holding 72% and 28% of its shares respectively. Other important partners are the Ile de France and Centre Regions, the Essonne department, and the Ministry of Research. Inaugurated in December 2006, SOLEIL is a public sector company with the status of Société Civile de Recherche (Civil Society of Research).

We have explained, at the beginning of this chapter, what synchrotrons are and how they operate. When it comes to research domains the synchrotron covers the fundamental research needs in a multiplicity of areas such as physics, chemistry, material sciences, life sciences (notably in the crystallography of biological macromolecules), earth sciences, and atmospheric sciences. In applied research, SOLEIL can be used in many various fields such as pharmacy, medicine, chemistry, petrochemistry, environment, nuclear energy, and the automobile industry, as well as nanotechnologies, micromechanics and microelectronics It offers the use of a wide range of spectroscopic methods from infrared to X-rays, and structural methods in X-diffraction and diffusion.

---

[37] European Commission on partners and networking https://ec.europa.eu/info/research-and-innovation/partners-networking_en

Figure 10 Synchrotron Soleil. (Source: https://www.synchrotron-soleil.fr/)

Figure 10 shows what the instrument looks like. The synchrotron radiation is light emitted by relativist electrons (at virtually the speed of light) of very high energy that spins in a storage ring of 354 m circumference. This is done tangentially to the trajectory in an extremely fine beam, and the trajectory of the electrons is curved with a magnetic field. A beam of electrons as fine as a strand of human hair, emitted by an electron cannon is first accelerated in a 16 meters long linear accelerator. After this initial acceleration, the electron beam is directed towards a second, circular accelerator called the Booster that brings the energy level up. Then the electrons are injected into the 354-meter circumference storage ring and spin for several hours. In the storage ring, magnetic devices control the trajectory of the electrons or make them oscillate. The electrons then lose energy in the form of light, the synchrotron radiation. This radiation is directed, selected, and stored by optic systems toward experimental stations called beamlines. Each beamline constitutes a true laboratory for biology, chemistry, and Earth sciences, equipped to prepare and analyse samples to be studied and process the information gathered. In this chapter we will focus our attention on only two of the beamlines of the synchrotron, the X-Ray beamline and the Infrared beamline. The reasons for this choice are the following: both beamlines have users in the field of biology, but Infrared is more multidisciplinary, which may lead to higher creative processes and outputs. Moreover, the corresponding instrumentation and methods are relatively new and emerging, since this type of beamline has been implemented at synchrotrons only for a decade and there are very few of them all over the world at the time of the interview.

The x-ray beamline is entirely dedicated to measurements in macromolecular crystallography. Biological samples take the form of crystals, the creation and growth of which are usually highly uncertain and may need years. Based on X-ray diffraction an experiment consists in collecting a huge amount of data for a given crystal in a couple of minutes (about 5 minutes today against several days 20 years ago). This type of experiments is an essential and well-established way to determine the 3D structure of proteins and other macromolecules. It is considered as one of the main pillars of structural biology. The users constitute the rather mature and well-organized scientific community of crystallography which has existed for decades. In this context X-Ray, operational since March 2008, delivers an intense, parallel and tunable X-ray beam for measurements at high resolution or from large unit cell dimension crystals.

The Infrared beamline we study is one of the two infrared beamlines at SOLEIL. The beamline is dedicated to microscopic analysis of a variety of samples, spanning from polymer films and multilayers, mineral inclusions, biological and biomedical studies, to archaeology. It is worth noting that the beamline at SOLEIL entails two branches (and two end stations), one of which is fully dedicated to biology and under the responsibility of a scientist of this discipline.

3.2 Empirical analysis

The method of analysis for this chapter consists on case study research and the use of qualitative data, more specifically interviews with key informers which in this case are users of the synchrotron as well as scientists and managers working at the synchrotron and involved in its operability. In addition to the interviews there is desk research which consists on the use of institutional and policy documents.

3.2.1 Why choosing a case study methodology?

The choice of research method is strongly determined by the nature of the research question and the amount of previous literature devoted to the topic. Here we ask “how” Large RI contribute to the production of a creative scientific outcome. We are interested in complex and dynamic phenomena. Additionally, the body of existing literature on the topic is scarce. As we have seen previously, research on scientific creativity is becoming popular but it is still emerging and therefore fragmented. When it comes to the study of Large RI impact on scientific creativity the literature is, to the best of our knowledge, inexistent. It is because of this, because we want to understand and because we want to know “how”, that we think that the appropriated methodology is a case study design (Wacheux, 1996; Yin, 2003). Case study research has been invoked as exploratory research. It is indeed a methodology that is appropriated to do research in topics that are not yet well known and there is very little to build up on. The use of case study methodology allows science to advance from the specific case towards the general one when entering a new research area. Although exploratory, the case study aims to fully understand a phenomenon and not solely to explore it. More specifically the typology of the case study is based on Eisenhardt (1989) and Yin (2003).

3.2.2 The data

Previous to the interviews there has been some desk research with the objective to better understand the European Research Infrastructure cartography as well as the functioning and history of synchrotrons and the disciplines that are linked to them. Table 3 collects the most relevant documents used.

Table 3 Consulted documents

What is a synchrotron?	Nature, volume 410, page 722 (05 April 2001)
Report on Roadmapping of Large Research Infrastructures	OECD 2008
A History of Molecular Biology	Nature Medicine volume5, page140 (1999)
100 years of X-ray crystallography.	Science progress, 2017 Mar 1;100(1):25-44
Biological applications of synchrotron radiation infrared spectromicroscopy	Biotechnology Advances Volume 30, Issue6, November–December 2012, Pages 1390-1404

The main source of information for this case study consists of interviews transcribed to verbatim. There is a total of 20 interviews that are used for this analysis. There are two kinds of interviews, eleven operators (from S1 to S11 in the table below) and nine users of the beamlines (from U1 to U9). We consider as operators all the people working at SOLEIL, at the synchrotron or beamline level. Users can be academic or industrial; the table below describes them more precisely. To avoid loss of information, each interview was recorded, transcribed in full verbatim, under the conditions of anonymity and confidentiality of information. Table 4 summarizes these interviews.

Table 4 Summary of interviews

Code of interview	Duration	Description
S1	3 hours	Soleil staff, management level
S2	1 hour	Soleil staff, beamline scientist
S3	1 hour	Soleil staff, beamline scientist
S4	1 hour	Soleil staff, management and beamline management
S5	1 hour	Soleil staff, general
S6	1 hour	Soleil staff, general
S7	3:30 hours	Soleil staff, beamline scientist
S8	1:30 hour	Soleil staff, beamline scientist
S9	1 hour	Soleil staff, beamline scientist
S10	1 hour	Soleil staff, beamline scientist
S11	2:30 hours	Soleil staff, beamline scientist
U1	2:30 hours	Researchers (French research laboratory): X-Ray
U2	2 hours	Researchers (European University): X-Ray
U3	2:45 hours	Researcher (French University): IR
U4	2:20 hours	Researcher (European hospital): IR
U5	1:30 hours	Industrial User (Pharma): X-Ray
U6	1:30 hours	Researchers (French research Institute): IR & X-Ray
U7	2 hours	Researchers (European research Institute): X-Ray
U8	1:45 hours	Researchers (European University): IR & X-Ray
U9	1:45 hours	Industrial User (Pharma): X-Ray

The content of the interviews (i.e. the topics discussed) is summarized in the following table.

Table 5 Interview frame

Soleil staff and beamline Scientists	Users of the synchrotron
Objective/ Goals	Modes of use, frequency
Recruiting Policy	Relevance for the research
Synchrotron research	Support from synchrotron scientists
Technical support	Collaboration with synchrotron scientists
Joint research with users	Collaboration with other users
Technological development	Communication with synchrotron scientists
Collaboration with other synchrotrons	Communication with other users

The following section describes how the data for the case study was analysed. All the interviews were transcribed to verbatim and therefore the data exist in a written form which allows for codification and classification into categories. This was conducted following three steps: the pre-analysis, the exploitation of the material and the treatment of the results (inference and interpretation).

3.2.3 Method of analysis

At the starting point of the analysis we have the two following research questions. How can synchrotrons favour creativity along the collective knowledge creation process? And, which mechanisms are behind this? To better answer these questions the previous analysis of literature suggested the most relevant factors that favour creativity, the underlying mechanisms, as well as the different kinds of creative scientific outcome. These are listed above and are the topics that we will be looking for in the interview analysis.

The analysis consisted in three phases. Because the interviews were semi-directed there were some general topics treated which resulted into more specific codes that went deeper into the subject and helped us understand the mechanisms that explain the different ways through which RIs favour creativity.

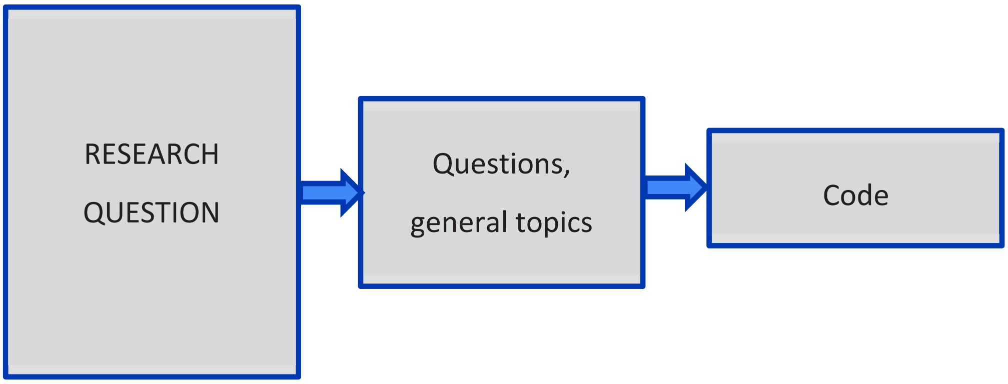


Figure 11 Coding process of interviews

3.3 Analysis and results

This section will describe the results from the interview analysis. It is organized by topics and in form of tables of contents. Each table will consist into 4 colons that will contain, in this order: the central concept suggested by the literature, the initial code that comes from the analysis of the interviews, the reference to the interviews associated to that code and finally a couple of verbatim excerpts from at least one of the interviews were the idea appears.

3.3.1 Favourable factors to creativity

3.3.1.1 Scientific autonomy and leadership

As explained earlier in this chapter, how research is managed has been found to be crucial for scientific creativity. Researchers should be able to decide on their own research agendas in order to be creative. Additionally, managers should be involved in R&D activities. Table 6 shows the results from the case study in this area.

Table 6 Scientific autonomy and leadership

Central concept	code	Interviews	Citation example
SCIENTIFIC LEADERSHIP	Managed by scientists that keep doing research	U2, U3, U4, U9	*« Ce ne sont pas que des techniciens, ce sont des scientifiques aussi qui, même en venant d'un autre domaine comprennent nos inquiétudes »* **(U9)**
		S1, S2	*« Oui on a en principe nos propres projets de recherche, c'est relativement difficile à faire parce qu'on est très sollicité par nos utilisateurs, bref on le fait, on est 3 pour faire tourner la ligne et on essaie toujours d'avoir un de nous trois qui est libre pour suivre dans le labo nos propres travaux »* **(S2)**
RESEARCH AUTONOMY	Research independence	S1,S2, S3, S7	*« Dans chaque ligne de lumière on a un responsable de ligne. J'en suis un. En tant que responsable de ligne j'ai ma propre autorité sur la recherche que je décide d'engager à condition évidemment malgré tout que j'ai une cohérence dans mon profil » **(S1)** « Les grandes thématiques sont généralement créées par des gens qui collaborent entre eux qui ont des spécialités complémentaires »* **(S7)**

One important factor for scientific creativity consists on **scientific autonomy** when deciding on the research agenda as well as **having leaders that are scientists themselves.** Research has shown that individuals that can decide on what to achieve and how to achieve are more likely to be creative. This is because being creative requires to be free to follow intuition instead of being restricted to some standards and rules. When it comes to science it is the same, research has shown that scientists who can decide on their own goals are more creative. It has shown as well that when managers are involved in research themselves the teams are more creative.

When it comes to the decisions concerning the beamline, beamline managers work in a framework of scientific independence where they choose their own priorities and research agenda. Indeed, beamline scientists are completely free to establish their own research agendas, they have complete freedom when deciding on their research strategy. For instance, at the IR, beamline scientists are continuously travelling to different research labs and institutions to promote the beamline and find interesting research that could endorse into collaboration projects with them. These, as well as the collaborations and joint projects are often based on the beamline scientist's intuition on what is interesting and what is promising. This research independence is, thus, connected to the previous factor, is because of their freedom that beamline scientists are able to promote and endorse in collaborations, joint projects and even long-term partnerships.

3.3.1.2 Scientific diversity

Next table will study scientific diversity and all the situations that bring this scientific diversity to the synchrotron. More specifically we focus in multidisciplinary research as well as variety of backgrounds of researchers (both users and SOLEIL staff) the end all the results will be interpreted and discussed.

Table 7 Scientific diversity at the synchrotron: Multiplicity of backgrounds & Multiplicity of scientific communities

General topic	Code	Interviews	illustrative extract of verbatim
VARIETY OF BACKGROUNDS	Multidisciplinary staff and management	S1, S3, S5, S9	*« Mais j'ai embauché en tant que scientifique non pas un physicien du synchrotron lorsque j'en ai eu besoin, j'ai pris quelqu'un qui ne connaissait rien du tout (à la physique) mais qui connaissait la bio spectroscopie et ça a changé tout. »* **(S3)**
	Multidisciplinary synchrotron users	S1, S3, S7, S9	*« […] on réunit des utilisateurs d'origines variées autour d'un instrument potentiel, de la mise au point d'un instrument potentiel... » **(S1)*** *« On a fait un diagnostic médical basé sur notre connaissance, sur les données que l'on a acquises au synchrotron […]. Pour moi c'est un exemple où la complémentarité entre le physicien, le médecin curieux. S'il n'avait pas été curieux ce médecin on ne serait pas arrivé à ça. »* **(S3)**
MULTIDISCIPLINARY RESEARCH	Share knowledge across disciplines	S1, S3, S5, S7,, S11 U3, U4, U8	*« La bonne approche c'est d'arriver à motiver et à convaincre des gens [de discuter] d'une façon interdisciplinaire, donc un biologiste ou un médecin avec parfois en intermédiaire un biochimiste, un bio spectroscopiste, ou voire un chimiste vont s'assoir et se mettre d'accord. »* **(S3)**
	Traditional physics methods are applied to biology problems	S1, S3, S7, U3, U6, U8	*« C'est particulièrement vrai en synchrotron puisqu'on a vraiment des méthodes physiques utilisées par des biologistes qui ne sont pas du tout physiciens, pas du tout spécialistes de méthodes... »* **(S7)**
	Joint projects across disciplines	S1, S3, S7, U4, U3,	*« Et donc des médecins se sont dit mais alors vous avez des outils d'observation de nos échantillons*

		U6	*biologiques qui dépassent nos connaissances qui est une connaissance pragmatique, par la coloration, l'observation, Qu'est-ce que vous pouvez nous en dire de plus. »* **(S3)** *« Je suis physicienne et je suis dans une équipe de neurologie, de neurobiologistes. Le domaine sur lequel on travaille dans cette équipe porte sur des outils ou des approches un peu innovantes, différentes pour répondre à des questions générales, très fondamentales ou très appliquées en biologie et en médecine dans le domaine des neurosciences. »* **(U4)**

From the relevant organizational conditions that are likely to lead to scientific creativity the most repeated one by the literature is diversity. Diversity of knowledge in all its possible ways has been found to be a key factor for scientific creativity. Because of the nature of the synchrotron we are specifically looking for variety in the form of multiplicity of backgrounds and multiplicity of scientific communities that participate in the synchrotron. The shape in which we have found these factors in the synchrotron consists on **variety of backgrounds** and **multidisciplinary research projects**.

The synchrotron is an instrument that can be used for multiple purposes and therefore is used by people coming from different communities and with a high diversity of backgrounds. Moreover, the technicians and scientists working at the synchrotron are also coming from different backgrounds. This does not happen only by accident; the synchrotron has a hiring policy based on this variety. When planning technological developments potential users from different backgrounds are invited to discuss and participate to the planification. Because of these two factors the result is a synchrotron user environment with a variety of backgrounds present.

Partly related to this variety of backgrounds we have as well the fact that there is multidisciplinary research taking place at the synchrotron. One representative example is the fact that traditional physics methods are applied to solve problems posed by clinical and biology research. Because of this, there are continuously research questions that touch different topics at the same time. As specified in one of the interviews (S3) the approach does not consist into, as a physician, go talk to a physicist and ask for technical advice. There is a continuous exchange of knowledge, both sides will nourish each other and eventually come up with the good research question and the good way to approach it. For this, communication is crucial, next section shows how effective communication takes place at the synchrotron.

3.3.1.3 The role of communities

The table below highlights the different communities meeting at the synchrotron, as well as their intersections.

General topic	Code	Interviews	Citation
COMMUNITIES	Participation in a community of synchrotron beamlines	S2, S5, S6, S7, U4, S8, S9, S11	*« […] on a choisi de travailler avec nos collègues dans les autres centres pour développer un logiciel de collecte qui facilite la tâche pour la bio-cristallo et qui est le même pour tous les synchrotrons, pour qu'un utilisateur, industriel ou autre vienne et voie la même chose même si l'équipement est différent »* **(S5)**
	Communities of user in different fields	S2, S3, S7, U3, U6, U9 , U7	*« Plus précisément en biologie à Soleil on peut distinguer la bio structurale type biochimie, vs plus bio cellulaire. […] Et toute cette partie bio structurale côté biochimie-biomol est bien structurée, et la culture de l'utilisation des sync est très bien établie. La partie bio cellulaire a besoin de bcp plus d'explications et de pub, parce qu'ils pensent qu'avoir un microscope dans son labo, c'est suffisant[…]»***(S7)***

	Research-technologists	S2, S3, S7, U3	*« Oui c'est la physico-chimie, on est d'origine des physiciens crystallographes, mais on a des compétences, on est tous passés par des projets de recherche dans la biologie structurale, bref on a des compétences des fois qui sont assez poussées dans certains domaines, mais on n'a pas un grand overview.»* **(S2)** *« Quand j'ai recruté un responsable de ligne, j'ai pris un biologiste - pas un physicien connaissant le synchrotron - parce que j'avais besoin de quelqu'un qui parle "biologie", sachant détecter le langage, poser les bonnes questions, comprendre les biologistes»* **(S2)**
	Crossing of communities	U9 U4	*« Oui, comme la communauté CCP4 - Collaborative Computational Project-, […] C'est une initiative anglaise, parce que c'est quand même en Angleterre qu'est née la crystallo, pour ce qui est informatique, avoir des outils informatiques pour tout ce qui est cristallo, synchrotron… […] Les plus grands cristallographes travaillent pour et avec CCP4, et développent des outils pour CCP4. […]»* **(U9)** *«J'ai toujours des biologistes qui sont sur un microscope, ils sont dans la salle d'à côté et j'ai des physiciens très souvent dans l'autre salle. Ils se parlent, ce qui est intéressant c'est qu'ils se parlent et j'ai vu*

			2 projets se monter par des discussions informelles sur des cellules de poumon en haute pression. .» **(S3)**

The results confirm the existence and crossing of various communities. One aspect for which the concept of community is central is the existence of a community of managers of synchrotron beamlines. This is a very active community at which researchers from different synchrotrons in the globe (not necessarily only Europe) communicate continuously and work together to develop new technologies of methods and improve the research that is done using synchrotrons. The existence of this community is likely to be a very positive factor for users' creativity. It is also a positive factor for quality, as they work together to continuously update the research equipment (see 3.3.3).

Moreover, the synchrotron staff caracteristics, namely their hybrid carriers and competences, their ability to create a common language with their users and of course, the fact that they work on a generic technology and have a lot of exchanges with other synchrotrons' scientists, all this make us think about Research-Technology i.e. an interstitial community. Knowledge sourcing from a given beamline towards its user community do exist, and vice versa. All this makes the synchrotron a highly innovative place, with many impacts on users' creativity, as will be confirmed in the section about the creative outputs. But this would not occur without the presence of the creative mechanisms described below.

3.3.2 The creative mechanisms

Literature suggests that one of the most important organizational mechanisms for scientific creativity is effective formal and informal communication. This type of connexion has been reported to be relevant within one same group and across groups. The reason why communication is important is, in first place that is necessary to understand the knowledge that is being transmitted. In other word it is necessary to have a common language, or to create one if it does not exist. Putting together a wide variety of knowledge as we have seen in previous section[38] is not useful unless actors can properly understand it. In second place, communication is important as it is sometimes the first

---

[38] This is also discussed in Chapter 1.

step towards collaborations and joint projects. We have observed how actors communicate at the synchrotron and the two following tables show the results of the empirical analysis in this area. More specifically the first table (

Table 8) shows the factors related to communications and transfers that are present in the synchrotron.

Let us consider the communication mechanism in more detail. Indeed, in several areas of the society, creativity has been shown to be found where the boundaries between groups are blurry or when there are bridges established to communicate between groups. It is not different in science; research is much more creative when exchanging information with other scientists. Boosting communication can be done simply by creating a climate propitious for knowledge exchange, in which people can discuss and exchange knowledge and ideas. From the interviews we have observed that at Soleil we can find several different situations in the spectrum of communication.

We first have simple **effective communication**, different individuals who communicate in the form of formal and informal discussions because they use a common space, or they share a common research interest. We have several common situations here, all present in several interviews. The first one consists in simple and quick exchanges that are done during breaks at the common areas of the facilities between different users; the friendly and easy-going atmosphere plays a relevant role here. In second place we have discussions between users and operators of the beamline, these ones are relevant because they develop trust and friendship, which has been specified by several interviewees as crucial for knowledge exchange. In both cases this communication can be done across disciplines, meaning that the different individuals involved in this communication do not necessarily share the same background. This is particularly true for the case of the Infrared beamline.

Eventually these informal discussions can lead to **collaborations**. The way communication sometimes turns into collaboration or partnership is documented in

Table 8 below. Indeed, another frequent case of effective connexion is **collaboration** **(**defined as intense and bilateral connexion). Collaborations are more common between users and operators, than between users. The former often begin with the unilateral provision of pure technical expertise, i.e. mono-directional transfer of knowledge from the beamline scientist to the users.

It can be noticed that there are important differences between the two beamlines at this point. An interesting result has emerged while comparing them. As highlighted in the table on collaborations below, the Xray beamline, which is the most mature, is engaged mainly in some kind of service provision and scarcely in effective collaboration with its users. By contrast the recent Infrared beamline tends to experience many more effective collaborations. Moreover the Infrared leader uses his personal beamtime to contribute to very promising projects of some users. As a consequence, the Infrared beamline scientists are most often associated to the authorship of their users' paper.

Interviews show several cases of collaboration and different levels of collaboration. The most common is to punctually help each other because the competences of two individuals or two groups are complementary: an advice to joint problem solving and even joint research projects, meaningful research questions that have been asked and solved as a result of joint thinking process. These collaborations can eventually turn into **long term collaborations** or partnerships, which is the last mechanism regarding connexion that was found in the interviews. When there are long run partnerships, non-SOLEIL researchers acquire the status of collaborators and they have access anytime to SOLEIL's facilities; it also means that they have a shared research agenda that is planned jointly.

It is relevant to note here that most of the collaborations, joint projects and partnerships would not have been possible without the investment of beamline scientists. All interviewees explained how, although the conditions for doing encounters are present, these encounters are fructuous and become something more than a simple chat thanks to the beamline scientists. Beamline scientists intentionally support and sustain an environment that is propitious for collaborations. They are as present as possible during the design and performance of the experiment and they build trust-based relationship with users.

Table 9 will focus on a bilateral and more intense type of connexion, which is collaboration.

Table 8 Effective communication at the synchrotron: formal and informal communication

General topic	Code	Interviews	Citation
EFFECTIVE COMMUNICATION	Formal and informal discussions	S1, S3, S4, S7, U3	*« […] personnes qui échangent, au travers de ces réunions, de ces séminaires, de ces participations à des conseils scientifiques et qui vont permettre de franchir des gaps de culture entre des communautés qui ont un tel gap de formation que ça constitue un gap de culture »* **(S1)** *« C'est vraiment, pour moi, la connexion est l'humain, la relation, la motivation entre différentes disciplines est quand même génératrice d'idées, de concepts. Il faut savoir se parler, ce n'est pas toujours facile. Bon moi je parle, j'essaie de parler avec des biologistes et parfois je ne comprends rien, il faut s'investir »* **(S3)**
	Friendly and easy going atmosphere	S3, S7, U4, U7, U9	*« c'est un aspect que l'on essaie de développer le plus possible ici parce que … la technique, on a de la très belle technique on a de la très belle technologie, c'est clair. Mais il y a un autre aspect qui est ce contact sur lequel, il faut qu'on entretienne nos relations, qu'on puisse leur permettre de mener leur expérience dans la plus grande technicité possible, dans la plus grande convivialité possible »* **(S3)**
	Knowing and trusting the staff	S7, U2, U4, U7, U9	*"I think in research you can have collaborations, but you can have also friendship. and I think you collaborate more with those people you trust. I trust Paul, and I will do work there. I think Paul*

			trusts me.” ***(U4)***
	Communication across disciplines and communities	S1, S2, S3, S5, U1, U4, U8	*« C'est-à-dire que, être au contact d'une autre culture, j'ai dit que la culture d'un physicien est différente de celle d'un biologiste et quand ces deux personnes se parlent, elles s'enrichissent toujours mutuellement. Dès qu'elles trouvent les éléments de jargon commun, l'enrichissement est immédiat. C'est comme quand on voyage dans un pays étranger, on découvre une autre culture, on s'enrichit forcément. »* **(S1)** *« Nous c'est vraiment les aspects structuraux de l'ARN. C'est clair on est dans une niche. Tous les gens ici sont dans des niches. On est quelques milliers dans le monde à faire ça. Même pas. On n'est pas nombreux. Par contre on parle aux autres niches. Les niches se nourrissent. »* ***(U1)***
	Informal discussions among users about what they do	S3, U2, U3, U4, U7	*“I’m trying to solve another problem using a different molecular replacement technique at the moment. And that was all suggested by somebody else when”* **(S2)** *« J’ai eu des collègues l’autre fois, ils ont scindé le faisceau en 2 sur IR. Souvent on discute la nuit entre les gens qui sont sur l’1 et sur l’autre »* **(U3)**

	Soleil staff provides with a crucial technical expertise	U1, U2, U3, U4, U5, U7, U8, U9	*« L'accompagnement scientifique au synchrotron c'est capital. Une ligne de lumière c'est un peu comme une voiture de sport, c'est comme la F1. On ne peut pas nous en tant que pilote on conduit notre Clio toute la semaine, on arrive dans la F1 le weekend, ce n'est pas possible. »* ***(U1)***

Table 9 Effective communication at the synchrotron: collaboration

General topic	Code	Interviews	Citation
COLLABORATION	More collaborations at infrared	S2, S3, S7, U3, U4, U6, U8,	*«C'est plus le cas sur des techniques émergentes [de prendre du temps perso pour tester des projets d'utilisateurs cf vis à vis], donc SMIS et DISCO que dans notre domaine qui est bcp plus balisé.»* ***(S7)*** « on a moins de 10% des publis auxquelles on est associé. Comme c'est plus "streamline", plus automatisé, on ne demande pas à être associé au papier.» *(S7)*
	Joint problem solving	S7, U2, U3, U4, U8, U9	*« On leur soumet nos difficultés de traitement. J'ai sans arrêts des mails et des discussions téléphoniques ou des réunions une fois que l'on a un peu dépouillé. On se revoit, on fait un bilan »* **(U3)**
	Spontaneous collaboration between users	S3, S2, U3	*« Donc en discutant on parle obligatoirement de ce que l'on connait. Je lui dis est que vous avez fait de l'AFM, non parce qu'on n'a pas l'AFM. Donc je lui dis si tu veux en faire pas de pb de venir chez nous. Là c'est une collaboration, c'était technique. Oui c'est faisable. Donc oui il y a des possibilités de collaborations. Mais en même temps on n'a pas vraiment le temps pour parler de ça. »* **(U3)**

	Joint research project between beamline scientist and user	S1, S2, U4, U6	*"I went there with this problem, they had this expertise, we discussed and we came up with an idea, you could call it a project, and we worked together on it"* **(U4)**
LONG TERM COLLABORATION	Long-term partnerships with other institutions	S1, S3, U6	*« Oui on a un très fort partenariat avec l'université X, le laboratoire de physique de Y, beaucoup de relations avec Z… Alors il y des partenariats où Soleil a essayé de contracter des associations de chercheurs et d'enseignants…Ce mixage est extrêmement essentiel. Il y a des ponts qui s'établissent »* ***(S3)***
	Joint projects and long-term collaborations with other institutions	S2, S3, S5, S6	*« Moi j'ai œuvré de mon côté pour avoir une collaboration forte, dans laquelle Soleil était visible, c'est développer un instrument qui nous permettrait un jour de démontrer que […]. »* ***(S3)*** *« On est associés à Soleil, on a des badges d'entrée, on fait du co-développement, on décide la stratégie ensembles »* **(S6)**

As we explained at the beginning of this chapter communities have a very important role when it comes to communication and knowledge exchange. During the fieldwork in this case study we have seen that communities are somehow permanently present. All the favourable factors to creativity have some relationship to the idea of community. When we talk about variety of backgrounds, we refer to people coming from different communities, when we discuss about the existence of collaborations these are often done between members of different communities.

One aspect for which the concept of community was central was the existence of a community of managers of synchrotron beamlines. This is a very active community at which researchers from different synchrotrons in the globe (not necessary only Europe) communicate continuously and work together to develop new technologies of methods and improve the research that is done using synchrotrons. The existence of this community is, if we follow what has been written previously by some scholars, a very positive factor for creativity. It is also a positive factor for quality, as they work together to continuously update the research equipment.

3.3.3 Synchrotron contribution to creativity: knowledge, technology and communities

We have, in the previous section, observed several conditions that allow us to think that the synchrotron offers the proper framework for scientific creativity. This section is devoted to the study of creative outcome at the synchrotron. As we have seen earlier in this chapter, creative outcome can take multiple forms, from the development of a new theory, to the discovery of a new empirical phenomenon and that without forgetting the development of new methodologies as well as opening new research paths. This section explores which of these kinds of creative outcome can be find at the synchrotron.

It is important to keep in mind, that the creative outcome is necessary related to the factors that are supportive for creativity and that, sometimes, it is difficult to establish the borders between what do we consider a favourable factor and what do we consider a creative result. This is illustrated later with examples.

Table 10 Synchrotron contribution to creativity: New knowledge and Technological development

General topic	Code	Interviews	Citation
New knowledge	The objective is to find new research topics	S7, U3, U8	*« Je considère que c ma mission, de trouver les bons sujets nouveaux […] C'est important d'anticiper les besoins, car les choses ne sont jamais éternelles. »* ***(S7)***
	Quality of scientific publications	S1, S2, U1, U5	*« On en fait de très bons papiers, de qualité. Dans les tops 10. Par exemple pour la 1ère fois de ma carrière je peux soumettre à Science ou Nature sans me dire que ça va être surfait »*

			(U1)
Technological development	Continuously looking for new applications of the technology	S1, S2, S3, S7, U3, U4	*« On a au début à la tête un physicien, qui monte cette ligne de lumière pour faire de la physique et traiter ses sujets à lui qui sont plus des sujets qui tournent autour de la chimie. Et voilà, il fait de l'infra-rouge, et puis il y a des gens d'ailleurs qui travaillent plutôt dans le domaine des sciences du vivant qui se disent « tiens l'infrarouge, je pourrais essayer pour résoudre mes problèmes » »* ***(S1)*** *« Nous quatre allons beaucoup dans les labos et les départements pour montrer ce qu'on peut faire avec SOLEIL aux gens. En général, le matin, on montre les potentialités de SOLEIL. Et l'après-midi avec ceux qui pensent avoir des thématiques qui s'y prêtes ou ont déjà des sujets mûrs, on essaye de voir si on peut monter un proposal, s'il faut faire des essais avant pour voir si c'est faisable… tout ce travail en amont. »* **(U6)**
	Technological development	S1, S2, S3, S4, S7, U4	*« Et donc pour faire cela, il faut un certain environnement, et donc on a développé avec eux un environnement pour un échantillon et on a la propriété intellectuelle partagée sur cette chose. »* ***(S4)*** *« Il y a une influence de nos utilisateurs pour créer les lignes de lumière, vous verrez le nombre de workshops qu'on organise pour définir les lignes de lumière, et pour être sûr qu'il faut la faire avec telles caractéristiques, y*

			a un travail de fond qui est fait par mes amis de la direction scientifique qui est quand même remarquable. » **(S1)**

The most representative kind of scientific outcome is the production of **new knowledge**. Science aims to produce knowledge and the synchrotron is a tool that is meant to advance in this direction. When analysing the interviews, we see this kind of scientific outcome several times. On one hand we have the continuous seek for new areas of research (new research topics, new research questions, etc) expressed by the operators (as well as some users) of SOLEIL, particularly for the beamline IR. New subjects and new areas of research are creative scientific outcome. This is strongly related to the factor for creativity: joint research projects on new topics. These projects are, somehow, a factor and a result itself as they have opened the way towards new research questions and new research topics. This production of new knowledge also comes from the publication of good scientific articles. Several users explain how their publications using the synchrotron are among their better ones. This quality of the publications, which is related to the quality of the technology at SOLEIL, are a representative example of creative outcome.

In second place we have the most relevant result find in the interviews, the continuous **technological development.** In a big number of interviews, we observe how both, users and operators of the synchrotron are continuously looking for new applications of the instrument. This is observed for both beamlines but the effect is a lot stronger in the case of the IR beamline. The technology is known to be able to respond to the needs of multiple disciplines and the operators of the synchrotron try to keep it that way as well as to expand the number of applications. This is done, we observe, by approaching directly the potential users and offer them to collaborate and jointly looking at what could be done together. The second face of technological development consists on, alone or jointly with other institutions and users continuously improve the instrument and the technology associated with it. There is a clear objective of being pioneers on the development of the technology and being always ahead of their time. Because a lot of its users are new, it is not yet a well-established technique and this technological development allows them to remain competitive.

3.3.4 Quality and impact

Table 11 Quality and Impact of synchrotron research

General topic	Code	Interviews	Citation example
QUALITY OF THE EQUIPMENT	The technology is cutting edge	S1, S2, U1, U2, U3, U4	*« Le rayonnement synchrotron est extrêmement brillant, il a bcp de photons qui passent avec des longueurs d'ondes bien plus intéressantes que les longueurs d'ondes que l'on peut avoir dans des labo »* ***(U1)*** *"If you do an experiment which is very much at the edge of what is possible, very technical, very challenging you want to have all these expertise"* **(U7)** *« C'est indispensable pour nous, la question est plus de savoir si on arrête l'équipement qu'on a ici. Se passer du synchrotron, ça veut dire passer beaucoup plus de temps à optimiser les cristaux. Maintenant, tous nos sujets ont besoin du synchrotron et le gain de qualité sur le synchrotron est sans commune mesure »* **(U9)**

QUALITY OF RESEARCH	Quality of accepted projects	U1, U2, U3, U5, U7, U9	*"Yes it's a premiere and it would be published in a high impact journal if somebody else doesn't publish it before which can always happen. We know we have competition in this project so you never quite know who would get first there. But yes, that will be published in a high impact journal"* **(U7)** *« Le problème c'est que les synchrotrons sont très demandés, donc il faut quand même s'inscrire à l'avance et avoir un bon projet pour avoir du temps de faisceau. »* **(U9)**

Creativity, as we have defined it in Chapter 1 and earlier in this chapter, depends strongly on the quality of the research, although it is not defined by it. For research to be creative it has to be valuable in addition to novel. We have found that most of the researchers interviewed have, among their best papers, the papers produced at SOLEIL. We were repeatedly told that because SOLEIL was often used for their most complex problems or their most novel ones, they often manage to publish SOLEIL's research in journals with a good impact factor. For some of the interviewees this was not only often but always. Systematically when using SOLEIL they knew they could submit their research to Nature or Science and be confident on their chances to be accepted.

Additionally, most of the interviewees consider that the quality of the equipment at SOLEIL is crucial. This quality of the facilities is a result of the numerous collaborations and joint projects for technological development that are developed at SOLEIL with collaboration of users.

3.3.4.1 An unexpected creative result: The creation of a new user community

During the first part of this chapter we have insisted on the idea that communities have often a central role in creativity and we have interrogated ourselves about the role of communities on the case of Large RIs and more particularly in the case of the synchrotron. But the communities are not only initial conditions or pre-existing organisational factors. An interesting and unexpected result of our field research is that they may as well constitute a creative output in itself. The next table summarizes our findings.

Table 12 The emergence of new communities of scientists

	Initial Code	Interviews	Citation
New communities	Existence of a new communities of users around Infrared beamline	S1, S3, S7, U3, U4,	*« Et maintenant on a une communauté de médecins qui travaillent avec des physiciens »* **(S1)** *« Oui on connait plus ou moins tous les gens qui utilisent cette technique dans le monde […] mais ce n'est pas une grosse communauté…elle a une 10aine d'années au grand max, c'est assez récent. »* ***(U3)***

At the beginning of this chapter we suggest that communities could have a relevant role in the construction of scientific creativity at the synchrotron. We already observed their presence when discussing the factors that contribute to scientific creativity. Indeed, behind all the collaborations between individuals with different backgrounds we find the capacity of communities to collaborate with one another, and to **create an entirely new community around the use of the beamline** by researchers who never went to synchrotron before. This is actually the case of the Infrared beamline.

Coming back to Literature suggests that one of the most important organizational mechanisms for scientific creativity is effective formal and informal communication. This type of connexion has been reported to be relevant within one same group and across groups. The reason why communication is important is, in first place that is necessary to understand the knowledge that is being transmitted. In other word it is necessary to have a common language, or to create one if it does not exist. Putting together a wide variety of knowledge as we have seen in previous section is not useful unless actors can properly understand it. In second place, communication is important as it is sometimes the first step towards collaborations and joint projects. We have observed how actors communicate at the synchrotron. More specifically Table 8 shows the factors related to communications and transfers that are present in the synchrotron.

In Table 12 we find several elements of what could constitute the development of a new community of scientists. There are cohesive social ties built on trust, continuous communication and joint projects. Additionally to this social aspect of communities we find common research topics and objectives. They face common scientific challenges. They also have in common that these pursued challenges are built around the use of one same technology. In summary, we have all the elements that constitute a community of scientists, namely: continuous communication and social ties, common topics and scientific challenges and common technology. All these elements could suggest the existence of a community that is emerging.

Hence an additional and relevant result of the case study is the emergence of new communities of scientists in the case of a recent beamline managed by a pioneer scientist. The point is all the more important since a new community is not an output ressembling any other one. Neither does it have the same nature and impact. Actually, a new community is an output belonging to a higher-level rationale: it is a dynamical concept by nature, a **process** that represents by itself a source of additional creative outputs (of any kind).

When it comes to identifying themselves as a community that has been built around the technology, this is evoked in some of the interviews. Some individuals refer to the existence of a community while others do not use that terminology but they explain the existence of common topics and cohesive social ties. What we can extract from this is that

SOLEIL synchrotron provides with the essential elements for a community of science to emerge. Despite all the positive signs, it is probably too soon to be able to say whether this new community will develop its own research topics as an autonomous sub-discipline with its own social interactions and norms, or whether the situation will stabilize at the current state.

All in all, these results confirm the relevance of our conceptual framework and some of our expectations. Most importantly, they show some light as refinements on the mechanisms hidden behind the factors that we identified in the literature review. Among them is the identification of a virtuous circle, i.e. a creative output that is in fact a new creative process that will in turn lead to more outputs, and so on and so forth...

4 Discussion of the results

In this section we summarize the results and offer a global view of them. This would show us how the synchrotron can impact the process of collective knowledge-creation and in which ways it makes it more creative. We also discuss the results by comparing them at the literature and we compare our findings to what we expected to find in first place. Which is the impact that a large instrument, more specifically the synchrotron, can have on scientific creativity? This is the question that opened this chapter. In order to answer it we divided this question into three more specific research questions, which favourable conditions for creativity does the synchrotron offer? Which are the mechanisms in action? and, which kind of creative scientific outcome do we find at the synchrotron? In this section we have answered to these questions.

4.1 Favourable factors and mechanisms for creativity at the synchrotron

We will here look back at the literature review, as well as the results, and compare the expectations that we had with out initial theoretical framework to the results that we found with the qualitative case study.

4.1.1 Beamline scientists

Literature suggests that it is important for scientists to have independence and for people making decisions at the management level to be scientists themselves. In the synchrotron, the top management consists on scientists. Additionally, the beamline managers are scientists themselves as well and not only technicians. This means, not only that they have a scientific background but also that they are actively performing science today. Beamline scientists have autonomy, this is, they can decide on their research agendas. This manifest in the ways of making decisions about partnerships, technological developments and their research agenda.

What we find in our interviews, which is not said by the literature, is that the rest of the factors for creativity: variety and synchrotron as a hub of communities, are strongly dependent on these scientists-managers and their autonomy. It is also the case for the mechanisms, they are facilitated not by the infrastructure conditions by themselves, but also by the human factor that is involved in them. It is the beamline scientists who actively search for collaborations, who decide to endorse into a policy of openness towards multiple disciplines and who facilitates the encounters between individuals. For the case of the Infrared beamline, we were told by the beamline scientist that he si aware of the benefits of putting a variety of people and communities together and therefore he does his best in order to achieve this.

4.1.2 Variety

Let us look at variety, which is suggested by all the literature on creativity and scientific creativity, as a key factor for creativity to occur. We expected the synchrotron, as a multidisciplinary facility, to offer variety in the form of different backgrounds with their methods and knowledge. This is indeed the case. The main way in which variety appears in the synchrotron are the existence of variety of backgrounds and the performance of multidisciplinary research. People from different background work in the same place and often work together and there is an exchange of heterogeneous knowledge. Because of this there is research that is performed with a multidisciplinary perspective putting together knowledge and tools from a variety of knowledge bases in order to solve a problem.

There is, however, a strong difference between the two beamlines studied. Although in both cases the interviewees recognize the existence and benefits of this variety, it happens at much less extend at the x-ray beamline than at the Infrared beamline. The reason is that the x-ray technology in the synchrotron is older and has very specific uses. There is a mature community using it, which is the crystallography community. Because is a well-stablished community and a well-stablished use of the technology there is lesser variety of backgrounds and profiles that meet.

4.1.3 The role of communities

If we have a look at the last of the favourable factors for creativity at the synchrotron which is the facility as a hub of communities, we encounter the same situation. At the synchrotron there is an encounter between multiple communities from multiple origins. There is the community of technologists to which the beamline leaders participate to, and there are, as well, the different communities of users depending on their discipline and research topic. Only because of the investment of the beamline managers into actively participating into this communities or actively communicating with them, that these communities are active at interacting with the synchrotron. This bring us to the mechanisms that these favourable factors are involved in.

4.2 The creative mechanisms

The creative mechanisms are those that allow the previously mentioned variety to interact in multiple ways. Here we discuss which ones we found at the synchrotron

4.2.1 Effective communication

For variety to be useful there is a need for effective communication. If communication is not enabled and even promoted the different knowledge bases will not meet and endorse into common research questions or at least exchange actively in order to feed each others' knowledge. The literature review let us to expect that, due to the multiple variety of backgrounds and disciplines that meet at the synchrotron should enable communication between the mentioned variety of backgrounds present.

We found this to be true. However the effect is even greater due to the continuous presence of beamline researchers getting involved in the activities that take place and facilitating the contact. Because users know and trust the beamline scientists (as he is always the same person, contrary to other users who they might meet only once) the communication is done mainly between beamline scientists and users. However often the users would ask the beamline scientists to put them in contact with other users.

4.2.2 From communication to partnerships

From communication arises collaboration. The exchange of knowledge may lead at some point to collaboration. This collaboration might go from a simple exchange of competences where two groups or two individuals work on separate parts of one same project, each using its own knowledge base independently; to the development of joint research projects. Where research questions and solutions are searched together, and the two knowledge bases are interacting on an ongoing manner. Once again, all cases the encounters among individuals that leads to these collaborations are not completely hazardous. Although some hazardous encounters are reported by the interviewees, most of them have been facilitated by the active investment of beamline managers into putting people in contact and facilitate their joint work. Additionally, an important part of these collaborations occurs between the synchrotron and the users.

Finally, these joint projects can become long term partnerships which means that two groups or two individuals continuously work together and when one research project is finished there is another that comes up and the collaboration continues. This has been shown by literature as a very positive factor for creativity. We find, at the interviews, that beamline scientists actively look for this kind of projects. It is part of their research agenda to assure that they exist and that they are the more and more important.

4.2.3 Difference between beamlines

An important remark for this section is that, although our results are consistent across interviews for all the central topics, they appear with different strengths depending on the beamline studied. While, for some individuals communication remains informal and even collaboration consists into a simple exchange of expertise, for other there communication is continuous and collaboration takes the form of joint research projects were the agenda is planned jointly. In both cases, communication and collaboration the strengths is higher in the case of Infrared than in the case of x-ray. The same happens with the presence of variety of backgrounds of users and operators, is present in both beamlines but much stronger in the case of infrared. Finally, for the cases of scientific independence, scientific management, quality of research and quality of the equipment we see that the strength is the same across interviews and for the two beamlines.

Next section will summarize and comment the results concerning the role of communities.

4.3 Creative outcomes

Table 13 Outcome

Topic	Central concept	Explanation
New knowledge	Production of creative scientific outcome	There is a continuous search for new research topics and original research questions. We can also observe the production of good scientific articles.
Technological development	Methodological and technological development	Synchrotron operators and users are continuously working together on new applications for the technology. They also work together for the continuous technological development that allows for experiments which are more precise and are done more efficiently.
Quality of research	Quality of research production	Most users express that the projects that they do using the synchrotron are among their best ones.
New communities	The emergence and fertilization of new communities of scientists	Several interviewees, both users and operators of the instrument refer to the existence of a community around the use of Infrared synchrotron technology for biology and medical sciences. In addition, we have several elements that support this idea

Table 13 shows the results when it comes to the question, which kind of creative scientific outcome can we find at the synchrotron? We first have the creation of new (creative) knowledge. Most interviewees express how, their better publications are those that are linked to the use of Soleil. By better they understand those who answer to the most complex problems and those which are published in the better journals. Many interviewees said that when they go to the Synchrotron to perform an experiment they can leave being pretty sure that this experiment will allow them to publish in a top ranked journal.

Additionally, to this we have the development of New technologies. This is shown by the interviews as the most relevant result. This is linked to the observed joint projects as positive factors for creativity. These projects are, at the same time, factors and results of creativity as they are at the same time a collaboration between different actors and a technological development which consists on the improvement of a technology and methodology.

4.3.1 Novelty and value

As we have seen earlier in this thesis for knowledge to be creative it needs to be new as well as valuable. What we can see some of the outcomes go more into one of these criteria and some go more into the other. The development of a new technology or method has a big part of novelty but a smaller part of value. When a technology or method is new is hard to value it, to evaluate well what can it provide to collective knowledge-creation process. However, its novelty comes with a lot of potential. On the other hand, outputs such as better ranked publications are more related to the idea of value. Impact factor and publication are well accepted rewards to scientific excellence. The access to them is, however, submitted to a certain number of rules and often these strict rules limit novelty and doing things differently, as we have seen in Chapter 1.

4.3.2 Unexpected result

We have, finally, one of the most relevant results and contributions of this chapter. Which is the creation of new communities of scientists. All the aspects described as factors that are favourable to creativity are, indeed, describing the existence of a community. We have a group of individuals that have common research interests, that share methods and technology and that communicate regularly. Some even refer as themselves as being part of a community. The relevance of this result lies on the fact that new communities come with new research questions, new methods and new approaches. It consists on the opening of completely new research tracks which is a way of pushing the knowledge frontier and advancing towards new challenges. On a way we are closer to explorative science and in rupture with the state of the art. This community was, however, at an early stage at the moment where the interviews took place (2013) and it would be very interesting to interview them again today and to see how the situation has evolved.

4.4 Summary of findings

To sum up, the results concerning favourable factors and mechanisms can be expressed as follows. Synchrotrons provide several favourable factors for creativity including: having scientists as managers, scientific independence of beamline scientists, existence of variety of knowledge and the existence of a hub of communities. However, all these factors depend on the strong involvement of beamline scientists and without their commitment none of the favourable factors would be as present and active as it is. When it comes to the mechanisms that the favourable factors are involved in, we find exactly the same situation. Effective exchange and connections through communication, collaboration and joint projects happens mainly because of the commitment of beamline scientists for this to happen. When it comes to the outcome of the process, it consists mainly into the development of new methods or technologies and the publication of good quality research.

The following figure shows how the collective knowledge-creation process is when organizations use synchrotrons. What we see is that, compared to a traditional knowledge creation process, here external knowledge appears at all the different levels of the combinatory process and not only very early at the level of literature review. This knowledge, in the case of the synchrotron is often technical and it concerns the methods of analysis. Additionally, due to the effective communication and the existence of collaborations, external knowledge feeds internal knowledge. Users do not simply draw from a pool of knowledge from other people’s knowledge, there is an active exchange. This exchange can even go to the point of creating a common pool of knowledge for the case of the new communities that emerge at the synchrotron.

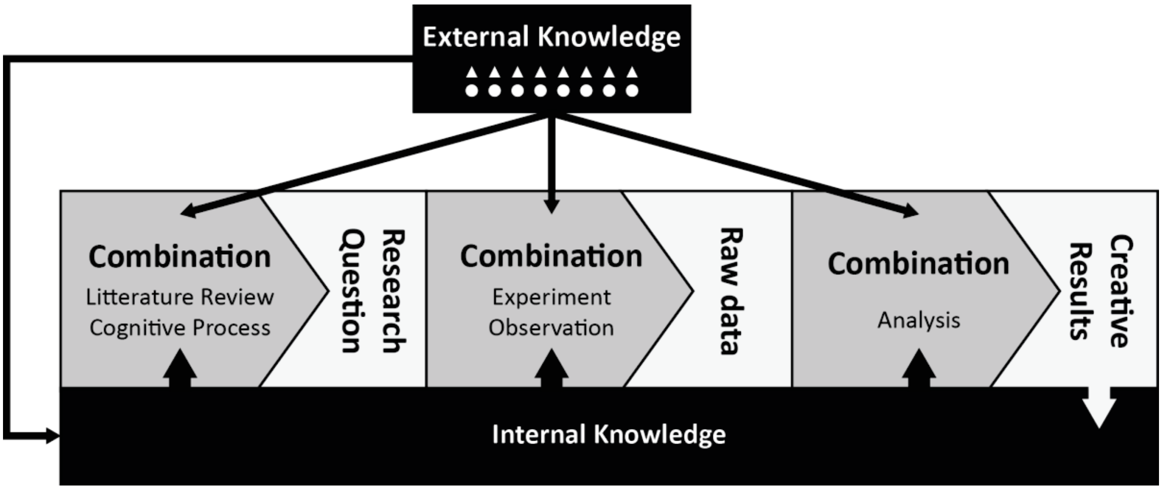


Figure 12 The collective knowledge-creation process at the synchrotron, results

4.4.1 Possible limits to creativity

In addition to the favourable factors for creativity that emerged from the interviews and the observation of creative outcome, there are also two factors that appear repeatedly and that could put creativity at risk. In first place, several interviewees, especially users of X-Ray comment on the fact that access must be planned a lot in advance and that the submissions of projects work in a different time scale that research. What this implies is that at some moment they might have some interesting samples without having booked time on the synchrotron. This lack of flexibility could eventually put problems to creativity. Users have, however, found a way to overcome this lack of creativity, they organize themselves in groups, called BAGs (Block allocation group proposal). They asked for time jointly and enough in advance to follow the times of the synchrotron and they decided later on how to share the allocated time.

The second potential limitation to creativity found in the interviews is the tendency towards automation of some experiments. Indeed, there are discussions about the possibility, in the near future, of doing the experiments remotely. If this happens all the aspects of creativity related to communication and spontaneous discussions would disappear. This is something that for now it is only being considered for the case of X-Ray beamline and it has not been established yet, but is something to keep in mind.

5 Conclusion

In this chapter we presented an original and relevant research question, can Research Infrastructure enhance creativity in science, and if yes, in which way? The originality lies as well on trying to answer to these questions for every stage of the creative process.

In order to answer to these questions, we decided to first focus our attention in one particular kind of research infrastructure which are big instruments. Because creativity is difficult to measure (as explained in Chapter 1 and better developed in Chapter 4) and also because it is a complex object, we decided to study creativity at the different stages of the knowledge-creation process. We first looked at the presence of organizational conditions that are favourable to creativity, then we studied the mechanisms that interact with these favourable factors and finally we had a look at the outcome of the creative process.

The results meet the expectations of our Theoretical Framework but they also surpass it. We find multidisciplinary. Synchrotrons are used with multiple research purposes and have users coming from multiple disciplines. These users will communicate with one another and exchange knowledge. This is one important factor than enables creativity. These discussions often turn into collaborations. Indeed, not only there is informal exchange of knowledge, there are as well joint projects and even long-term partnerships. What is particularly surprising is that all these exchanges happened only because of the commitment of the beamline managers and they will to create a place of exchange.

We find, however, big differences among the two beamlines studied with suggests that there might be a lot of creativity at the beginning and less creativity as the technology and the disciplines associated to it become mature. Indeed, we know that the Infrared beamline doesn’t have, yet, a well-established community of users and it is used in a multiplicity of manners and with multiple objectives. The X-ray beamline, on the other hand, offers an improvement of a technology that has been working for a few decades with the same community of users. Going deeper into the understanding of these differences could be the object of future research.

When it comes to the creative results, we have some hints of the existence of creative outcome. This comes mainly in 2 ways: traditional knowledge creation in the form of the discovery of a new phenomenon or biological or chemical material, or development of new methods and techniques. This aspect, however, is to be developed further. Chapter 4 approaches the possibility of treating scientific outcome in a quantitative manner.

Finally, we have the emergence of a new community of scientists. We see several conditions that make us think that a new community of science is emerging. Future research should focus on a follow up of this community as the interviews were done a few years ago and the situation has probably evolved. It would be very interesting to study the creation of a new community around the use of an instrument.

Chapter 3: Large Bio Medical Databases as drivers of creativity: An analysis of the case of the Pharmaceutical Industry.

1 Introduction

Science has always produced data. Since scientists started observing the world around them, they started taking notes on their observations and experiments where they would describe, measure and count. They have always as well-preserved samples and specimens. Even before the existence of science as we know it today humans would take notes on lunar cycles, write about harvest cycles or write about important events for the community. These collections of notes are data and they have been accumulated for centuries and part of them is still used today. With each technical improvement, the quality and accuracy of observations was improved as well. More recently (from the 1950's) several instruments have been developed which allowed the massive collection of data. With the development of these instruments the quality of data has grown together with the size and number. Some known examples are the current telescopes in astrophysics or the sequencers in genomics. This phenomenon is, however, present in a multitude of fields and we have mass data being produced in energy, environment or social sciences among others.

The growth on the size of this data has been exponential and recently its importance has led several domains of science to a complete change of paradigm as it has changed the ways scientists work and disciplines are conceived. Science has experienced, through history, some of these paradigm changes. First science was based on the observation and description of the natural world. Science was at that moment focused only on observable phenomena and it consisted mainly on what we know today as in-vivo science. A first paradigm change consisted on the development of theories based on abstraction and on the generalization and interpretation of observations. Another paradigm change came with the generalization of in-vitro science which consisted on experimentation. More recently, from the middle of the 20th century the use of computers allowed for the generalization of simulation of theories and a science based partly on prospection. Now with the cumulation of the results of those simulations together with the generalization of instruments that collect observations with a multitude of parameters associated science is generating very big amounts of data (André, 2014). These data have become central for the scientific activity. It has been called the fourth paradigm change and it consists on science which is based on the exploitation of these big amounts of data (Hey et al., 2009). The use and analysis of data has always been an important part on the production process.

Now this part of the process is becoming the more and more important and the amount of data being used as well.

The ability of the community to use such a big amount of data requires a collective effort in collecting this data. There is a need as well for an effort of public institutions to foster international collaboration and to make these big amounts of data available by means of RI, as no single laboratory could handle the storage capacity that is needed. Let us look at the case of biological databases. Today biological databases collect data produced by research in biology everywhere in the world. These databases are crucial in many fields of life science research, including medical research (Attwood et al., 2011). Biology is generating very large amounts of data. Managing this data has become a complex matter over time as it requires very large amounts of storage space and computing capacity. This is why there is an increasing need of publicly available databases that analyse, integrate and summarize the available data, providing an invaluable resource for the biological community (Bolser et al., 2012). The data that is being produced is integrated constantly and we can observe a continuous increase in its size and use (Gong et al., 2011). The main research areas concerned by these databases are genomics and its different subfields. The information contained in these databases includes gene function, protein structure, location and clinical effects of mutations among others. The discipline that exploits and manages these databases is bioinformatics, which is an interdisciplinary field. There are numerous databases offered by both commercial and public institutions. However, when we talk about Large databases there are only a few relevant players. Most researchers needing these kinds of data will use resources offered by one of the following publicly funded Research Infrastructures (RIs): NCBI, DDBJ and EMBL-EBI, which come from the United States, Japan and Europe respectively (Zou et al., 2015) In every case they are provided by public institutions, they are publicly available and open access.

The development of these biological databases has come with a growth on the use that medical research does of them. During the second half of the 20th century medical research had still an important part of randomized experiment and the few data used consisted of small in-house produced data silos. During the last decade the use of large datasets for health research and in particular in the case of pharmaceutical research has become the more and more extended. Data driven research has become highly important in the health care industry in general and few have looked at the impacts of this change.

The pharmaceutical industry has been relying the more and more on the use of these databases. Having data that are consistent, reliable, and well linked was one of the biggest challenges the pharmaceutical industry in the past, as it was for the bio medical research in general and there was a focus on the use of internally generated data. Now, with the increasing availability of high-quality databases, the pharmaceutical industry has incorporated genomics as a source of drug targets: data is becoming the foundation upon which the discovery happens. Databases and bioinformatics are today key aspect of drug discovery, contributing to both target discovery and target validation[39]. In this chapter we explore whether this change of paradigm has had an impact on scientific creativity.

The Pharmaceutical industry has faced a crisis of innovation for more than 30 years. Even though the investment in R&D activities has not stopped increasing, the number of new drugs that are approved every year didn't increase significantly. Data from the United States suggest that the number of new drugs approved per year has remained constant for the last 30 years[40]. Because the number of compounds registered has increased in a similar way as the increase in R&D expenditure some scholars have explained the innovation crisis with a possible focus of pharmaceutical companies on more complex problems. They argue that easy problems have been solved and that the current challenges of the sector are more complex and therefore scientific advance requires more time. Following the same line, another explanation pointed out is that large companies are investing more into fundamental science and pursuing long-term goals and challenges (Munos, 2009). If this is the case, the pharmaceutical industry is in need for some creativity in order to successfully face the current scientific challenges. One can expect that the move towards a data-based science is motivated by this need for creativity. This change could also be motivated simple by a change in paradigm in the scientific field in general. Regardless which situation we are facing, the use of databases could be a way, in the following years, to overcome the crisis of the pharmaceutical industry through an increase in creativity.

To sum up, large databases are growing as well as their relevance in science. Pharmaceutical industry is shifting the way they do research towards a data driven

---

[39] Drug discovery in the 21st century https://cdn.intechopen.com/pdfs-wm/25223.pdf

[40] http://www.phrma.org

research. These two factors could show a solution to the well acknowledge crisis of the pharmaceutical industry. It is for these reasons that this chapter aims to understand which is the impact that the use of large biological databases has had on scientific creativity for the pharmaceutical industry. To do so, this chapter is organized as follow. Section 2 builds up the theoretical framework that helps understanding which kinds of elements of the creative process we can expect to find for the case of databases. We recall the most relevant concepts studied in Chapter 1 and introduce new concepts that are particularly relevant for the case of databases. This allows us to identify the research gaps as well. Section 3 presents the methodology that we use in order to fill those research gaps and complete and concretize the theoretical framework. We perform a qualitative case study with users of the EBI database. Section 4 presents the results of the case study and discusses them. The chapter will end with section 5 which will consist on a conclusion of the findings and a discussion of the perspectives for future work.

2 Theoretical framework

As we have seen in Chapter 1, there is a growing body of literature aiming to understand scientific creativity. It is, however, still an understudied topic. When it comes to the specific case of databases and scientific creativity there is, to the best of our knowledge, no previous research that aims to understand that relationship. The objective of this section is to identify in the literature the concepts that are important for the understanding of the role of databases in scientific creativity. The concepts that we identify are used to bright up our understanding of the mechanisms, drivers and inputs potentially involved in the creative processes of science production. We focus particularly on those concepts that could apply to the case of biological databases and the use that the pharmaceutical industry does of them.

2.1 Findings on scientific creativity: the importance of variety

As we have seen in Chapter 1, knowledge production is the result of a combinatory process. Scientific discovery can be viewed as a form of human problem solving, a process which involves combination. Nelson and Winter (1982) said that “the creation of any sort of novelty in art, science or practical life – consists to a substantial extent of a recombination of conceptual and physical materials that were previously in existence." In other words, to create new knowledge scientists use existing knowledge pieces and combines them. Following the same idea Arthur Koestler (1964) talks about bisociation which consists on combining two frames of thought (with concepts, ideas and perspectives) in order to build new creative knowledge. We study, in this chapter, creativity as a process where several components are combined at different stages.

Creativity in science is defined as the creation of knowledge which is new and valuable. We have seen in Chapter 1 that these two criteria are crucial for creativity and therefore they should both be always present when studying creativity. Novelty is involved during the earliest parts of the process, where the combinatory dynamics happen. As we have seen in Chapter 1, we consider science production as a collective process with multiple actors involved. Literature shows us that scientific creativity is strongly influenced by communication. When communicating strongly within an organization and across the borders of the organization scientists exchange knowledge, ideas and viewpoints and this has a positive effect on creativity. The reason is that they introduce diversity in their combinatory process of science production and increase the chances of doing novel combinations. Following a similar logic, works on network brokerage argue that people who are placed at the intersection of heterogeneous social groups have an increased likelihood of drawing upon multiple knowledge sources, leading to the generation of new ideas. For example, managers who occupy brokerage positions are more often than others the source of good ideas (Burt, 2004; Rodan and Galunic, 2004). Diversity has been recognized as a driver of scientific creativity. Diversity comes in the form of different ideas and concepts and it can be applied to the kinds of knowledge that are put together during the process of science production. Researchers have shown how research that uses knowledge from a variety[41] of fields and involves scientists from a variety of backgrounds is more creative. Research organisations that allow for multidisciplinary research across departments, foster collaboration and promote mobility of researchers tend to be more creative (Heinze and Bauer, 2007; Hollingsworth, 2002; Zuckerman, 1967). Similarly, when teams include a variety of backgrounds or research is organized around problematics rather than disciplines, there is a tendency towards more creative

---

[41] Diversity and variety are terms used to refer to the same idea in the literature on scientific creativity. They also appear as synonyms in the Collins English Dictionary. https://www.collinsdictionary.com

outcomes (Heinze et al., 2009; Lee et al., 2015). Literature does not explain whether biological databases have an impact on variety but there are reasons to think that they do. Firstly, because of the massive growth that these databases are experiencing we can expect the variety of elements available is growing as well. Secondly, these biological databases are very large and include a variety of elements and parameters associated to those elements (gene function, proteins, results of experiments and simulations, 3d structures).

The previously explained combinatory process has as a result new knowledge. Following Heinze and Bauer (2007), we can understand the last part of the creative process as an outcome that consists always on new knowledge but can take different forms. Firstly, there is the formulation of new ideas or new sets of ideas that open a new cognitive frame, brings theoretical claims to a higher level of sophistication or challenge existing paradigms. An example of this kind of creative science is the Theory of specific relativity in physics by Einstein. In second place there is the discovery of a new empirical phenomenon that stimulates the building of new theory. A famous example of this kind of creativity would be how the observation of biodiversity led to the Theory of Evolution by Darwin. Thirdly there is the development of a new methodology. A new methodology, despite not being a scientific result by itself has the potential to solve theoretical problems that could not be empirically tested yet. Closely related to the previous one there is the invention of novel instruments that open up new research domains and new research questions that we could not imagine before. Finally, there is the new synthesis of formerly dispersed knowledge. It consists of putting together ideas and connecting phenomena that were considered separately before, and putting them together into one same cognitive frame. It is at this last part of the production process of science that the notion of value appears. This new knowledge that is created will be confronted to the evaluation by the peers when science is performed by scholars and aims to be published. When performed in private organizations this new knowledge will need to prove its value. For instance, a new promising compound in a pharmaceutical company will have to go through pre-clinical trials in order to prove it is useful.

What we observe is that the idea of diversity or variety is common to all the literature that focuses on the study of the determinants of scientific creativity. This idea is present as well in many bodies of literature that study creativity in general. In the next section we will study the concept knowledge distance which is strongly related to diversity of knowledge and its role on creativity.

2.2 Insights on knowledge distance and variety

As we have seen variety is crucial for the study of scientific creativity and the notion of knowledge distance very insightful and strongly related to variety. Science is done by recombining knowledge (Schilling and Green, 2011). The knowledge used when creating new knowledge can be more or less varied and this will have an impact on how creative the result is (Uzzi et al., 2013). There are two ways of accessing a variety of knowledge. The first one, as explained before, consists in communication within and outside the boundaries of an organization. It consists of having individuals with different backgrounds, experiences and competences discuss and work together. This is achieved by means of collaboration as well as by participating in conferences and seminars. Another, more traditional way to use external knowledge when performing science is to look at the knowledge that is in the public domain mostly in the form of books and scientific articles. In both cases knowledge distance will limit the degree of variety of knowledge that is used.

The access to a variety of knowledge when recombining it to build new one is not easy. Let us look to public knowledge in general. The publication of knowledge doesn't make it accessible for the society. This knowledge must be found (the researcher needs to know it exist) and understood. Knowing this knowledge exist is already difficult and even when found it must be scanned, interpreted and learned. When exchanging knowledge with other individuals we find the same issue. It is not always easy for individuals to communicate when they have very different frames of thought (i.e. different languages). The ability to identify and understand external knowledge is absorptive capacity (Cohen and Levinthal 1990). Absorptive capacity refers to the capacity of an organization to acquire and assimilate knowledge coming from other organizations or individuals. Absorptive has been described as depending greatly on the prior acquisition of related knowledge and the existence of a diversity of backgrounds in the organization.

The reason why absorptive capacity is needed is the existence of a distance between one’s own (and an organization’s own) knowledge and the knowledge that the individual wants to learn and use. Audretsch and Feldman, (1996) explain that the closest the knowledge explored is to one's own knowledge; the easiest it will be to face the challenges of finding it and learning it. This is why through the process of knowledge production, actors often tend to use only knowledge that they are familiar with. It consists of knowledge that is in their own domain of expertise or in very related domains. Absorptive capacity helps firms to go further than their domain of expertise when looking for external knowledge. To explain the importance of Absorptive Capacity Nooteboom, (2000) uses the concept of cognitive distance, defined as "a difference in cognitive function which can be a difference in domain, range, or mapping". The bigger the difference in “mental schemas” between two or more individuals, the greater is the cognitive distance that separates them. This distance between their knowledge bases has the potential to create very novel connections that are valuable. As seen before, major discoveries come very often from combining a variety of disciplines and backgrounds. Nooteboom explains how greater absorptive capacity allows for a greater cognitive distance when using external knowledge. This applies both, to the use of formalized knowledge such as publications, books and patents; and to the use of other’s knowledge through communication.

We have, therefore, the idea that excessively close actors might have little to exchange after a certain number of interactions and that in order to have creative results there is a need for combination of more distant knowledge bases. By staying in close domains at some point agents’ risk to start to always recombine the same kind of knowledge and it becomes redundant and less valuable and it leads to lock-in processes (Arthur, 1989; David, 1985). However, using a wider knowledge base effectively requires scientists to do a higher cognitive effort. This cognitive effort is based on understanding the new domain and the new elements of knowledge rather than on the recombination of elements. The capacity to cross the borders of one’s own discipline and explore unfamiliar elements is therefore limited (Q. Li et al., 2013). The more the knowledge distance between the new domain and one’s own domain, the more the exploration is limited.

We see, therefore, that distant knowledge is hard to use because finding it and understanding it takes more time due to the lack of familiarity with it. This distant knowledge is, however, crucial for creativity as the combination of distant pieces of knowledge is more likely to lead to novel results. To be able to use this distant knowledge, organizations will develop absorptive capacity. However even with absorptive capacity some knowledge will remain laborious to use and companies as well as research teams will have to make a choice between the use of distant knowledge (which requires large amounts of time and effort) or the use of closer knowledge (which is easier but is less likely to lead to creative results). Because of large databases include a very large amount as well as a large variety of knowledge one can expect that they would make distant knowledge more accessible. This chapter will explore the role (if some) that databases have on the access to distant knowledge. With our qualitative analysis we aim to study whether databases contribute to the access to a wider variety of knowledge through the ability to access distant knowledge. Moreover, we want to study the precise mechanisms that could allow this to happen.

2.3 Insights on the role of explicit knowledge on creativity

To do creative science, teams need to be able to exploit distant knowledge. This exploitation of distant knowledge is not easy and requires what we call absorptive capacity. Indeed, scientific teams need to build the ability to find and understand distant knowledge. It is not easy, and it comes often at a high cost. We think that databases can have an important role in building absorptive capacity, and we show here the literature that suggests that. To the best of my knowledge, existing literature does not approach the impact that the extended use of Large Databases has on scientific creativity. For this reason, we explore the findings on the role of explicit knowledge on creativity as databases can be understood as a kind of explicit knowledge. What we can observe in the literature on explicit knowledge and creativity is that there isn’t an agreement on which is its impact.

Explicit knowledge is often described as codified. Most explicit knowledge is technical or academic (which includes data) or information and uses a formal language (for example manuals, mathematical expressions, copyright and patents). It can be shared via written documents as well as by oral means such as conferences. Explicit knowledge is codified and articulated as it is the case for databases. Explicit knowledge is also described as the one we formally learn at school and university, it is knowledge which can be expressed in words, numbers or equations and it is easy to share (Koskinen et al., 2003). This ability of being shared is what could give explicit knowledge a special role in creativity. Indeed explicit knowledge facilitates learning and therefore the acquisition of new knowledge and competences that, when combined with one’s own knowledge, have the potential to increase creativity (Smith, 2001). One could thing that databases as well, because of some common aspects with explicit knowledge, can facilitate learning and therefore creativity.

However, explicit knowledge has often been described as the kind of knowledge that is produced and stocked at the end of the creative process and not as the type of knowledge that enables creativity. Explicit knowledge is related to organized tasks and routine. It assumes a predictable orchestrated environment. It is also aligned with a specific way of thinking which is logical, based on facts, that use proven methods and convergent thinking. This would mean that is not the kind of knowledge that enhances creativity as creativity requires divergent thinking and the use of uncommon concepts and ideas. Takeuchi and Nonaka, (1995) propose four basic patterns for creating knowledge in organizations. These are:

1. From tacit to tacit: learn by observing, imitating and practising, or become an expert into a specific way of doing things, such as learning from mentors and peers.
2. From explicit to explicit: combines separate pieces of explicit knowledge into a new one, like using numerous data sources to write a financial report.
3. From tacit to explicit: converting tacit knowledge into explicit knowledge means finding a way to express what used to be difficult to put into words. It involves stating one's vision of the world and coding it. It as an intellectual process and there is creativity involved. It is as well a way to transmit the knowledge.
4. From explicit to tacit: Re-frame or interpret explicit knowledge using one's own references. This is considered a creative process and it consists in finding ways to broaden, extend or re-frame a specific idea.

Within this framework explicit knowledge has an important role in learning and understanding which makes us think it could be crucial for creativity. For Takeuchi and Nonaka, (1995) as far as we are in a process of transforming explicit knowledge into tacit knowledge or tacit knowledge into explicit one there might be creativity. We propose, however, that the use of explicit knowledge is, by itself, relevant for creativity. This framework shows that explicit knowledge is a key tool for sharing and learning. Because knowledge is built up into previously acquired knowledge making knowledge explicit allows for a better understanding of this knowledge and therefore for being able to use it.

We can therefore imagine that databases, as explicit knowledge, are easy to use and understand and therefore combination is made easier. If we look at the literature on knowledge distance, we see that understanding knowledge with which we are not familiar, can be an important challenge for firms and in this chapter, we think that databases could ease this challenge.

2.4 Research gap

Existing literature on scientific creativity throughs some light into the understanding of which the role of biological databases on scientific creativity may be. It leaves, as well, some open questions and highlights some gaps in the literature that we aim to fulfil in this chapter.

Novelty is crucial for creativity, in order to find novelty, scientists need to draw from a variety of knowledge ideas and concepts, if possible distant ones. We will investigate whether the databases can offer that.

We see, indeed, how variety in the knowledge that is combined is important for scientific creativity. We have some reasons to imagine that the growth of the use of biological databases has brought a growth in the variety of knowledge that is used. Biological databases integrated several kinds of compounds as well as a multitude of parameters associated to these compounds. One can easily imagine that this brings a variety of knowledge that is higher than the one companies had when working exclusively with in-house produced data. We don't know, however, the extent of this effect and how can it precisely affect novelty in the combination process. It is for this reason that we will investigate in first place, **do they have an effect on the variety of knowledge combined when performing science? How??**

We see, in second place, how it is important to use knowledge which is distant to an organization's own knowledge. In order to use distant knowledge, organizations need absorptive capacity. Databases could have a positive effect on absorptive capacity, and therefore on creativity, by facilitating learning. Indeed, the literature shows how explicit knowledge (a category that databases fit into) can facilitate learning and hence creativity. For this reason, we ask: **can they facilitate access to distant knowledge? How?** This would, as well, play a role on the ability to use a variety of knowledge explained before (which would include distant knowledge).

Finally, as creativity has a dimension of value as well, we will investigate which is the role of databases on providing value, more specifically, **how do databases offer value to its users?**

In this chapter we will explore whether this happens, and which mechanisms are involved in this part of the process.

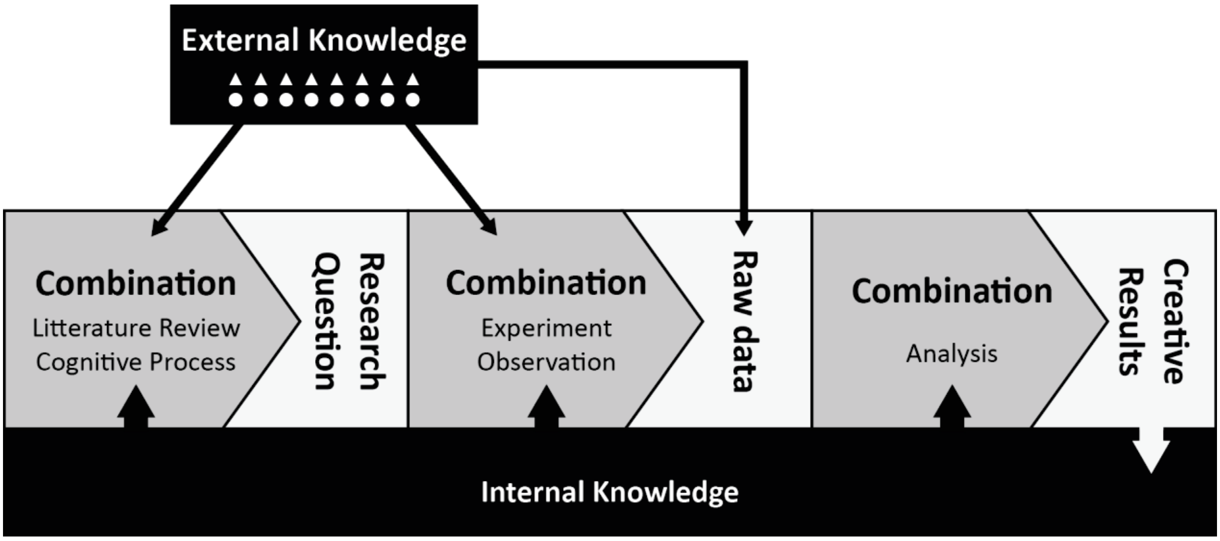


Figure 13:Collective knowledge-creation process with data

The figure above shows the collective knowledge-creation process and the possible impact that the use of databases can have of it.

As we saw in Chapter 1, the process of knowledge creation is characterized on one side by its collective dimension. Indeed, science is the more and more done in group, often by means of collaborations and always with interactions with the rest of the individuals of the organization and with the external world. The figure shows as this dimension and it shows as well how a logic of recombination of different pieces (of knowledge, ideas, concepts, methods) is present all along the process. It is a combinatory process of concepts and ideas that allows the research question to emerge. It is as well a combinatory process that allows for the generation of data. It combines techniques and methods for observation and experimentation, as well as some formal and tacit knowledge, in order to generate data. This data is later analysed and interpreted again through a combinatory process and ends up generating results. These results feed the process as they become the organization's internal knowledge. They might consist of formal knowledge as well as of methods or techniques.

The organization as well as the external world feeds the process creation with knowledge. Traditionally we considered that external knowledge consisted only in formalized knowledge such as written papers. There are, however, different kinds of knowledge and data are among them. Indeed, as explain early in this chapter data are the more and more complex, with an increase in parameters and detail. We think that databases introduce variety into the external knowledge that is used during the collective knowledge-production process.

In order to understand whether this is the case and which exact mechanisms allow databases to introduce variety we perform a qualitative case study at a EBI, a European bioinformatics Institute which is one of the biggest providers of biological data in the world.

3 Presentation of the case study, EBI databases and methodological approach

To empirically explore the impact that databases have on scientific creativity and the suitability of our theoretical framework, we consider the case of the European Bioinformatics Institute (EBI). EBI is one of the main providers of public biological databases. As explained in Chapter 1, we use a qualitative case study methodology with semi directed interviews with key actors. Case study methodology is recommended when looking at processes and trying to understand and explain complex phenomena (Yin 1994, Eisenhardt 1989 and 2007). It is also recommended when doing exploration of an understudied topic. Although we are not doing pure exploration and we do not use grounded theory, our approach nevertheless includes an exploratory dimension. Below we will see the details on the history of EBI and the nature of the databases that it provides, the details on the profile of the interviewees and the reason they were chosen and finally the strategy to be able through interviews to answer our research question.

3.1 Description of the EBI databases

The databases chosen for this research are the ones offered by EMBL-EBI, the European Bioinformatics Institute. EBI is a research institute but also a provider of databases containing biological compounds such as proteins or genes information. This database is the largest bioinformatics resource provider in Europe and is used all around the world. The access to these databases is completely unrestricted, free and without need of registering. The users of these resources are both public and private researchers. Our research focuses on private medical research, in particular we focus on the case of Pharmaceutical companies.

EBI origins lie in the first Nucleotide Sequence Data base that was established in 1980 at EMBL in Heildelberg, Germany. The initial goal was to create a central database of DNA sequences submitted to academic journals. It began with very modest aspiration of simply abstracting information from literature but soon it started receiving data directly. Universities and laboratories all over the word upload the results of their experiments and observations directly into the EBI databases. This reduces the job to EBI to that of verification. It poses however some challenges as the amount of data being inserted is growing at high speed. In addition, the magnitude of the database grew in scale when the Human Genome Project finished in 2003 and all the produced data had to be integrated. It was an international scientific research project with the goal of determining the sequence of all the three billion nucleotide base pairs that make up human DNA, and of identifying and mapping more than 100.000 genes of the human genome from both a physical and a functional standpoint. It is still today the world's largest collaborative biological project. EMBL played an important role on this project and when EBI database was established it integrated rapidly all the results from the Human Genome Project. This gave EBI more visibility and therefore more use and more popularity. Additionally to the accumulation of data produced by other institutions, and offering them, EBI employs scientists and produces research of its own, exploiting the database as well as producing data and introducing them.

EMBL-EBI started with two databases, one on nucleotide sequences and another one for protein structure but with time it has diversified, and it provides now resources in all the major molecular domains. It gives access to freely available data from life science experiments, performs basic research in computational biology and offers an extensive user training programme, supporting researchers in academia and industry. The services entail not only data archiving but also data curation and integration. They allow users to query EBI's large biological databases programmatically, eventually to build data analysis pipelines or to integrate public data with users' own applications. The six core data resources are operated by relatively large teams of 15 to 20 people (scientific curators, software engineers, bioinformaticians, and visitors including PhD students). [42]

3.2 Research Strategy

In order to have an answer to our research questions, the users from the pharmaceutical sector were interrogated directly about the use of these databases. The specific target was the heads of Bioinformatics departments in big Pharmaceutical companies. Why this specific sector and why only big companies? As we have seen there has been an innovation crisis in the pharmaceutical Industry and it is interesting to ask whether the change on ways of working within this sector is likely to bring some creativity and, in the end,, innovation. The chosen companies belong to what is commonly known as "big pharma" which are very large pharmaceutical companies. The reason to focus on Big Pharma and not small companies is because it is precisely within big pharma that there has been an innovation crisis. The pharmaceutical market is dualized, with a few very large companies and a big number of start-ups. Within start-ups innovation is the key to their survival, and we can expect, therefore, that creativity will be present. For Large companies, however, where the strategy lies more on the exploitation of well-established products, being creative is more challenging (Orsenigo and Malerba, 2015). Finally, an important reason for the choice is that the Pharmaceutical industry is concentrated; most of the market is controlled by a reduced number of firms and it is therefore relevant to focus particularly on them.

The first step, before the design of the questions and topics to be discussed during the interviews, was to learn about the context. The aim of this was to better understand the past and present use of data in the pharmaceutical companies as well as which kind of

---

[42] fhttp://www.ebi.ac.uk/

data are used and how are they used. This was done by means of desk research, mainly scientific literature discussing advances on methods and resources and European reports. The kinds of documents studied during this part of the processes are some policy reports concerning the provision and use of databases, OCDE and European Union reports on science, as well as the annual reports of the providers of those databases. There was as well a process of literature review with articles published in some of the mainstream scientific journals that discussed scientific policy. Table 14 shows a list of the most relevant documents used during this part of the process.

Table 14 Desk research: Institutional reports and scientific journals

Desk research
Institutional reports
Data Availability Policy Final report of Public consultation on Science 2.0, Open Science, European Commission. *Sharing data from large-scale biological research projects: a system of tripartite responsibility Excellent Science in the Digital Age,* European Commission.
Scientific journalism and news
Computational Biologists: The Next Pharma Scientists? Science magazine. Michael Price April 13, 2012 *An Explosion Of Bioinformatics Careers*. Science magazine. Alaina G. Levine June 13, 2014 *A decade's perspective on DNA sequencing technology*. Nature Elaine R. Mardis 2011 470: 198-203 *Science after the sequence*. Nature News. Declan Butler reports June 2010

3.3 Choice of informers

The users studied participate in a Partnership Programme with EBI[43]. This partnership was created by EBI in order to keep informed of industrial users' needs. The participation in the partnership is open to any company that is ready to pay the participation fees. The participation on this Partnership and its impact on creativity are not studied in this Chapter and for anonymity reasons we cannot disclose the names and companies of the

---

[43] This Partnership is The Industry Programme. It was used only as a way to identify industrial users from the Pharmaceutical Industry. The creativity that is the result from this partnership is not studied. The members of the Industry Programme may change from one year to another. https://www.ebi.ac.uk/industry

interviewees. The reason to chose them was, firstly, that most of them come from what is commonly known ad “big pharma” and this kind of companies are the focus of our research. Secondly, the existence of this partnerships allowed us to meet them informally a few times during their quarterly meetings and build some trust. This trust was crucial as the pharmaceutical sector has traditionally been very closed and secretive.

The interviewees are all managers of bioinformatics departments of their companies. This position gives them a privileged view as they are the intermediaries between the researchers who use the databases daily and the databases themselves. This means they are in daily contact with both the users’ needs and the data and technical possibilities. The second reason for choosing these individuals is that they constitute a homogeneous profile and this can give us an in-depth view on the sector. Since the methodology used is a case study methodology, it is appropriated to focus on a specific profile of user and seek for an in-depth knowledge of it.

3.4 Interviews

The interviews were semi directed (see topics treated in the following section). There was a set of topics, but no specific questions were included in order to avoid influencing the answers of the interviewees. They were encouraged to talk freely about their vision on the effects of these databases on creativity. In total, 9 companies have been studied. All the interviewees are at the head of the bioinformatics departments of their respective companies. There were two waves of interviews as well as informal discussions during a 2-day meeting organized by EBI in between the two waves of interviews. The format was face to face interview (with one exception of a Skype interview). The reason why there where two waves of interviews is that, during the first wave, companies were interrogated mainly about the value dimension of creativity. Some questions concerning the novelty part of creativity were done already but they were mostly exploratory. Creativity consists on both, value and novelty. Although novelty often comes earlier in the production process of science, value is a dimension that is more often studied and therefore it is easier to identify as there are more widely accepted tools and parameters. This dimension is also easier to observe because it is present very often at the outcome level. The first wave of interviews allowed us to understand the use that is done of these databases, how do they intervene in the production level and at which points can we find the notions of variety and distance. These two notions are the ones that are related to novelty. The second wave of interviews was more focused on these dimensions of variety of knowledge distance and on trying to identify novelty.

The interviews lasted between 30 and 120 minutes. After analysing them a second informal encounter was set to discuss some missing and incomplete information. This encounter happened with the entire group during one of the Industry Programme meetings. Afterwards a second wave of interviews was set, with three of the already interviewed people. In this case interviews lasted longer, from 78 to 115 minutes. All this information is summarized in Table 15. All interviews were recorded and transcribed verbatim, under the conditions of anonymity and confidentiality of information. Anonymity conditions here implies not only not disclosing the name of the people and companies involved.

Table 15 Overview of interviews

Company pseudonym	Informal discussion	1st wave	2nd wave
Ph1	Yes	90 min	80 min
Ph2	Yes	90 min	115 min
Ph3	Yes	60 min	78 min
Ph4	Yes	65 min	No
Ph5	Yes	60 min	No
Ph6	No	100 min	No
Ph7	No	30 min	No
Ph8	No	50 min	No
Ph9	Yes	No	No

The first wave of interviews was conducted between March and April 2015. Most of them took place at EBI during an Industry Programme meeting. All the interviews started with the general information on the company (size, employees, size of R&D department). The subsequent topics focussed on the main goal of the interview: characterising at the level of each company the role of bioinformatic databases in general and of EBI's ones in particular. More precisely our questions concerned the kinds of databases they used other than EBI (for instance, some private databases or public databases from other providers). We asked as well about the way they used them (how many people in the company were concerned by its use, frequency of access, etc.). We also asked about the importance of these databases (i.e. dependence for R&D activities). The same questions were asked about bioinformatic tools related to the databases. Then the final part of the interview concerned the evolution of the role of these databases during the last decade, in order to understand the changes, they possibly brought to the R&D processes of the companies, and/or their contributions to the efficiency of R&D activities. Table 16 summarizes all this information. In order to ensure that there had not been any misunderstanding and that what we had understood what we had been told during the interviews, we contacted the interviewees and asked for confirmation of agreements with our interview notes. This interaction was done via exchange of e-mails and was completed by December 2015.

Table 16 Content of first wave of interviews

Question	Issues explored
Presentation of the company.	General numbers (Number of employees, number of employees in R&D, number of employees in bioinformatics, annual revenue, R&D budget and bioinformatics budget
Which is the intensity and relevance of use of EBI and other databases?	Which databases are used, how often are they used, which is the level of integration between in-house databases and EBI databases and degree of dependence
Which is the intensity and relevance of use of bioinformatics tools in general?	Which kinds of related tools are needed, how often are they used, which is their role in the production process
Which have been the changes and improvements in the way research is done and the results you have?	How bioinformatics in general and EBI in particular have changed the way they work and the research possibilities, comparison to how science was made before and quantity of resources they have access
How did databases bring those improvements?	Exploration of several topics

After completing this first wave of interviews, during December of 2015 some informal conversations took place. We went to a workshop of the Industry Programme to discuss with the participants about our conclusions and go a bit more in depths on some topics that needed further discussion.

Table 17 Content of informal discussions

Topics discussed
Importance and dependence of today's pharmaceutical research on Databases
Evolution of the use, current challenges and future possibilities
The importance of spaces such as the Industry Programme for the community

The second wave of interviews consisted in the confirmation of the knowledge learned in the first wave of interviews as well as achieving a more detailed and articulated vision on how databases can facilitate the access to a higher variety of knowledge as well as how this variety impacts creativity. As we can see in Table 17, we asked specific questions which were answered with extensive detailed explanations. Firstly, we asked interviewees about sources of novelty and types of distant knowledge. To go deeper into these questions' interviewees were asked whether this meant an access to knowledge coming from other communities, other disciplines and other methods. After these questions, we assessed the connecting property of databases. How do Large Biological Databases connect pieces of knowledge that remained unconnected before? Finally, creativity and kinds of creativity were discussed, as well as the possible risks that the dependence on databases might bring.

Table 18 Topics treated in 2nd wave of interviews

Question	Issues explored
Which is your perception of the creativity in R&D?	Kinds of creative process: introduction of new methodologies, new scientific paths or new fields of research
Do databases provide access to a higher variety of knowledge?	Databases provide access to a higher variety of pieces of knowledge that are introduced in the research. Including knowledge from other disciplines and communities.
Do they have an impact on the amount of disciplines and methods used?	The large amount and variety offered by the database requires from new hiring policy. Multidisciplinary teams and variety of methods.
Do databases connect distant knowledge bases?	Databases and bioinformatics facilitate the connection of pieces of knowledge that were difficult to connect otherwise
Thoughts on risks	Lock in effects, fashions and trends.

4 Analysis and results

As explained before, the interviews have been written down as verbatim and analysed. Because we part from an abduction logic, the methodology used for the analysis is to ask questions to explore the topics identified by the literature review (variety, knowledge distance, novelty and value) and go from the general to the more specific in order to understand how databases can contribute to these different mechanisms that have traditionally been associated to creativity. For the parts of the interviews that could not be associated to one of our large themes we attributed it new codes, that emerged fully from the fieldwork.

4.1 Analysis of transcript (first wave of interviews)

The first wave of interviews shows how EBI databases offer several improvements to companies when performing science. These improvements appear in the form of access to a higher quantity and quality of data which results in speeding their production process. Databases have also changed completely the way they perform science, which is more data driven than it used to be. When asking about the origin of these changes the idea of access to an integrated set of data which are curated and do not need much treatment is highlighted. The idea of variety emerges as well by itself when interviewees discuss how the existence of some tools related to the databases allows them to access more varied knowledge.

Table 19: Analysis of transcript (first wave of interviews)

Topic	Code and explanation	Interviews	illustrative extract of verbatim
DEPENDENCY ON PUBLIC DATABASES IN GENERAL	**Weak dependency on EBI** (EBI as the preferred database but not crucial as long as there is access to other public resources)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9	*"I do like EBI but if it was down for some reason for an extended period I could easily use NCBI or DDBJ"* **Ph2**
	Strong dependency (Some interviewees considered public databases as crucial and claim most of the research activity depends on the use of those databases)	Ph2, Ph4, Ph5, Ph6,	*"The way we do science today requires public databases. Is not a matter of price, private providers could never offer the quality of resources such as EBI one's [..] The advancement of science requires these services "* **Ph2**
	Weak dependency (Some interviewees considered public databases as important but not crucial)	Ph1, Ph7, Ph8, Ph9	*"We did not have large databases in the past and we performed science. Of course we like to be able to use EBI databases but if they didn't exist we would do science in other ways"* **Ph3**
	Growing dependency (All interviewees considered public databases dependency in their companies growing)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9	*"We definitely use them the more and more and our research is the more and more depending on them"* **Ph2** *"Yes, of course, every year there are more projects that depend on them*

			and they are bigger" **Ph8**
IMPROVEMENTS	**Quantity** (Large amount of data contained in the databases)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"The quantity of data we can use now could never be offered by any private provider"* **Ph3** *"The quantity of data we allow us to do research that we would not have thought it was possible to do 20 years ago"* **Ph8**
	Quality (Quality provides first of all a feeling of reliability as well as complete information which allows users to save time on vitrification and completing data and use it on treating them instead. Quality has been also described as how the data are presented, the interface and the updates.)	Ph1, Ph2, Ph3, Ph4, , Ph6, Ph7,	*"The quality of data is very important for us. It has fast updates, right entries, structure... I haven't seen any private source that equals this"* **Ph1** *"Good quality of the data allows us to do what we do without spending countless hours to make sense of what the data actually are telling us"* **Ph6**
	Speed (Some interviewees insisted on speeding the research process as the most relevant impact that the use of EBI databases has had on their companies. Firms do not longer need to invest efforts on searching, verifying and building databases.)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9	*"The speed of target identification and validation has increased because efforts can be dedicated to exploit the databases instead of building them"* **Ph2**
CHANGES	**How science is done** (Something mentioned by an important number of interviewees and given a big emphasis is that their way of working depends on having access to that kind of database)	Ph1, Ph2, Ph3, Ph4, Ph7, Ph8, Ph9,	*"We would not be able to do what we do without the data that is provided by the EBI. The efficiency of the data is a key enabler. Being able to download the data in its raw format they can work with it straight away. There are many cases that the efficiency of the data is really important to us, up to the point that we wouldn't do a project without the data being standardized and available."* **Ph3**

ORIGIN OF CHANGES	**Access to variety** (Each of the interviewees agreed on saying that with EBI they have access to a wider variety of data that they would without. The idea of variety emerges on its own when studying this topic)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"Whenever we ask scientific labs for public data they have published in the 90s, in most of the cases they do not give us the data. Without the central repositories we would not have access to a lot of the data."* **Ph3**
	Aggregation (Data produced by science, even when public, are dispersed in terms of form and location and it would be very costly and almost impossible for firms to collect all these data and integrate them.)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"Getting this data from the different dispersed primary sources would be a nightmare and typically we wouldn't do it. Officially the data is out and has been published but most of the time we would not be able to find this information, or it would be too costly."* **Ph3**
	Curation (The data that EBI offers have been curated which means that they have been verified and there is no or little doubt of their authenticity. This allows firms to use them without having to engage into a validation process to ensure the data are reliable.)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"Data are very reliable, we couldn't have such a reliability by buying the data to private suppliers. The curation job done by EBI and the users is very valuable"* **Ph2**
SOURCES OF NOVELTY AND VARIETY.	**Standards** (EBI has standards in the way the different compounds are expressed, described and named. All interviewees agreed on saying that is the key to making all those already public data. without those standards it would be impossible to compare data coming from different sources and sometimes even to find those data)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"Standards are the key to what we do today. Scientist would rather share a toothbrush than gene names. This made it very difficult before to know, even with published papers, which gene they were talking about so the data used was limited."* **Ph2** *"Standard are fundamental foundations which are now used so widely that it is difficult to imagine not having them."* **Ph1**
	Ontologies (This is the second factor considered as crucial, and it is related to the previously mentioned Standards. It is a tool offered by EBI that allows for a query	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"It allows us to access harmonized data and at the same time keep our own version without having completely changed gene terms used internally which are often useful because it is related to the use the company does of them"* **Ph4**

	done using any of the possible given terminologies to one unique element. This allows scientists to search data related to a specific compound even when they do not know the standard nomenclature.)		
	Variety (Interviewees highlight the possibility of using other data than the ones that are internally produced. They insist on the fact that even if using private providers of external data, it would never be as varied as it is with EBI ones.)	Ph1, Ph2, Ph3, Ph4, Ph5, Ph6, Ph7, Ph8, Ph9,	*"Without EBI or any other public database doing what EBI does we would go back to the use of internally produced data"* **Ph3** *"We would buy private data but it could never be as rich and complete as the data from the EBI of course (...) Here we have everything"* **Ph4**

This first wave of interviews consisted on the exploration of five general topics. This exploration had, as a result, the emergence of the 13 important concepts that we observe in the Table. These five general topics are: dependency on EBI and public databases, the improvements that the use of these databases has allowed for, the changes in the way that science is performed, the origins of those changes and how these databases can be a source of variety and novelty in science.

Dependency on EBI and other public databases

When asking about the dependency of the research activity on EBI all the interviewees said that, although it is their preferred source of data, they would easily continue their normal research activity by using the databases offered by some of the alternative public providers of biological databases. When asked about the dependency on public databases in general (and not only EBI ones) we can observe a polarization in the answers. On side we have those interviewees that consider that the advancement of science requires the existence of public databases. The most common argument that we find here is that no private provider could offer the quantity and accuracy of the data offered by EBI and the other big public database providers. If all of these databases were to stop existing, the advancement of science would suffer greatly. Other interviewees, however, claim that although without large databases science could not be done as it is done today, their company could continue doing research. This research would simply be done in other ways. They most common argument is that it is only recently that they use these databases and not long time ago they could easily do without them, with in-house databases or small databases provided by private companies. Finally, all of them think that the dependency on public databases is growing and they have become the more and more important in their companies for the last years.

Improvements offered by the use of large EBI databases

The second topic of discussion consisted on the improvements offered by the use of public databases and more particularly, EBI ones. There is rather uniformity concerning the answers of the different people that were interrogated. The improvements that EBI concern the quality and quantity of data that is available to do research. This has as a result an increase on the speeding of the research processes. With EBI the data they can access is considered to have better quality than the data offered by private companies but also more reliable than the data that used to be produced in-house. This increase in quality comes in the form of accuracy of the data, number of parameters and also reliability (it is less likely to find mistakes). Concerning the quantity, the amount of compounds that are available is simply very big and cannot be compared to those offered by private providers or to the databases that are built inside the company. These two characteristics have allowed companies to save up some time and speed up the production process.

Changes on the way science is done

Most of the interviewees agreed that the way they perform science today is determined by the existence of these databases. It has changed the way that problems are posed and solutions are changed. The projects that are proposed today depend on the existence of large biological databases and without them those research projects could not be performed. The kinds of research questions asked would simply be different without the access to these kinds of databases.

Origin of changes

The changes on the way that science is performed come, in first place, from the quantity and quality of the available data. We asked which other reasons have made databases become as important as they are today in the production process of science. Three main factors were highlighted: access to variety, aggregation and curation. The access to a variety of compounds emerges in the interviews by its own before asking specifically about it. Most interviewees explain how EBI databases allow them to identify and understand a wide variety of data (i.e. variety of compounds and variety of parameters available). They claim that to have this effective access is not enough by having the information available and open. They need this information to be efficiently aggregated and curated (which means that the accuracy is verified and that the parameters are correctly written). EBI databases offer this.

Sources of novelty and variety

Because variety is important and because the topic emerged easily and naturally in most of the interviews, we asked what kind of variety do these databases offer and more precisely which factors allow for this variety. Concerning the kind of variety available the answers were rather vague and they simply talked about data that comes from other organizations than their own and data that are different than the ones they usually produced. This vagueness came also from our lack of understanding of the subject and motivated the second wave of interviews. What we found when exploring this topic is that what allows the access to this variety of data are the standards and ontology tools that EBI offers together with their databases. EBI's databases use some standards to express the multiple information and parameters of the data. Thanks to this standards the data can easily be understood, even when they are very different from the data one is used to produce. Ontologies are language tools and their role in access to variety follows the same logic. One of the traditional problems that scientist faced in the area of biology to use external knowledge is the names of compounds, particularly proteins and genes. Because there are hundreds of thousands of proteins and genes there was, up to now, no a common way to name them. Because of this, when looking for information on a protein, for example, scientists needed to know the name used by different groups of scientists to name that same protein. Ontologies are language tools that allow overcoming that limitation and have been defined by several interviewees as the most relevant contribution of EBI databases.

4.2 Analysis of transcripts (2nd wave of interviews)

During the first wave of interviews the topic of variety of knowledge could not be studied in-depth. We found, during the interviews, that variety was important. We could not, however, concretize that knowledge and understand exactly in which way knowledge accessible with EBI databases is more varied. This was due to a lack of knowledge in the area of biology and bioinformatics. This is why, in order to better understand this, we did a second round of interviews with a focus on creativity in general, types of variety and the possibility to connect distant knowledge. We can see the results in the following table.

Table 20 Analysis of transcripts (2nd wave of interviews)

Topic	Code and explanation	Interviews	illustrative extract of verbatim
CREATIVITY IN GENERAL	**Creativity** (Each of the people interviewed responded affirmatively to the question "Do you think that the use of Biological databases such as EBI favours creativity?" Something else highlighted is that EBI allows them to work on understanding disease, rather than simply trying to cure it. Understanding disease however is complicated and requires creative thinking.)	Ph1, Ph2, Ph3,	*"With bioinformatics we need to understand the disease, but we are free to decide how do we do to understand the disease and because disease is complicated to understand you need to be creative. The more creative you are the better research proposals you will do and the better chances to understand disease. Better you understand it, better you can intervene."* **Ph3**
TYPES OF VARIETY	**Varied Team** (The backgrounds of people working together when these databases are used are heterogeneous with different academic backgrounds expertise. The use of these Large Biological databases has forced firms to have teams of people with a variety of backgrounds.)	Ph1, Ph2, Ph3,	*"Mixed background. 15 people 10 to 12 different backgrounds> physics, biologist, chemist, biochemist, physicians, computer scientists"* **Ph3** *"Only if you bring all these very different people together you can reach all the different topics we need to work on because you need tech geeks, IT geeks, methodological geeks, people that understand the disease to ask the right questions to the data"* **Ph3**
	Variety of methods (Here interviewees confirm using methods that come from different communities and different disciplines thanks to the need of treating heterogeneous data. Again, as it happens with the team, the use of a multiplicity of methods is required to be able to exploit the data)	Ph1, Ph2, Ph3,	*"The process is not standardized at all; every day is different. Not routine. We do re use methods but Development of new methods come very fast. New data so need for new tools and new analysis times. For instance, new sequencing technology needs new ways to analyse the data"* **Ph1**

	Varied origin of data (Here interviewees confirm using data coming from multiple origins. More specifically they highlight the use of data that comes from different communities and different disciplines thanks to the existence of databases such as EBI. The existence of EBI has led to the introduction of bioinformatics as an essential discipline in pharmaceutical research and it is a very open discipline.)	Ph1, Ph2, Ph3,	*“Bioinformatics is a very open community and pharma companies have become more open. Bioinformatics share more methods, tools, data, etc. Bioinformatics was one of the first disciplines within pharma that does that, and it has spread. The other departments do the same now.”* **Ph1** *“Communities are getting closer to each other. More and more communities use bioinformatics and this makes them be closer to each other. More communities do bioinformatics. Pharma companies now always do bioinformatics before not all of them”* *“Cancer data to answer questions about cardiology because the data is there. But there is also cross talk, cross fertilization on methods”* **Ph3**
	Varied typology of data (This was a topic all interviewees agreed on and considered it as one of the major impacts of the use of Bioinformatics and the availability of databases. It included mechanisms such as connecting distant knowledge bases and combining knowledge that otherwise we would not be able to combine)	Ph1, Ph2, Ph3,	*“Yes, data are more diverse. And it is becoming more and more diverse. It started being about sequencing data but now there is also, epigenetic data, methabolomics data, data on bla bla and also the sequencing data is becoming more diverse.”* **Ph3** *“Yes, it does (provide access to a wider variety of pieces of knowledge) and this variety is important. There are a lot of mechanisms in a cell, proteins are just a part of it but the larger the varieties of data on the different biological compounds, the better processes are understood.”* **Ph1**
POSSIBILITY TO MAKE CONNECTIONS	**Connecting data (**To the question of whether this access to variety also meant that data that were distant could be connected all interviewees answered affirmatively.)	Ph1, Ph2, Ph3,	“*Complementing protein structure databases with other databases and connecting them all will tell more information*” **Ph2**

KINDS OF CREATIVE OTPUT	**More promising compounds** (The variety of data together with the possibility to connect it brings more promising compunds)	Ph1, Ph2, Ph3,	*"The amount of data and the expertise of our teams allows us to use sophisticated algorithms that"* **Ph3**
	New research questions: (Another aspect was that all the interviewees agree is the fact that the questions that can be asked now are questions we can only think about because we have the data. One of the most important creative results of the existence of these databases is precisely the possibility to ask these new research questions)	Ph1, Ph2, Ph3,	*"The sole existence of the data sets brings questions that we would not know that could be asked or that we never asked to a specific set of data before. Some of the new research Questions are different in its nature, they are research questions that we would not know we could ask. The existence and availability of data allows as well for the development of new methods to answer the questions that we have always had".* **Ph2**
EMERGING UNEXPECTED TOPICS	**Serendipity:** (This concept is described as a happy accident or how, given a big enough number of people working on a subject the possibilities of someone, even by accident, making the right connection or a very original connection are bigger. This is a topic that was brought up by the first interviewee himself and confirmed later on)	Ph1, Ph2, Ph3	*"There is a democratization of data that enables hundreds of thousands of researchers in different areas, wealth categories, settings to do something interesting. When you through a billion bioinformaticians into a dataset instead of one you have serendipity, somebody will come up with a finding or an innovation."* **Ph1**

During the second wave of interviews, the interviewees were asked more specifically about creativity in general as well, the variety of data they can access, and which are the creative outcomes they consider to have gained thanks to the use of publicly available databases. There were four questions that were asked from which a total of nine topics emerged. The questions asked consisted on: the impact of the use of databases on creativity, the types of variety that are allowed by the use of the databases, the possibility to connect distant knowledge and the kinds of creative outcome. There was as well the emergence of an unexpected topic: serendipity. The reason to ask first for creativity in general and later on go further into the detail on the variety of knowledge that is used, the use of distant knowledge and the creative outcome is the difference between which is our theoretical conception of creativity and what do the interviewees understand as creativity.

The impact of the use of databases on creativity

When asked about creativity, the three people that were interviewed answered directly by asking what we consider as creativity, our definition was challenged but we always found some common ground for discussion. Most interviewees considered that EBI database help them to be more creative. Because of the existence of such a large amount of data, research is not randomized anymore and there is a need to understand the theory behind the disease. Then they have to design algorithms that will explore the data and hopefully find some promising compounds. This way of working forces them to think in different ways which they consider to be creative. They also consider that the existence of teams with mixed backgrounds (for instance data scientists and biologist) forces them to put together different kinds of knowledge and different kinds of expertise and therefore it makes them more creative.

The types of variety

As seen during the first round of interviews, this second round confirms the fact that EBI databases provide access to a wider variety of data and that this has an impact on creativity. More specifically this variety is expressed in terms of variety in the origin of data (particularly data coming from a wider variety of disciplines, and from multiple disciplines) and typology. Additionally, this variety of data also imposes a varied background in the team and a multiplicity of methods in the firm to process them.

We were explained how; traditionally researchers in biology and pharma used only data coming from their close circles. This means for example data coming from the same community of researchers (i.e. community of cancer research, community of heart disease research, community of Alzheimer disease research, etc.). This was, as explained during the first wave of interviews, because the way they named things forbids them to identify the research that concerned some specific genes or proteins. Thanks to the existence of ontologies, researchers can use data coming from a variety of communities and origins. Indeed, not only the research community would determine how the data are named, the country or the laboratory of origin of the data had an impact as well. EBI allows as well to access to a wider variety in terms of typology of data (i.e. protein function, protein expression, protein sequence, protein/gene interaction, etc.). This is allowed thanks to the existence of standards which permits researchers to understand easily the data and its parameters.

Additionally, this variety present in the databases themselves forces companies to have variety in their research team. Because of the multiple kinds of data and the complexity of the problems that this data can be used for, research teams require a variety of backgrounds. They need to have computer scientists, data scientists, biologists, experts on different kinds of compounds, etc. This variety of teams working together will bring a variety of methods, tools and expertise that will all be put together in order to solve complex problems.

The possibility to connect different knowledge

Interviewees explained how, these databases have allowed them not only to access a varied typology of data, but also to be able to make connections among these data. This means, for example, to put together protein structure data and protein function data and use algorithms that will explore both of them together.

The kinds of creative outcome

We asked the three interviewees which kind of creative outcome comes from the use of EBI databases. It is important to notice that here we are not yet talking about drugs; the outcome from the research process is at a much earlier stage than the drug development. Indeed, large biomedical databases play a role in the very early stage of research and the creative results consist in promising compounds, new research questions that are not on the continuity of previous research, opening new scientific paths and the development of new methodologies. When using EBI the compounds found to be promising are more likely to be validated. This means that the promising compounds are more likely to continue to the stage of pre-clinical trials. A surprising result consisted in considering that not only they found promising compounds. Thanks to the use of EBI databases researchers can open new research paths and ask questions that they would not have thought about without the existence of the data.

Unexpected emerging topic: serendipity.

Finally, an unexpected result which emerged from the interview was the role that these databases have in terms serendipity or happy accident. Indeed, we found that the existence of a critical mass of people facing a critical mass of data is crucial for someone coming up with the right idea. This is something that came up in the first of the interviews performed and that was later confirmed, although with less enthusiasm, by the other two interviewees. Interviewees explain that, while facing the “full picture” with an “open mind” it is more likely to have “happy accidents”. They all explained how the unexpected findings became more common as the use of databases spread. Asked for examples on this kind of effect the most common ones consisted on doing an unexpected finding while looking into a completely different research question. In one example, while trying to find a method to detect the DNA of a future child in a pregnant woman’s blood, researchers found a method to detect cancer in people’s DNA. It is, however, not the only type of serendipity. In another case, after months of trying to figure out how to identify the genes associated to a specific kind of cancer they found the solution in an unexpected way. Several algorithms were being used to target the genes that were thought likely to be associated with that certain kind of cancer. Simply by observing at some data and confronting it with their own knowledge, that group of researchers understood that they had to modify the algorithm they were using. That modification in the algorithm let them to find the genes associated to cancer.

4.3 Discussion of results

In this section we will summarize the results and put them together in order to build answers to the questions that were asked at the beginning of the chapter. This chapter aims to understand the driving mechanisms for scientific creativity that are facilitated by the use of databases. More specifically we wanted to answer to three main questions:

- How do databases offer value to its users?
- Can databases have a positive effect on novelty?
    - Do they have an effect on the variety of knowledge combined when performing science? How?
    - Can they facilitate access to distant knowledge? How?
- Which kind of creative outcome comes from the use of large public databases?

Finally, **we find an unexpected result which is that databases allow for serendipity**, as it exposes a big amount of information to the entire scientific community and that cannot be fitted in the figure due to its hazardous nature.

We develop all these aspects in the following lines to better answer to the research questions that were asked before.

- **How do databases offer value to its users?**

The literature review did not give any specific clues on what the impact of the use of large databases in value could be. Value is an important notion of creativity and we usually put it at the end, as an outcome. We did indeed expected that, if found, value would be present in the form of a marketable drug. However, we found the notion of value already at the early stages of the process. Databases provide value in terms of good quality resources. We see that large public databases offer some value to companies that use them because they are changing the way they do science and they are the more and more dependent on large databases. Without them science could not be done in the way that is done today. The quality and quantity of the data is the reason for this choice. No private company could offer such an amount of data and such quality in terms of accuracy and completeness. Having databases has allowed them to speed the production process.

- **Can databases have a positive effect on novelty? Do they have an effect on the variety of knowledge combined when performing science? How?**

Literature suggested that, because of the large amount of data present we did expect novelty particularly in the form of more varied knowledge available. We find what we expected but at a much larger scale. These databases collect (almost) all the information available for (almost) all the existing data in the field of biology. There are several factors that play a role on the possibility to use such a large variety of data. The effective accessibility to this large amount of good quality data is allowed by its aggregation and curation. In other words, the fact that all the information is found together in the same platform and by using the same interface is a great facilitator of access. Additionally, the accuracy and validity of this data is continuously verified and does not create problems to the companies that use them. Indeed, we were told that the reliability of the data allowed scientists to trust the data blandly and do not limit themselves to the areas of knowledge they are familiar with in order to assure the validity of the data This variety of data comes in the form of variety in its origin (mostly data produced by other communities) and on its typology (kinds of compounds and parameters associated)

- **Can databases have a positive effect on novelty? Can they facilitate access to distant knowledge? How?**

This question, strongly related to the precedent one, asks whether the big variety of data available today is also more distant. Because the formalization of knowledge that it comes with the transformation of knowledge into data, we expected this to happen. Literature suggested indeed that databases would allow access to more distant knowledge. The interviews confirmed this and once again what we did not expect was the extent of this effect. The idea of distant knowledge concerning biological compounds has virtually disappeared. Because all this knowledge is put together into one single database and using the same codes to navigate through them all these data are easy to use by companies in their production process. The feature that is considered the more relevant from these databases by all the people interrogated is that they provide standards and ontologies, which are tools that allow for the understanding of the information related to each biological compound. Ontologies are knowledge tools and consist into translation but also putting some uniformity on the multiplicity of ways these compounds are named and its parameters expressed.

- **Which kind of creative outcome comes from the use of large public databases?**

Because we are at the very early stages of the drug production process the kind of outcome that we are facing is not, yet, a commercial one. Databases help, in first place, to find more promising compounds. Because of the big amounts of information concerning genes, proteins and other compounds, scientists can run algorithms that allow them to better predict which of these compounds are more likely to become a drug. Interviewees told us that, when using EBI databases, the targets identified (potentially drugable compounds) are more likely to be validated and go into the pre-clinical trials. A second surprising kind of creative outcome consists on being able to ask new research questions and open new research paths. This last aspect comes from the fact of thinking about a possible question, a possible project, after being inspired by the data

Unexpected result

An unexpected result is the fact that databases facilitate serendipity. Serendipity is often defined as a “happy accident” and it consists on discoveries that have a big part of chance in explaining them. Thanks to these databases scientists seem to be more likely to come with hazardous discoveries. Indeed, creativity, although depending on chance, can be induced by some favourable factors such an open mind of researchers or the accumulation of knowledge that allows scientists to identify what is odd and unexpected. In a similar logic, databases offer a complete view of the available data to the whole scientific community. This means that, on one side, scientists have a large number of elements to combine and a lucky discovery is more likely to happen. At the same time, the number of scientists potentially doing this is big which increases even more the chances of an hazardous discovery.

Our findings can be summarized as follows. Companies are the more and more dependent on large databases and without them science could not be done in the way that is done today. The quality and quantity of the data is the reason for this and additionally it allows for speeding the production process of science. The effective accessibility to this large amount of good quality data is allowed by its aggregation and curation. In other words, the fact that all the information is found together and that its accuracy and validity is continuously verified and does not have to be verified by the companies. The feature that is considered the more relevant from these databases is that they provide standards and ontologies, which are tools that allow the access to a wider variety of data. This variety of data comes in the form of variety in its origin (mostly data produced by other communities) and on its typology (kinds of compounds and parameters associated). Additionally to this variety, databases allow scientists to make new connections thanks again to the existence of standards and the aggregation of data. Concerning the kinds of creative outcome that one can find, it consists on one side of promising compounds and on the other side of the opening of new research questions and new research paths. This last aspect comes from the fact of thinking about a possible question, a possible project, after being inspired by the data. This aspect is related to our last finding.

The following figure illustrates these results:

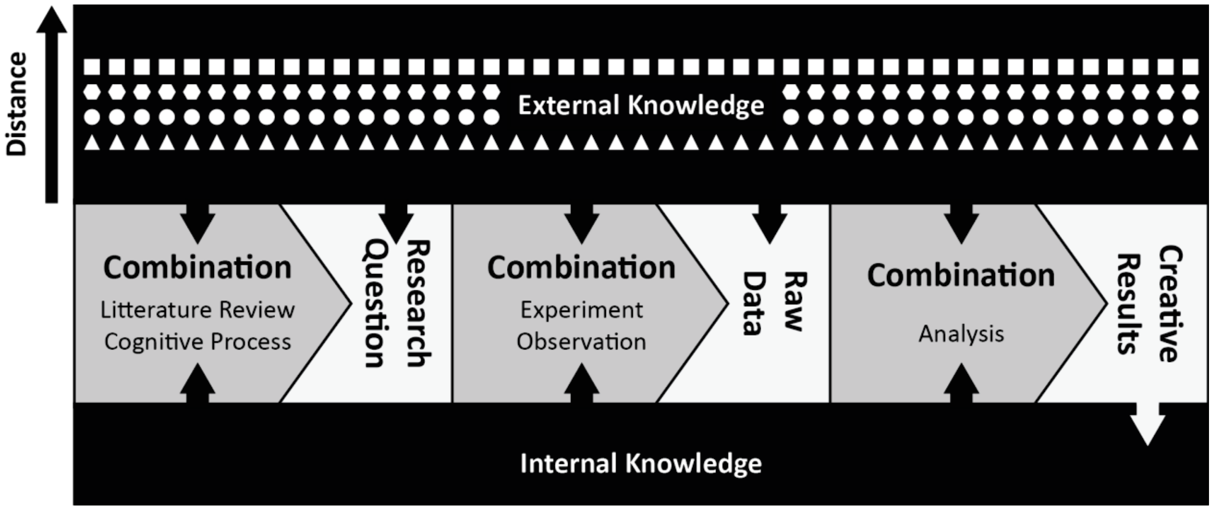


Figure 14: Collective knowledge-creation process with data (results)

The figure above shows a collective knowledge creation process. Traditionally the external knowledge that entered the production process was limited in size because of how costly it was to get. Indeed, organizations used very little external knowledge and it consisted mostly in formal knowledge in the form of publication. This knowledge was present mostly in the early stages of science production. As we expected the amount of knowledge that is accessible thanks to databases has become larger as well as its variety. However, what we have found additionally to this, is that in the domain of biology virtually all the existing knowledge tat has been produced and formalized is accessible. This increases not only the share that the external knowledge represents compared to the internal one, it also increases the variety of pieces of knowledge that can be combined in order to produce new knowledge. Additionally, this knowledge is now present at all the stages of the science production process.

4.4 Serendipity

Serendipity appears as an unexpected result. We have briefly seen in Chapter 1 how creative science is often explained in terms of serendipity. We did not expect databases to have an impact on serendipity and therefore its theoretical implications were not studied earlier in this chapter. We discuss here what literature tells us about serendipity and how to relate it to our empirical findings.

Serendipity in science defines the notion of unexpected and beneficial discoveries. Historically it has played an important role in explaining how big breakthrough discoveries are made (Merton and Barber, 2004). The description which often refers to the notion of "happy accident" is rather imprecise and therefore a multiplicity of phenomena can be described as serendipitous. This is the reason why Yaqub (2016) has discussed the heterogeneity of the phenomena and created a typology of the different kinds of serendipity and the mechanisms that lead to them. By exploring the most famous cases of discovery through serendipity as well as the literature that studies them, he created a typology.

His paper classifies serendipity according to how it shows up or in other words, what makes a discovery serendipitous. This leads to 4 kinds of serendipity. The first one, **Walponian** serendipity consists in targeted search leading to an unanticipated discovery. Researchers were looking into one problem and made a discovery into another. This happened sometimes as the result of an accident or simply because some effect (which wasn't the one they were looking for) showed up during the experiment. This is the most known kind of serendipity which led to the idea of a "happy accident". The second kind of serendipity is **Mertonian** and it occurs when targeted search solves a problem via an unexpected route. It consists in researchers trying to solve a problem and finding a solution in an unintended way. For instance, by accidentally mixing some components that were not expected to lead to a solution but do. These two kinds of serendipity correspond to the examples explained in the previous section of discoveries that were allowed thanks to the use of databases and were serendipitous.

There are still two additional kinds of serendipity that have been identified by the literature. There is, for instance, the **Bushian** type, where an untargeted research (often exploratory research and sometimes an activity which is not research at all) leads to an important discovery. Some examples are the x-rays or the re-purposing of drugs during clinical trials. Finally, there is the **Stephanian** kind of serendipity which consists in untargeted search making a discovery that solves a later problem. Stephan defines it as “finding answers to questions not yet posed” (Stephan, 2012, p232). It is the discovery of an interesting phenomenon that later will solve a problem. Because serendipity was a topic that emerged during the interviews and that we had not explored the literature before, we do not know if these kinds of serendipity can apply to the case of biological databases, we think, however, that it is a topic to be studied.

A common point to the four kinds of serendipity described above is that at some point there is a connection between two elements and an effect which is observed. This capacity of observing this effect is, according to Yaqub, very relevant for serendipitous discoveries to happen. We have seen, in the results on serendipity that have been explained earlier, how databases increase the possibilities of making connections between knowledge bases. In a second stage of his research Yaqub looks at the underlying mechanisms of serendipity. Which elements allow the humans to turn an accident or an unexplained phenomenon into a discovery and how do these elements work together. These mechanisms are often all present at the same time and do not exclude each other.

Firstly, there is **Serendipity and theory**. Theory or experience allows any given observer to identify the serendipitous episode as being incongruous with predictions and expectations. It might be an unexplained phenomenon that raises a question about what we know. The growth of theory may guide the observer on where to look, restricting the scope for their possible observations and inferences. Because we have theory, we know what is strange and that leads us to the discovery. In second place there is the **serendipity and the individuals**. It consists in having the skills, talents or experience to see how the unexpected discovery could be used. In third place there is **Serendipity and the tolerance of errors**. Mistakes play an important role on discovery and so does curiosity after an error happens. Finally, there is **serendipity and networks**. Networks are very relevant for serendipity; they play an informational role, bringing discoveries to the attention of researchers who can exploit them. They also play a teamwork kind of role. The exploitation and observation may require skills and resources of multiple people.

The second part of Yaqub's categorization shows how luck is important for serendipity to occur but it is a subordinate to researchers' skills, knowledge, curiosity and ability to share and communicate information. What they all have in common is the existence of an open-eyed or watchful state by the community. This is what makes researchers aware of what is out of the norm, which phenomena might be interesting.

To sum up the point, we found that the use of databases to had an impact on serendipity by enhancing some of the effects described here. It would be interesting to investigate further this data-driven serendipity. For instance, we may try to identify which types of serendipity, among those identified by Yaqub, are actually at stake in the case of large and varied databases. And also, what are the main underlying mechanisms that explain such "digital" accidental discoveries?

5 Conclusion and perspectives

As we saw at the introduction of this chapter, the use of large databases and the dependence of pharmaceutical industry on them is becoming more and more relevant in research. In this chapter we show that this use does not only provide with a gain on productivity as one could thing, it also offers supporting conditions for creativity. Indeed databases put together knowledge coming from multiple disciplines and communities and make it accessible to everyone. They help to overcome problems related to knowledge distance.

All of this is very relevant because databases are becoming crucial, not only for pharmaceutical research, but also for several other fields of science. The main contribution of this chapter has been to understand in which way the access to large public biological databases could favour scientific creativity. Databases can be, not only one more tool but also the resource that allows communities to share their knowledge and more likely be more creative. This is particularly relevant in a world where medical research is depending more and more on data. The use of large databases in medical research has a big potential in solving long term unsolved scientific challenges.

Another relevant result is the existence of favourable factors to serendipity. Serendipity has always been considered as a happy accident which allows for discovery. Although this “accident” was enabled by the fact that the person observing it had enough knowledge (or intuition) to consider it relevant, it was still considered a matter of pure luck. Our research suggests that databases could boost these kinds of accidents. The reason is that if we consider serendipity as the likeability to connect two heterogeneous pieces of knowledge which turn into a discovery, the more the pieces of knowledge an observer can understand, the larger the possibilities that she makes a connection. Also, the more observers analysing this pool of knowledge, the more the likeability that one of them makes a fructuous connection.

Another relevant contribution to theory is to show knowledge codification as a possible source of creativity rather than the end of creativity. Codified knowledge is easier to understand and allows for the understanding of distant knowledge which is one of the biggest problems when it comes to knowledge management.

There are however some drawbacks. One could think that because of the existence of standards and the presence of automation some interesting ideas could remain outside of the standards and produce an effect where some creativity is prevented because there is a lock in situation. Also, the existence of algorithms could generate a self-reinforcing effect where the same results (for instance combinations of pieces of knowledge) come out once and again and novelty is impeded. The access of everyone to all the available data could also create an effect where all researchers concentrate around the same topics, the low risk ones which allow for easy or fast results.

Further research should focus firstly on the long-term effects of Databases on creativity, for example the effect that databases have on the development of final products. When the interviews were done most of the companies were going through a process of depending the more and more on the use of external public data. A lots of projects that had been the result of the use of data were starting to arrive to the pre-clinical trial face and companies were, in general, at a moment of change between two different ways of doing science. If the same study was to be done today, there would probably be a lot more to say when it comes to scientific outcome.

It would be interesting as well to focus on the use of data at other stages of medical research such as clinical and pre-clinical trials. Finally, an exploration of the impact of large and varied databases on academic research would also be of high interest.

Chapter 4: Measuring RIs' impact on scientific creativity. The case of the synchrotron

1 Introduction

Previous chapters have shown how RI favours scientific creativity of users by explaining the favourable factors, mechanisms and creative outcome. Large RI provide, indeed, with some favourable factors for scientific creativity to occur and an appropriate environment for creative processes. Scholars that have studied scientific creativity show that a central element for creativity is variety in the elements that are combined in order to produce new knowledge. More specifically the idea of crossing disciplinary boundaries has been highlighted (Amabile, 1983; Chamorro-Premuzic, 2017; Heinze et al., 2009; Paulus and Nijstad, 2003; Simonton, 2004). Chapters 2 and 3 have shown us that these favourable conditions for creativity are offered by RIs. In the specific case of the synchrotron (Chapter 2) we know that it can be a proper place for cross-disciplinary collaborations. Indeed, we observed how different communities exchange and how, as a result, disciplines adopt methods that up to that moment were exclusive to other disciplines. We also observed, by means of qualitative studies, that creative outputs are likely to occur when using Large RI.

The objective of this chapter is to focus on the last part of the process and find a methodology to measure that impact on creativity of RIs at the level of the output. This chapter aims to be a complement to the previous ones. We have, up to now, seen the kind of factors that are favourable to creativity, the mechanisms through which they work and how they interact with each other. Here we will focus on the outcome and its measurement. The reason why we consider this research necessary is that to the best of our knowledge there are not wide spread methods to evaluate creativity in RIs (or other kind of Science policies for that matter). The impact of RIs is, however, studied. It is studied at several different levels, from purely scientific to societal impacts but without focusing on creativity. Thinking this is odd, considering that creativity is at the bases of knowledge creation

To do this we proceed as follows. Next section studies the state of the art of the literature that aims to measure creativity. To do so it will start by defining creativity in a way that can allows us to find a quantitative approach. After that it

2 Measuring scientific creativity

This section constructs a theoretical framework that we use to develop a proper methodology for evaluating the impact of RI on scientific creativity. There have been few attempts to measure scientific creativity. These are rather heterogeneous in terms of goals and features but as well in terms of which discipline developed it. We could consider it a patchwork of methodologies. Therefore, the first objective of this section is to map the various approaches to the measurement of scientific creativity and to assess how they can be used with a view toward impact evaluation. In order to achieve this, we will first remind how we define scientific creativity in this thesis and discuss the aspects of the definition that are relevant for the development of a quantitative approach to its evaluation. Secondly, we will briefly review the different attempts undertaken in previous literature to measure scientific creativity. Creativity being a broad concept we will focus on the measurement of a creative output. The reason is that we follow an impact logic and we seek to understand how RIs contribute to produce creative science. The second objective of this section is to discuss and choose a proper methodology for the systematic evaluation of creativity in science, a methodology later used in our empirical analysis regarding synchrotron impact.

2.1 Finding an operational definition of creativity

As it has been defined previously in this thesis, we consider creativity in science as the creation of a piece of knowledge that is new and valuable (Hollingsworth, 2004). Knowledge creation is therefore the output that we will consider here. Scientific output however can manifest itself in multiple ways, from the observation of a new natural phenomenon to the development of a new theory (Amabile, 1983; Ford, 1996; Woodman, 1993). Going from the concept of creativity to an operational definition that allows for empirical analysis is a challenge in studies focused on creativity in science (Amabile 1982). For that reason, empirical study of the creativity in science is an emerging field of research. Recently Heinze et al (2009) propose a list of creative scientific outputs:

1. Formulation of new ideas that open a new cognitive frame or bring theoretical claims to a new level of sophistication.
2. Discovery of new empirical phenomena that stimulate new theorizing.

3. Development of a new methodology by means of which theoretical problems could be empirically tested.
4. Invention of novel instruments that open new search perspectives and research domains.
5. New synthesis of formerly dispersed existing ideas into general theoretical laws.

This classification helps to better understand which kind of output we consider. More specifically when it comes to a unit of measure most scholars have agreed on considering count of publications, count of citations and impact factor of journals as the most representative ones. Any of the above-mentioned types of scientific output can (and probably will) become a scientific publication. New knowledge is creative if it provides new insights on relevant scientific problems. This means that creative research needs to be original. However, creativity is not only defined in terms of newness. For research to be creative it must also be valuable (Hollingsworth, R, 2004; Simonton, 2004). In this respect citations have often been considered as a good indicator of creativity because they consist in the recognition by peers of the value of a given scientific production. However, as we have seen in Chapter 1, among the previous criteria the one that is the most often disregarded is the criteria of novelty. Usefulness on the other hand is often acknowledged and rewarded via publications, grants and citations. Finding a single measure of creativity has been shown to be a complex task because of the gap between how the criteria that define creativity are rewarded.

For this reason, we argue that because creativity is composed of two attributes (novelty and value) these components should be both included in any measure of creativity. Moreover, we argue that novelty and value are attributes that, although having some degree of correlation, they often pull research in separate directions and have their own causal relationships. (Fleming, 2001; Lee et al., 2015; Yong et al., 2014). For these reasons a few authors (whose work is described in section 2.3) have recently decided to focus on the idea of novelty and offer novelty indicators as a complement of the widely spread indicators of excellence that are based on the count of publications, the use of citations and the journal impact factor (JIF thereafter) of the journals where the articles are published. However, before focusing on novelty, the next subsection examines the most important quantitative approaches to measure creativity of scientific production, including the limitations of the methods that have brought scholars to focus exclusively on novelty.

2.2 Metrics for the study of creativity in science

The early attempts to measure creativity in science were based on historical accounts by studying very well performing scientific institutions. The approach consisted in the identification of creative science, which allowed the identification of creative institutions, which finally allowed the researcher to study the organizational conditions within which creative research happens. The two main aspects of this approach are the following. First, it considers creativity in science as a binary attribute. Science can be creative or not creative. The methodology therefore seeks to identify the science which is creative. Literature using this approach will consider only the most creative science in a given field. This means that they focus on a reduced number of path breaking discoveries. The second important aspect is that this identification represents only an intermediary step needed to find creative institutions and study their organizational structures.

When identifying creative science some authors used prize winners as a measure of reference (for instance Zuckerman (1967) focuses on Nobel laureates). Laboratories with a large number of Nobel laureates were considered creative and their scientific institutions were studied to try to identify organizational structures favouring creativity. Because the study of Nobel laureates reduces the focus to a few cases which do not represent all the creative research, Hollingsworth (2004) extends the scope of creative research by including works who have been nominated not only for the Nobel but also other kinds of prizes. Both mentioned studies have had an important impact in the study of creativity in science and are very relevant to understand the organizational structures that are likely to foster creativity. However, these historical accounts focus mainly on the first half of the 20th century. The organization of science has gone through major changes and not all of what has been learned can be applied to today's science. To overcome this limitation Heinze and Bauer (2007) propose a methodology that is more likely to allow us to study the recent scientific developments. They identified creative science by doing a survey where they asked scientists to name research articles that they consider creative in their field. This allows them to identify and study creative research structures. This methodology has two main limitations. The first one is methodological, related to the fact that it does not allow for a systematic measurement of creativity. There is the requirement of fieldwork which means identifying the scientists of a field, creating a sample, contacting them and sending the survey. This makes it difficult to implement such a method on a systematic way (for instance continuous monitoring of creativity in the future). Additionally, and as it is the case for the mentioned methodologies, creativity is considered a binary attribute. Research is either creative or not. We can identify the most creative research in the field but given a specific piece of scientific output we cannot determine to which point it is creative. For this reason, the methodology is not suitable for impact evaluation logic. The second problem with this methodology is that by letting other scientists decide which research is considered creative and which research is not considered creative, we are facing every problem related to the delayed recognition of research. Indeed, scientists themselves have problems accepting big degrees of novelty that can potentially question the established paradigms.[44]

The next table summarizes the three quantitative approaches to the measurement of creativity in science that have been discussed above.

Table 21 Quantitative methods for the study of creativity

Work by	Aim of the study.	Identification method	Main results
Zuckerman (1967)	Organizational structures and managerial principles that foster creative research	Laboratories with a large number of **Nobel laureates**	Communication Cross-disciplinary collaboration [45]
Hollingsworth (2002, 2004)	Organizational structures that foster creative research	Laboratories with a large number of **nominees to other prizes**	Cross disciplinary departments Team variety (in terms of backgrounds)

---

[44] This has been largely studied in Chapters 1 to 3. See chapter 1 for more details.

[45] Other results that are not relevant for our study consisted in the independence of researchers when fixing research objectives

Heinze and Bauer (2007)	Organizational structures that foster creative research	**Surveys** asking scientists about creative research in their field	Team variety Cross institutional collaboration

2.2.1 Traditional focus on value through the measurement of impact

Impact assessment has traditionally been used to evaluate or monitor the performance of RIs. Impact assessment is driven by several rationales, from which three are highlighted by the literature as the most relevant ones (Bach et al., 2014). In the first-place policy makers want research organizations to monitor and understand their performance, both in terms of impact on the public in general as well as in terms of impact for academia. Secondly there is a need for accountability. Indeed, governments (but also other stakeholders and society in general) want proof of the value of research to justify the expenditure[46]. Finally, there is the will of understanding the methods that are leading to improve the results (both in terms of findings and societal benefits). What all three rationales have in common is the will to, by understanding impact, understand how to better allocate resources in the future (Donovan, 2008; Hanney and González-Block, 2011; Stegmann, 1997). When trying to understand the impact that public policies and more specifically RI have on science performance, most research is focused on metrics such as number and quality of publications which have traditionally been used to measure quality of research. Novelty, which is an important factor of creativity, which should be studied as part of the performance, is often disregarded. This is due to the well-known tension, which we developed in chapters 1,2 and 3, between the criteria of plausibility and novelty when evaluating science.

---

[46] European Science Foundation. Evaluation in National Research Funding Agencies: approaches, experiences and case studies, 2009. http://archives.esf.org/fileadmin/Public_documents/Publications/moforum_evaluation.pdf

2.2.2 Focusing on the measurement of novelty

The methodologies previously explained are a first step towards the quantitative measurement of creativity; however, they are not suitable because of their binary character and the difficulty to systematize them. Traditional methodologies for the evaluation of science focus more on impact which is only one aspect of creativity. In what concerts the other aspect, novelty, there is a bias in science against it in science. Novel science receives delayed recognition and often goes unnoticed for long periods of time because it takes a longer time to integrate the findings of this kind of research than those of incremental one. Novel science can be considered as revolutionary contrary to the rest which would be "normal" science (Barber, 1961). It is common for disruptive novel discoveries to be neglected or even attacked and rejected because there is very little to build up on or because they defy existing paradigms (Barber, 1961; Garfield, 1977; Stent, 1973; Tobias, 2009). This bias against novelty creates a need to understand how to recognize novel science and some authors have recently faced this problem.

2.2.2.1 The idea of combinatory novelty

Scientific discovery can be seen as a problem-solving activity where we combine perspectives, methods and tools in order to find solutions. In science we combine knowledge in order to create new knowledge (Stephan, 2012; Simonton 1989). How this recombination of previously acquired knowledge is done determines how creative the research will be. We consider novel science as science that is recombining knowledge in novel or unusual ways. This often means to combine multiple pieces of knowledge that have not been combined before or that are not very often combined. This view of novel combinations has been suggested by scholars in various disciplines (Mednick, 1962; Simonton, 2004; Weitzman, 1998). For example Nelson and Winter (1982) argue that new combinations of concepts or materials that already exist are the bases of novelty in art and science. Following a similar logic, the literature on network brokerage argues that individuals that are situated at the intersection of heterogeneous groups are more likely to generate new ideas. Because of its heterogeneity the knowledge inherent to these different groups is not often combined. For the individual that is in this brokerage position, the probability of being exposed to multiple knowledge sources is greater (Burt, 2004; Rodan and Galunic, 2004). When it comes to research, a way to foster this kind of brokerage position is to allow effective communication across boundaries of research disciplines.

This can be done for example by allowing researchers to work and communicate with people from other departments or by building teams of researchers with different backgrounds (Hage 2006; Hollingsworth 2004).

2.2.2.2 Using combination of citations

Building up on these ideas, Uzzi et al (2013) consider that what makes science creative is the balance between the use of atypical combinations and conventional concepts. To observe this, they propose an empirical methodology which consists in observing how previous knowledge is combined using citation data. They compare the observed frequency of the combination of citations with the expected frequency (if this choice was random). This gives each possible combination a score that shows whether a given pair is common or not. With this information they propose to observe for each scientific article the median of this measure for all its pairs of combined citations. This median enables them to characterize conventionality in the paper's main mass of combinations. They crossed this with the observation of the 10th percentile left tail tendency (of the distribution) of pairings in each paper where the rarest combinations are. Each paper has therefore two scores: one in median conventionality and another one on left tail novelty. They compared this information with the later success of the paper and found out that, as expected, a successful paper (which is considered by them the same as a creative paper) needs a central tendency that indicates conventionality combined with a small number of very rare combinations. Papers that use conventional well accepted combinations together with very novel ones are those with the highest long-term impact. The authors show that creativity requires the use of mainly highly conventional well-established knowledge but always combined with a small amount of highly novel work. The methodology for the empirical study of creativity can be summarized therefore in the use of a combination of indicators to evaluate creativity: an indicator of novelty and an indicator of commonness.

Lee et al (2015) study the increasing pre-eminence of team science and the need to understand the effects of team composition on the creativity of scientific results. More specifically they focus on the impact of the size of the team and of the variety of tasks on novelty. Building up on Uzzi et al (2013) they propose a measure of novelty defined by the commonness of pairwise combinations of prior work[47]. The highest the value of this indicator, the less novel the scientific output is. This method is built from the observation that combining certain pairs of knowledge domains is more novel than combining other pairs, and that a given paper draws from a variety of prior knowledge producing a knowledge combination distribution profile for each paper. When aggregating papers, they propose to use measures of centrality such as the median or the 10th superior percentile. They find that novelty relates to team size following an inverted U shape. This occurs because team size is correlated with the variety of fields of origin of the team members; hence variety, rather than team size, is the main factor for creativity. Indeed, the bigger the team size - thus the variety of backgrounds - the bigger the degree of novelty of the scientific results. Moreover, they find that team size and variety not only influence novelty of the scientific results but also has a positive relationship with impact, which leads to the idea of having more creative science.

Building on the aforementioned work, Wang et al (2017)[48] perform a study that examines the relationship between the two faces of creative research, namely novelty and impact of the results. Continuing the view of science as a combinatorial process of previous knowledge, they measure novelty in science by examining whether a published paper makes first-time-ever combinations of referenced journals. To do so they develop a measure of novelty and apply it to Web of Science (WoS) research articles across all scientific disciplines. They collect every journal article inventoried in WoS for the year 2001 as well as their cited journals. They then make pairs of journals that are co-cited in a paper and determine whether these two journals have already been mentioned together before or not. For each paper they build a measurement where each new combination of pairwise journals scores 1. Next, they compare the total score to the citation distribution variance. They found that highly novel papers, defined as those using newer combinations, deliver results that have a higher impact. More specifically, they are more likely to rank in the top 1% highly cited papers in the long run, to inspire follow-up highly cited research, and to be cited in a broader set of disciplines and that not only in

---

[47] Since we use this indicator in our methodology a detailed explanation is provided in the empirical section.

[48] For each paper they create a matrix with every referenced journal and examine whether the co-appearance is new. With this information, for each paper they construct a measure of combinatorial novelty.

the discipline of origin. The authors also found novel research to be riskier, meaning it shows a higher variance in its citation performance. They also obtained strong evidence about the delayed recognition of novel papers as valuable papers already mentioned in the literature. In the short-term novel papers are less likely to be top cited. Furthermore, they found that novel research is more likely to be cited in other fields of research than in the field of origin of the paper. Finally, when it comes to journal impact factor (JIF), novel papers are published in journals with a lower JIF than more conventional papers. These findings suggest that science policy and more specifically funding decisions which rely on bibliometric indicators based on short-term citation counts and JIF might be biased against novel research.

2.2.2.3 Using keyword combination instead of citations

Building up on all these studies, Carayol et al. (2018) propose a novel approach to the measurement of novelty based on the use of pairwise keywords as a proxy for the combination of knowledge. More specifically they develop an indicator of commonness for each paper. They consider all pairwise keywords combinations by papers published in a given year and a given research field. Research fields correspond to the subject categories of the WoS. Keyword combination frequencies are calculated within subject categories because the degree of novelty of the vocabulary of a publication is likely to be interpreted within a given discipline or community, not across disciplines. They study the set of all research articles published from 1999 to 2013 in the journals referenced by the WoS, which includes more than a million papers. There has been a growing concern that, though scientific production is globally increasing, the creativity of published research may be actually shrinking because of all the risks that come with the production of novel research. They found, however, no evidence of decay in scientific novelty during that period. The results of this paper are consistent with the work of Uzzi et al (2013). They also find, as in Lee et al (2015), that high novelty increases the chances of having a high impact (being highly cited) but a higher variance as well. Finally, they observe that highly successful papers are novel at the moment at which they are published but not that much a few years after. They claim that this occurs because the research questions are still active a few years after being novel. All these results are consistent across scientific fields.

All the mentioned studies have in common the consideration of novelty as a continuous attribute which allows for the evaluation of the science that has been produced as a result of a specific scientific policy such as the use of Research Infrastructure. The indicators used are therefore suitable for an evaluation approach. However, we still lack some tools that enables the comparison between two bodies of scientific output. If we want to assess synchrotron impact, we have to determine whether research produced at the synchrotron is creative or not and in relation to what. Additionally, this methodology is presented as a complement to traditional measures of impact and it evaluates only one of the two dimensions of creativity. Table 22 below summarizes the three quantitative approaches to the measurement of novelty in science that have been discussed above.

Table 22 Quantitative methods for the study of novelty

	Work by	Aim of the study.	Main results
Combination of cited work	Uzzi et al (2013)	Which kinds of novel combinations lead to creative science	Need for balance between the use of atypical combinations and conventional concepts
	Lee et al (2015)	The impact of team size and variety of tasks on novelty	Novelty increases with team size
	Wang et al (2017)	Whether novel science is or not likely to become high impact science	Novel science is likely to become relevant but also has a higher variance
Combination of keywords	Carayol et al (2018)	Whether science has stayed creative	Science has indeed stayed creative despite the fears that it would not.

2.2.2.4 Mapping the quantitative approaches to measure creativity

We have seen up to here the most relevant works when it comes to analysis of creativity. These studies can be divided into two main groups. The first one considers creativity as a binary attribute and focuses on the identification of the most creative achievements in science. This identification serves as a mean to study the conditions that are favourable to scientific creativity. The second group of analysis considers creativity of scientific output as the union of value and novelty. Value being an attribute that has well established methods to be studied[49] some scholars have recently focused on the analysis of novelty. The concept of novelty is based on the idea of science as a combinatorial activity. Thus, measurements of novelty have been focused on the novelty of the combinations done by science. This is done by using indicators where novelty is considered as a continuous attribute which allows for evaluating the degree of novelty of any scientific output.

---

[49] As explained before, these methods consist in count of publications, impact factor of publications and citations.

Table 23 Summary of quantitative methods to measure creativity and novelty

Goal		Identification Method	Work by
Identify most **creative** scientific output	Historical accounts	Nobel laureates	Zuckerman (1967)
		Nobel nominees and other prize laureates	Hollingsworth (2002, 2004)
	Study of recent science	Recognition by peers (survey asking for most creative scientists)	Heinze and Bauer (2007)
Evaluate the degree of **novelty** of scientific output.	Study of science at any time[50]	Combinatory novelty of citations	Uzzi et al (2013) Lee et al (2015) Wang et al (2017)
		Combinatory novelty of keywords	Carayol et al (2018)

3 An empirical exercise on synchrotron beamlines

Our working definition of scientific creativity is derived from the sociological idea that something is creative if it is both: on the one hand somewhat novel or unusual and, on the other hand, praised or at least accepted as potentially valuable by the peers (Heinze et al., 2009; Simonton, 2004). We first investigate how beamline papers are perceived by the scientific community and then we turn to the novelty aspect of creativity by looking at the extent to which citation given by a paper are novel. We will be working on the case of a

---

[50] As long as there is bibliometric information available.

synchrotron and more specifically we will be analysing the creativity at two of its beamlines (See Chapter 2 for the definition of Synchrotron and beamlines).

3.1 Methodology

We will use scientific articles as our subject of study and use them to measure knowledge creation or scientific "output". Considering our definition of creativity as knowledge which is novel and valuable our methodology will consist of two parts. In a first part of the analysis we evaluate the impact of the papers done at the synchrotron, as it represents the value perceived by the peers. In a second part we observe the degree of un(commonness) of citations as we consider it as a measure of novelty. Each of these two parts is composed by two steps or two units of measurements which gives us a four-step analysis and is done for each of the two beamlines.

3.1.1 Part 1: impact analysis

The first part consists on the impact analysis and is done in two steps: in a first step we compare journals where synchrotron science is published with other journals of the discipline. In a second step, we compare the synchrotron papers to other papers of the same journal. In other words, we examine whether synchrotron papers are published in "good" journals, and once published, whether these papers perform "better" than the other articles of the journal in terms of citations. We will always do this by separating the two studied beamlines: X-Ray (XR) and Infrared (IR).

Let us start by observing that each peer-reviewed scientific article should be considered somewhat creative because any peer-review process should guarantee that the accepted paper contains at least some novelty and is of interest to at least someone. But not all papers are equally interesting and novel. The journal in which a paper appears may be indicative. Journals differ in various aspects such as visibility, quality assurance, peer-review process, etc. All these aspects influence the (self-)selection of papers into journals and are arguably also somewhat correlated with journal impact factors. The social organisation of science is such that researchers prefer to publish and to do peer-review in high-impact journals rather than low-impact journals. This creates a self-sustained hierarchy of journals in which high-impact factor journals are able to publish the "better" papers. One may argue what exactly “better” means. As the common currency in science are citations received, "better" often means the prospect of earning more citations. The prospect of earning citations is probably more closely tied to relevance than novelty because of the already discussed problem of delayed recognition of novel science[51]. Therefore, we are interested in whether beamline research is published in the “better” journals, i.e. journals with higher impact factors, or not.

Beamline Impact

To investigate that we start by looking at beamline impact and we do as follow[52]:

1) We calculate for each journal and year an Impact indicator:I_{jt} where j refers to a journal and t refers to a year. This indicator consists of the WoS JIF normalized by the average of that field and that year.

$$I_{jt} = \frac{JIF_{jt}(WoS)}{(Field\ average)_{jt}}$$

When the Impact indicator is higher to 1, the impact factor of the journal is higher to the average of its field, when the indicator is equal to 1, the impact factor of the journal is exactly the same as the average of the field and when the impact indicator is smaller than 1 its impact factor is smaller than the average of the field.

2) We calculate the Impact indicator for each beamline and field: I_{bf} where b=beamline (XR or IR) and f=scientific field. To do so we do the sum of all the impact indicators by journal and year weighted by their frequency in the beamline

$$I_{bf} = \sum I_{jt} \cdot f_{jt}$$

f_{jt} is the frequency of journal “j” in year “t” for a given beamline, it is calculated by dividing the total number of beamline papers in journal “j” and year “t” (F_{jt}) by the total number of papers of beamline in year “t” and field “f” (F_{ft}).

---

[51] See Barber (1961), Garfield (1977), Stent (1973) and Tobias (2009) in Chapter 1

[52]

$$f_{jt} = \frac{Total\ number\ of\ beamline\ papers\ in\ journal\ \text{"j"}\ and\ year\ \text{"}t\text{"}}{Total\ number\ of\ papers\ of\ beamline\ in\ year\ \text{"t" and field "f"}}$$

$$f_{jt} = \frac{F_{jt}}{F_{ft}}$$

When I_{bf} is higher than 1, we consider the impact of beamline papers as higher than the average of the field, when I_{bf} is equal to 1 we consider the impact of the beamline papers as equal to the average of the field and when I_{bf} is smaller than 1 we consider the impact of beamline papers as smaller than the average of the field.

Number of citations

As an alternative measure of impact, we look at the number of citations. Here we will be observing which fraction of beamline papers receive more citations than the median number of citations of the same journal and year. This will tell us whether papers published using the synchrotron beamline are more or less cited than the rest of the papers in the same scientific journal.

3.1.2 Part 2: novelty analysis

For the second part of our study of creativity in science, we will study the second aspect of creativity which is novelty. This phase, as the precedent, is done in two steps. We continue using scientific journal articles as scientific outputs and the logic continues being to compare synchrotron papers to papers that did not use this synchrotron. Building on the idea of science as a combinatorial process we assume that scientists combine pieces of already acquired knowledge. These pieces of knowledge are identified as the citations. We may imagine that the citations represent the knowledge that has been used to produce science. We study novelty by using a commonness indicator. In a first step we use this indicator to study how often synchrotron papers cite “rare” journals and in a second step to study how often they do rare combinations of cited journals. All calculations are done by year and presented aggregated by discipline.

We follow the general idea that what is novel or rare depends on the audience. For example, in the 16th century porcelain has been very rare and novel in France – but not so much in China. The same holds for scientific instruments or methods. For example, 20 or 30 years ago, many econometric methods heavily used those days in economics had not been used (yet) widely in management research. We start from the idea that a journal tends to focus on both together a certain audience and/or research topic. Then, a paper inside a certain journal provides (hopefully) some insights for that audience advancing their research. References given by a paper may be interpreted in the same spirit. If a paper cites a paper of a certain journal, then the citing paper uses some insight related to the cited journal's audience or topic.

Journal Commonness Indicator

We look, therefore, at the citations of journals or, in other words, the references given. We first create an indicator that shows to which point beamline papers give uncommon citations. This commonness indicator is calculated as follows:

1) We download from WoS core collection all the papers for the journals and years where synchrotron research is published
2) For each paper of the universe we identify the number of citations going to each one of the journals of the WoS core collection
3) For each journal we calculate the probability of being cited as times cited over total number of citations (this is calculated by domain and year)
4) We attribute to each paper a vector of probabilities with one entry to each cited journal
5) We sort the vector from the least to the most likely citation
6) We take the value of the 10% less common citation given by the journal
7) We compare that value to the median value of the scientific field: Journal 10% value – Field median 10% value
8) We listed by beamline and field and we obtain the median of the commonness indicator that tells us whether beamline papers cite rarely cited journals or not. If the value is negative this means that synchrotron papers cite more rare journals than the rest of the papers in the field. If the value is positive this means that the beamline papers cite more common journals than the rest of the papers of the field.

The relative commonness indicator described above states whether the first 50 percent of beamline papers cite more or less common journals (and journal pairs) relative to all other papers in respective scientific category. If the relative commonness statistic is negative, the majority of beamline papers make relatively uncommon citations. If it is positive, only a minority of beamline papers makes relatively uncommon citations. What this indicator is telling us is whether papers use "rare" knowledge, or knowledge that is not often used in the field, or whether they use "common knowledge" or knowledge that is often used in the field. It is because of this reason that we consider it as a measure of novelty.

Pair Commonness Indicator

In a second step we look at (un)commonness of citation combinations. The idea of knowledge recombination suggests that novelty may also be obtained by combining different topics. Our "common pair" indicator provides information on unusual connections (again unusual with respect to the domain). We follow the lines of Lee et al. (2015) but have some differences: 1) our universe is not the whole WoS data set but rather the domain (journal and scientific field); 2) we remove citations to journals that are cited only once for two reasons: Firstly, unique citations are already considered by the "common journal" measure. Secondly, the importance of these citations gets blown up as it introduces many `unusual' pairs. The pair commonness indicator is built as follows:

1) We download from the WoS all papers from the fields and years that concern our investigation (fields and years where synchrotron research is published).
2) We list all the reference of each journal and pair all the journals that appear cited together.
3) We pool together all the journal pairs by year and field
4) We define the commonness of each journal pair as follows:

$$C_{ijt} = Commoness_{ijt} = \frac{observed\ number\ of\ pairs_{ijt}}{expected\ number\ of\ pairs_{ijt}} = \frac{N_{ijt}}{\frac{N_{it}}{N_t} \cdot \frac{N_{jt}}{N_t} \cdot N_t}$$

$$= \frac{N_{ijt} \cdot N_t}{N_{it} \cdot N_{jt}}$$

$$C_{ijt} = \frac{N_{ijt} \cdot N_t}{N_{it} \cdot N_{jt}}$$

5) For each paper we list all the present pairs of cited journals and we order them from the least to the most common. Once this is done we take the value of the 10% least common cited pair and record it. As done in the previous indicator, we compare that value to the median value of the domain: Journal 10% value – Field median 10% value

6) We aggregate by beamline and field and we obtain the pair commonness indicator that tells us whether beamline papers cite rarely cited journal pairs or not. If the value is negative this means that synchrotron papers cite more rare pairs than the rest of the papers in the field. If the value is positive this means that the beamline papers cite more common pairs than the rest of the papers of the field.

What this indicator is telling us is whether papers use "rare" knowledge pairs, or knowledge combinations that are not often used in the field, or whether they use "common knowledge pairs" or knowledge combinations that are often used in the field. It is because of this reason that we consider it as a measure of novelty.

[53]To summarize: The methodology consists on the analysis of the two faces of creativity: impact and novelty. When it comes to impact, we will study whether synchrotron papers are published in higher impact journals than the rest of the papers of the field and whether they are among the most cited within this journal. When it comes to variety, we will study whether synchrotron papers cite journals that are rarely cited and whether they do combinations (pairs of citations) that are rarely done.

---

[53] The relative commonness indicators described above state whether the first 50 percent of beamline papers cite more or less common journals (and journal pairs) relative to all other papers in resp. scientific category. If the relative commonness statistic is negative, the majority of beamline papers make relatively uncommon citations. If it is positive, only a minority of beamline papers makes relatively uncommon citations. In order to see whether this pattern should be considered systematic, we do a simple resampling test. The resampling test is performed as follows: Consider the beamline papers in each category as the original sample of size N. Create one resample by drawing from the original sample N beamline papers with replacement and calculate the commonness statistic (for the category) of that resample. We do that many times, here 1000 times, and note the fraction of times our statistic is greater zero. This is the significance level of rejecting the null hypothesis that the majority of beamline papers includes relatively common journal citations (pairs). We do this resampling test only for categories with 10 or more beamline papers

3.2 The data

3.2.1 Beamline papers

Beamline papers have been identified from the Infrared [54] and X-Ray[55] beamline websites. Bibliographic information for all beamline papers has been downloaded from Clarivate's Web of Science[56]. Both beamlines officially started operation in 2008. The analysis starts with papers published in 2010. We can be confident that at this point of time both beamlines have been fully operating and corresponding papers report research done with the instrument, rather than research on the instrument itself. 2017 is the latest year with full coverage. Furthermore, we kept only peer-reviewed scientific articles (i.e. for which the WoS field "Document Type" (DT) equals “Article”) and dropped other types of publications, mostly conference proceedings and reviews, from our sample.

Overall, between 2010 and 2017, there have been 615 peer-reviewed published articles based on research at the X-Ray beamline and 158 articles based on research at the Infrared beamline. There is no overlap between the set of IR beamline papers and the set of X-ray beamline papers, i.e. no paper is in both sets. Henceforth “beamline papers” refer to this selection.

3.2.2 Journal universe and scientific categories

Journal impact factors have been downloaded from WoS JCR Incites database[57]. Impact factors are based on the citations received by articles published in that journal and are available for all journals in the WoS Core Collection (CC)[58]. In any given year, the impact factor of a journal is the number of citations received in the two preceding years divided by the number of articles published in that journal during the two preceding

---

[54] https://www.synchrotron-soleil.fr/fr/publications?field_lignes_de_lumiere_tid=37

[55] https://www.synchrotron-soleil.fr/fr/publications?field_lignes_de_lumiere_tid=28

[56] www.webofknowledge.com/

[57] http://jcr.incites.thomsonreuters.com/JCRJournalHomeAction.action? , downloaded 03.09.2018

[58] $JIF_t = \frac{Citations_{t-1}+Citations_{t-2}}{Publications_{t-1}+Publications_{t-2}}$

years. Scientific categories of journals come from Clarivates Essential Science Indicators database (ESI). ESI ascribes each journal uniquely to one of 21 research fields or, if not possible, denotes it as multidisciplinary. In the analysis we use this categorization of scientific fields rather than the WoS core database fields “Web Category” (WC) or “Subject Category” (SC) because the former is broader and hence easier to grasp. Thus, we use information on impact factors from the WoS Core Collection and scientific category of journals from the ESI by joining both sets. The two sets of journals are not identical however. ESI covers only a (large) subset of the journals in the WoS core collection (CC). Journals found in both ESI and CC tend to have higher impact factors than journals only in CC[59].Nearly all beamline papers are in ESI. Only three beamline papers (all from X-Ray) have been published in two journals that are not in ESI but only in CC. For these journals, we complemented scientific category[60]. Impact factors are missing for 12 beamline papers (2 Infra-Red and 10 X-Ray papers) published in nine journal-year instances.

3.2.3 WoS notices

Finally, we downloaded (most of) the bibliographic records of the papers that appeared in journals where beamline research appeared between 2010 and 2017. This download was used to compare beamline papers with non-beamline papers. In detail, beamline papers appeared in 172 journals. For six journals we have no downloads at all (for no year): RSC ADVANCES, SCIENCE, SCIENTIFIC REPORTS, SOFT MATTER, SPECTROCHIMICA ACTA PART A-MOLECULAR AND BIOMOLECULAR SPECTROSCOPY, and VETERINARY RESEARCH. The reason is that these journals are very large in the sense that they publish a sizable number of articles per year compared to the rest of journals of our sample. This poses computational problems and might also bias the results if the size of the journals has some unexpected effects. For the other 166 journals present in our sample we downloaded most of the papers, in some big journals we downloaded papers for at least one journal-year. In total we obtained

---

[59] Although distributions are overlapping, median impact factor of journals in both ESI and CC is 1.32 while only in CC and not in ESI is 0.66.

[60] Journal CHEMISTRY & BIOLOGY has been ascribed to the scientific category BIOLOGY & BIOCHEMISTRY and ANTIBIOTICS-BASEL to category PHARMACOLOGY & TOXICOLOGY.

1,004,825 papers from 166 of these journals and when classifying them into journal and year we obtain 1551 journal-year sets. In cases where no downloads were available for a given journal-year of a beamline paper, we removed respective beamline papers from the sample. These are in most cases papers published in PLoS ONE, a very large journal. Sample selection may bias our results. However, we have complete coverage for the year 2010 (including PLoS ONE).

3.3 Analysis and results

In this section we analyse the data to see whether synchrotron science is (or is not) more creative than the rest of the science done in the same disciplines where synchrotron science is published. To do so we will compare our sample to the universe at two levels, the first level being impact and the second level being novelty, which are the two main attributes of creativity. This analysis will systematically be aggregated by beamline and by research field.

Beamlines tend to be specific for certain analyses typically used within specific research areas. Thus, papers from two different beamlines are likely to contribute to different research fields. Table 24 Descriptive statistics and Impact indicator provides us with some descriptive statistics on our data that will be useful to better understand the results on the impact indicator. The column “papers” shows how many X-Ray and Infrared papers contributed to each broader scientific category. Here we first focus on the distribution of papers over research categories. We see that both beamlines are mainly used for research in biology and chemistry albeit with some different focus. For instance, X-Ray papers are mostly published in journals focusing on BIOLOGY & BIOCHEMISTRY as well as MOLECULAR BIOLOGY & GENETICS while Infra-Red papers are mostly found in CHEMISTRY journals. Although Infra-Red papers are less numerous, they are distributed over more research fields. The columns “IR FRACTION” and “XR FRACTION” show the fraction of papers that belong to each of the categories. Table 24 contains as well information in our first indicator, impact, which we analyse in section 3.3.1 below.

Table 24 Descriptive statistics and Impact indicator

Scientific category	XR PAPERS	XR FRACTION	XR IMPACT	IR PAPERS	IR FRACTION	IR IMPACT
AGRICULTURAL SCIENCES	---	---	---	3	0.02	1.97
BIOLOGY AND BIOCHEMISTRY	218	0.36	1.79	12	0.08	1.27
CHEMISTRY	56	0.09	2.84	69	0.44	1.66
CLINICAL MEDICINE	---	---	---	3	0.02	1.17
COMPUTER SCIENCE	---	---	---	---	---	---
ECONOMICS AND BUSINESS	---	---	---	---	---	---
ENGINEERING	---	---	---	1	0.01	1.46
ENVIRONMENT AND ECOLOGY	3	0.01	2.64	---	---	---
GEOSCIENCES	---	---		6	0.04	1.52
IMMUNOLOGY	1	---	1.51	---	---	---
MATERIALS SCIENCE	1	---	1.83	3	0.02	2.47
MATHEMATICS	---	---		---	---	---
MICROBIOLOGY	61	0.1	1.9	1	0.01	0.62
MOLECULAR BIOLOGY AND GENETICS	108	0.18	0.98	1	0.01	0.87
NEUROSCIENCE AND BEHAVIOR	1	---	2.36	3	0.02	0.72
PHARMACOLOGY AND TOXICOLOGY	8	0.01	1.23	2	0.01	1.1
PHYSICS	4	0.01	0.94	20	0.13	1.13
PLANT AND ANIMAL SCIENCE	7	0.01	3.06	4	0.03	2.71
PSYCHIATRY AND PSYCHOLOGY	---	---	---	---	---	---
SOCIAL SCIENCES, GENERAL	---	---	---	---	---	---
SPACE SCIENCE	---	---	---	14	0.09	1.35
Multidisciplinary	137	0.23	3.61	14	0.09	2
Total	605	1	2.16	156	1	1.56

3.3.1 Scientific field and (relative) impact of beamline papers

In this section we are interested in the first aspect of creativity, impact, and we analyse whether beamline research is published in "better" journals, i.e. journals with higher impact factors, or not. Table 24 Descriptive statistics and Impact indicator provides a corresponding indicator "impact" which is based on WoS journal impact factors[61]. This impact indicator will be higher than 1 when, in average, synchrotron papers for a given beamline and discipline are published in journals with a higher impact factor than the average of the field. If the indicator is smaller than 1, then the synchrotron papers for a given beamline and discipline will be published, on average, in lower impact factor journals than the average of the field. See section 3.1 for the details on the calculation of the indicator.

Let us start by X-Ray papers in BIOLOGY & BIOCHEMISTRY, the main research category. We can observe that they fall in relatively high impact factor journals with our indicator "impact" calculated to be 1.79. This means that, in average, synchrotron X-Ray papers from this category have an impact factor that is roughly 79% higher than the average impact factor of the category. X-Ray papers in MOLECULAR BIOLOGY & GENETICS, the second most relevant field for the beamline, have an impact indicator of 0.98, which means that tend to be published in "average" journals. For MICROBIOLOGY and CHEMISTRY, the impact indicators are as well quite high: 1.9 and 2.84, although they represent a relatively small (10% each) of the total of the beamline publications. The big success story for X-Ray beamline is probably the high impact of multidisciplinary journals, obtained through articles in high impact journals such as Science and Nature. The rest of the categories represent a small number of synchrotron papers (less than 10%).

Let us now look at the case of the Infra-Red beamline and start by the papers in CHEMISTRY, the most relevant discipline for the field. The impact indicator is relatively high: 1.66; this means that synchrotron Infra-Red papers in CHEMISTRY are published

---

[61] Journal impact factors are calculated, as explained in previous section, on the citations received by articles published in that journal and hence vary across research fields and over time. This necessitates normalizing journal impact factors by research category and year, if we wish to determine at which end of the spectrum beamline papers (or rather respective journals) are positioned.

in journals that have, in average, an impact factor that is 66% higher than the average impact factor of the field. In PHYSICS, which is the second most relevant discipline for the beamline the impact indicator is 1.13; which means that the papers of the synchrotron are published in somewhat better journals than the average of the field. Finally, for BIOLOGY AND BIOCHEMISTERY, the impact indicator is 1.27; meaning that Infra-Red papers are published in journals whose average impact factor is 27% higher than the average impact factor of the field. The rest of the categories represent a small number of synchrotron papers (less than 10%).

Overall X-Ray users seem to publish in higher impact journals (normalized by year and scientific category) than Infra-Red users do and both tend to publish in relatively "good" journals when comparing them to the rest of the journals of the category.

We have just seen that synchrotron research is published in papers with a higher impact factor which means they are recognized as valuable by the community. To continue the analyses of the value of synchrotron research we still have to examine how synchrotron papers are cited compared to the rest of the papers in the same scientific journal (i.e. once in a "good" journal, is synchrotron research more cited than the rest of the research of that journal). Table 25 Fraction of beamline papers receiving more citations than the median provides the fraction of beamline papers that received more citations (until July 2018 when the data was downloaded) than the median number of citations received by papers of the same journal and year. We see that beamline papers do not systematically receive more citations than other papers in the same journal and year [62]. Synchrotron research is "average" compared to the rest of the journal. To illustrate, papers in BIOLOGY AND BIOCHEMISTRY for the X-Ray beamline were published in quite good journals, compared to the average of the field. However, once in a good journal only 38% of them receive more citations than the median number of citations of the journal.

---

[62] Because we do not have papers from all journal years where beamline research is published, our sample of beamline papers is slightly smaller (see above).

Table 25 Fraction of beamline papers receiving more citations than the median

Scientific category	XR CITATIONS	XR PAPERS	IR CITATIONS	IR PAPERS
BIOLOGY AND BIOCHEMISTRY	0.38	220	0.55	11
CHEMISTRY	0.4	58	0.43	67
ENVIRONMENT AND ECOLOGY	0.67	3	---	---
IMMUNOLOGY	1	1	---	---
MICROBIOLOGY	0.44	61	---	1
MOLECULAR BIOLOGY AND GENETICS	0.43	113	1	1
NEUROSCIENCE AND BEHAVIOR	1	1	0.67	3
PHARMACOLOGY AND TOXICOLOGY	0.56	9	---	2
PHYSICS	0.25	4	0.40	20
PLANT AND ANIMAL SCIENCE	0.17	6	1	4
AGRICULTURAL SCIENCES	---	---	0.33	3
CLINICAL MEDICINE	---	---	0.25	4
ENGINEERING	---	---	1	1
GEOSCIENCES	---	---	0.5	6
MATERIALS SCIENCE	---	---	---	3
SPACE SCIENCE	---	---	0.64	14
Multidisciplinary	0.36	83	0.17	6
Total	0.40	559	0.45	146

Table 26 provides the same statistic over years. We see again that synchrotron research receives an average number of citations if we compare to the rest of the journals. It is often inferior (although very close) to the median of the field. One may expect that more radical research findings tend to attract citations later than incremental (This idea has been commented at the literature review). It is important to notice that time pattern is consistent with the idea that beamline research attracts citations later than reference papers. Older articles (which had more time to get accepted by the community) get better results than recent ones in terms of citations. See for instance X-Ray papers. For the ones published in 2016 only 35% receive more citations than the median of the journals where they are published. For the papers published in 2010, which had the time to be recognized by the community, however, 48% receive more citations than the median of the journals where they are published. For the case of Infra-Red a very similar pattern is drawn. For the papers published in 2017 only 44% received a higher number of citations than the median of the journals where they are published. When we look at the papers published in 2010, which had more time to be assimilated by the community, 67% of them receive more citations than the median of the journals where they are published. These results could suggest that late recognition of creative research is behind this result. However, this effect is not very strong.

Table 26 Fraction of beamline papers receiving more citations than the median aggregated by year

	X-Ray	Infrared
2010	0.48	0.67
2011	0.43	0.55
2012	0.48	0.41
2013	0.39	0.53
2014	0.36	0.55
2015	0.39	0.47
2016	0.35	0.16
2017	0.40	0.44

To summarize, we found: i) that beamline papers are published in “better journals” or in other words journals that tend to attract relatively many citations and have a high impact factor, and ii) that beamline papers do not attract systematically more citations than other papers published in the same journals. We observe some pattern of increasing relative number of citations over years. Although not strong, this pattern is consistent with the idea that beamline papers provide greater novelty than other papers because papers with more novelty tend to attract citations later and novelty is a crucial attribute of creativity. Therefore, the next subsection investigates the novelty of papers, using bibliometric indicators based on the references given in the papers.

3.3.2 (Un-)commonness of journal citations and citation combination in beamline papers

This section consists of the second step for the study of creativity and we will study the second aspect of creativity which is novelty. This step, as the previous one, is double-folded. Because we understand science as a combinatory process we will assume that when doing research scientists combine pieces of already acquired knowledge that are expressed in terms of citations. We assume that the citations represent the pieces of knowledge that have been used to produce science. We will study novelty by using a commonness indicator. In first place we will use this indicator to study how often synchrotron papers cite “rare” journals and in second place to study how often they do rare combinations of cited journals. This analysis will be done by journal and year and aggregated by discipline.

3.3.2.1 (Un-)commonness of journal citations

As explained above, we follow the general idea that what is novel or rare depends on the audience. Journals tend to focus on both together a certain audience and/or research topic. Then, a paper inside a certain journal provides some insights for that audience advancing their research. References given by a paper may be interpreted in the same spirit. If a paper cites a paper of a certain journal, then the citing paper uses some insight related to the cited journal's audience or topic.

Citations to journals that are relatively rarely cited by papers in a certain journal may therefore introduce some rare or “uncommon” topic (or type of knowledge, or domain). Our “common journal” indicator exploits that idea. This indicator which is based on the work of Lee (2015) is explained earlier at the beginning of section 3. In Table 27 (X-Ray) and Table 27 (Infrared), we report for each scientific category the median of the "common journal" indicator over all beamline papers. A negative median means that most beamline papers in that category have relatively many uncommon journal citations, a positive median means that most beamline papers have relatively few uncommon journal citations. The tables show 5 colons: number of papers, which indicates de size of the sample for a given field and beamline; commonness journals, which is the result of the commonness indicator that has been explained before; significance level journals, which is the significance level of the commonness indicator; commonness j.pairs which is the commonness indicator for journal combinations and finally sig level j.pairs which is the significance level for the commonness of cited combinations.

Let us start by examining the case of the X-Ray beamline (Table 27). As we can observe, the tendency is to cite papers that are not necessarily very rarely cited, in other words there is not a lot of citations of rarely cited journals, citations are rather common. This is true for the globality of X-Ray papers but it changes if we look at each discipline. The most represented disciplines: BIOLOGY AND BIOCHEMISTRY and MOLECULAR BIOLOGY AND GENETICS follow this same described pattern. Their result is positive, which means that for more than half of the papers of the beamline and discipline the commonness indicator is positive as well, which means that they usually cite more common papers than the median of the field (category). However, in CHEMISTRY, for example, we can see that X-Ray papers cite relatively uncommon journals. The indicator is negative, which means that for more than half of the papers of the beamline and discipline the commonness indicator is negative as well, meaning that the median of the journal commonness indicator of the sample is smaller (so less common) than the median for the universe.

Table 27 (Un)commonness indicators (journal and pairs): X-Ray beamline

	Number of papers	Commonness journals	sig level journals	Commonness j. pairs	sig level j. pairs
BIOLOGY AND BIOCHEMISTRY	220	0.0003	1	-0.0086	0.073
CHEMISTRY	58	-0.0001	0.009	-0.0378	0.007
ENVIRONMENT AND ECOLOGY	3	-0.0002	---	0.0939	---
IMMUNOLOGY	1	-0.0002	---	-0.0367	---
MICROBIOLOGY	61	0.0001	0.977	-0.0288	0
MOLECULAR BIOLOGY AND GENETICS	113	0.0002	0.997	-0.0429	0
NEUROSCIENCE AND BEHAVIOR	1	0.0004	---	0.1539	---
PHARMACOLOGY AND TOXICOLOGY	9	0.0002	---	-0.0302	---
PHYSICS	4	-0.0007	---	0.9854	---
PLANT AND ANIMAL SCIENCE	6	-0.0002	---	-0.0063	---
Multidisciplinary	83	0.0002	1	-0.0599	0
Total	559	0.0002	1	-0.0301	0

Table 28 shows the same results for the Infrared beamline and we can see a completely different pattern. We see in first place that for almost every represented discipline (except ENGINEERING which contains only one paper and GEOSCIENCES) the citation analysis shows that Infra-Red publications tend to cite relatively uncommon journals. This means that synchrotron's X-Ray papers introduce, in the journals where they publish, knowledge from other journals that the publication journal rarely cites. This means as well that they introduce in the journals where it is published some new concepts (new for the audience/discipline).

Table 28 (Un)commonness indicators (journal and pairs): Infra-Red beamline

	Papers	Commonness journals	sig level journals	Commonness j.pairs	sig level j.pairs
AGRICULTURAL SCIENCES	3	-0.0001	---	-0.0104	---
BIOLOGY AND BIOCHEMISTRY	11	-0.0002	<0.000	-0.0995	0.035
CHEMISTRY	67	-0.0002	<0.000	0.0046	0.563
CLINICAL MEDICINE	4	-0.0001	---	-0.0358	---
ENGINEERING	1	0.0004	---	-0.1577	---
GEOSCIENCES	6	0	---	0.083	---
MATERIALS SCIENCE	3	-0.0002	---	0.066	---
MICROBIOLOGY	1	-0.0004	---	-0.0253	---
MOLECULAR BIOLOGY AND GENETICS	1	-0.0005	---	-0.0225	---
NEUROSCIENCE AND BEHAVIOR	3	-0.0002	---	-0.0327	---
PHARMACOLOGY AND TOXICOLOGY	2	-0.0001	---	-0.0171	---
PHYSICS	20	-0.0006	0.022	-0.0068	0.451
PLANT AND ANIMAL SCIENCE	4	-0.0003	---	0.0188	---
SPACE SCIENCE	14	-0.008	0	-0.2623	0.001
Multidisciplinary	6	-0.0001	---	-0.1956	---
Total	146	-0.008	0	-0.2623	0.024

3.3.2.2 (Un-)commonness of combinations of journal citations

As explained above, the idea of knowledge recombination suggests that novelty may also be obtained by combining different topics. We will use a “common pair” indicator that provides information on unusual connections (again unusual with respect to the domain). We follow the lines of Lee et al. (2015) as explained in section 3.1.2 but have some differences: 1) our universe is not the whole WoS data set but rather the domain (journal and scientific field), 2) we remove citations to journals that are cited only once for two reasons: Firstly, unique citations are already considered by the “common journal” measure. Secondly, the importance of these citations gets blown up as it introduces many “unusual” pairs. The relative commonness indicator described above state whether the first 50 per cent of beamline papers cite more or less common journal pairs relative to all other papers in respective scientific category. If the relative commonness statistic is negative, most beamline papers make relatively uncommon combinations of citations. If it is positive, only a minority of beamline papers makes relatively uncommon combinations of citations.

Results are interesting in that one type of un-commonness does not imply the other type and results vary by field. Let us observe the case of X-Ray beamline (Table 27) which did not behave particularly well in terms of citations of uncommon journals. When it comes it comes to novel combinations of cited journals it does show some novelty. Indeed, when it comes to commonness of cited pairs the result is the most often negative (and always negative in the most frequent disciplines). This means that more than 50% of the X-Ray beamline papers cite relatively rare pairs of journals compare to the median of the field.

For the case of Infrared beamline (Table 28), this performed well in terms of citing uncommon journals; we see that it does as well uncommon combinations of citations. For most of the fields the result is negative, which means that more than 50% of the Infrared beamline papers cite relatively rare pairs of journals compare to the median of the field. However, we observe that for the most frequent category (the one that includes most of the beamline papers) this indicator is positive, meaning that

To sum up, the results of the empirical analysis show us the following. We aim to study scientific creativity, which is defined in terms of value and novelty. As a proxy for value we study impact and we observe the following. Research done at the synchrotron is published in journals that have a high impact. More specifically it is published in journals that have a higher Journal Impact Factor than the average of the field. When it comes to number of citations (another measure of impact) journals published using synchrotron have an average number of citations compared to the rest of the papers of the same journal. In other words, synchrotron research is published in relatively good journals but then it only gets an average number of citations compared to the rest of the papers of the journal. The second part of the analysis consists on the study of novelty. As a proxy of novelty of combined knowledge, we use novelty of citations. More specifically we look at whether synchrotron research cites journals that are rarely cited in the journal where it is published. In the case of the X-Ray beamline we observe that it rather common journals but it combines these citations in an uncommon manner. Therefore, X-Ray research doesn’t use a lot of new knowledge, compared to the rest of the research in the field, but it combines knowledge in an uncommon manner. When looking at the Infrared beamline we see that it cites rarely cited (uncommon) journals and it generally does well as well in terms of combination of citations, except from one category (which is the most relevant category of the beamline).

We can conclude that, according to our methodology and the indicators used, synchrotron research performs well in terms of both, impact and novelty of the findings. The results, however, vary slightly from one technology to another.

3.4 Discussion of results

As discussed earlier, the results seem to show that the research done at the synchrotron is more creative when comparing it to the rest of the research of the field. It has more impact and it is less common from what we can infer than it is more valuable and more novel. The validity of this affirmation depends however on the validity of the indicators. Are our indicators valid as indicators of value and novelty? When it comes to value of academic research the count of publications, classing them by impact factor of the journal and the count of citations are widely spread indicators of research value[63]. As for novelty, we use the well spread idea of considering science as a combinatory exercise where previous knowledge is combined and take the citations given by a journal as a representation of this combinatory exercise. One could argue that considering a scientific journal as representing a current of knowledge (sub-discipline, topic, and approach) is not valid or at least could be questioned. Additionally, the specific indicator used for these purposes comes from a relatively new paper (Lee 2015), however it has been used by several other papers and cited several times which indicates that there is at least some agreement on its validity.

The weakness on measuring value through impact factor and count of citations is that these kinds of indicators, although widely used, are not without criticism and some alternative indicators have been built to overcome the weaknesses of JIF. It would be an interesting track to do the same analysis with the Eigenfactor (instead of the JIF). The Eigenfactor newspaper ranking was developed by Jevin West and Carl Bergstrom at the

---

[63] Although the Journal Impact Factor presents severe limitation for certain disciplines (among others: a short two-year lag between publication date and citation of a article, a bias toward "hard" sciences and American journals, the non comparability of different disciplines) , these limitations are not effective in the case of life sciences and chemistry that are at the core of our analysis. They would nevertheless question our approach if we wished to extend and apply it to Social Science, Humanities or Mathematics.

University of Washington (Eisenhardt, 1989). It avoids many JIF defaults through a random search of the entire citations network using a Page Rank algorithm (at Google).

Concerning the first commonness indicator, an important flow is that it treats all cited journals equally regardless of their size when classifying them as common or uncommon. This could make us think that small journals (e.g. those publishing a small number of articles per year) are less likely to be cited and therefore considered always as less common (or in other words more novel). Although this is true and the issue deserves some thought, we believe our commonness indicator is still a representation of novelty as, regardless of the reason, there is novelty on citing a small rarely cited journal.

Considering our methodology in its entire, it can be considered as the first contribution of this chapter. So far, most research had focused on impact of research rather than its variety. After literature (Garfield, 1977; Heinze et al., 2009; Hollingsworth, 2002; Simonton, 2004) pointed out that novelty is not always rewarded in science and that often science fails to recognize very novel research, some authors have focused on the study of novelty in science (Carayol et al., 2018; Lee et al., 2015; Uzzi et al., 2013; Wang et al., 2017). What makes our methodology original is the fact that we look at both aspects, value and novelty (which are calculated as impact and uncommonness).

To summarize one of the main contributions of this chapter is methodological, as it offers an original methodology which stands on widely used and accepted indicators and approaches which gives the methodology some reliability. The reliability of the results stands on these same grounds. These indicators, however, are not without criticism. Because of this calculation should be done again using alternative indicators.

We have commented up to now the internal validity of the results by discussing whether the indicators are valid to measure what we want to measure. We discuss here the external validity of the results or, in other words, the possibility to extrapolate what we have learned in this chapter for other cases (for example another RI).

Concerning the methodology itself, a particularly relevant flaw remains and should be addressed in the near future. We propose a methodology which is suitable for creativity evaluation. This methodology however does not provide with a single indicator but four of them. One may ask, what can be done when these four different indicators point in different directions? There are several alternatives here. One could argue that for something to be creative it should be considered as both valuable and novel. We could also think that if one of the two attributes (valuable or novel) scores highly, the other should be allowed to score under the average. What we propose here is that this indicator allows its users to have freedom when deciding under which priority rules the criteria should be implemented. However, further research could investigate a single indicator that would include all the aspects that have been studied here.

A second problem that this methodology poses consists in the need of scientific publications. Indeed, to be able to apply this methodology using an impact evaluation rationale we need to have a list of publications associated to the public policy (for instance the RI) that we want to evaluate. This information is often not available. From what we have learned during the fieldwork done for Chapters 2 and 3, when they publish their results the RI users do not always acknowledge properly the RI contribution for conducting their research.

To finish, let us look at the results for the empirical analysis as such. We observe that synchrotron research is more creative than research done without using the synchrotron. To which point can we extrapolate the results here to other synchrotrons or even other RIs? The answer is, as in any case study, this empirical analysis is valid only for the synchrotron beamlines studied. As observed, the results vary from one beamline to another (although they generally go roughly on the same direction), we can expect results to vary as well in all the other beamlines of the synchrotron, other synchrotrons or other RI. However, observing that Synchrotron research is more creative gives us reasons to think that other RIs that share characteristics might be creative as well. This should be validated by applying the methodology developed here to other RIs.

4 Conclusion

This chapter has proposed a methodology to measure the impact that Large RIs have on scientific creativity, which consists on a two-part process that analyses the two attributes that define creativity, namely value and novelty. The first part studied whether research done at the RI is published in higher impact factors journals than research that is not done using the RI and whether it receives a higher number of citations. The second part looked at novelty. More specifically and following a logic of combinatory novelty, we analysed whether synchrotron cites uncommon (rarely cited) journals and whether these citations are combined in an uncommon manner. The different indicators included in this methodology are based on widely accepted indicators to the study of bibliometric databases.

We analysed the case of two beamlines of [64] a European Synchrotron (a type of RI). To this end we applied our original methodology to all the scientific articles that are listed at the beamline's websites. We downloaded information on these articles from the WoS, as well as information on the other articles published in the same journals in order to have a universe to compare to. What we observed is that synchrotron research seems to be more creative as it scores well in both, impact and uncommonness of citations. For impact evidence shows that synchrotron research is published in relatively high impact factor journals. However, in terms of citations received they have an average performance which could be explained by delayed recognition on novel research, as we see that some years after being published their citation count (always compared to the universe) improves. For novelty we observed a difference between the two studied beamlines. The Infrared beamline showed itself as producing more novel research as we could observe how it does uncommon citations and it often combines these citations on an uncommon manner. X-Ray research, however, does not cite particularly uncommon journals, however it does combine the citations in a rather uncommon manner.

Taken together our results suggest some impact of synchrotron on value and novelty of research, and therefore on creativity. Further research should enlarge the study to other synchrotron beamlines as well as to other RIs. Regarding the methodology, it could be used as well to the study of other scientific policies as far as publications are easily identifiable but it should be as well contrasted by testing it with other kinds of indicators.

---

[64] Research at synchrotrons is organised by beamlines. A beamline is a set of equipment that modifies and brings the synchrotron light beam to the material being studied and records what happens. The studied synchrotron hosts 29 beamlines. Chapter 2 develops with more details.

Chapter 4

Appendix Chapter 4

Descriptive statistics of journal references are presented in this table. Descriptive are sums or averages over years within scientific category. In detail: papers is the number of papers, citations is the total number of references in all papers, matched citations are citations with identified WoS journal, avg. citations is the average number of citations per paper, Coeff.Var.citations is the average (over years weighted by number of papers) of the coefficient of variation of citations per paper (with coefficient of variation being standard deviation over mean), avg. journal citations is the average number of journals cited in a paper, Coeff. Var. journal cit. is the corresponding coefficient of variation (again averaged over years).

Category	set	papers	citations	matched citations	avg. citations	Coff. Var. citations	avg. journal citations	Coff. Var. journal cit.
AGRICULTURAL SCIENCES	IR	3	117	100	39		20	
AGRICULTURAL SCIENCES	XR	0	0	0				
AGRICULTURAL SCIENCES	cRef	25077	883259	760391	35.22	0.32	19.24	0.38
BIOLOGYAND BIOCHEMISTRY	IR	11	511	471	46.45		28.82	
BIOLOGYAND BIOCHEMISTRY	XR	220	10809	10367	49.13	0.3	26.08	0.28
BIOLOGYAND BIOCHEMISTRY	cRef	123697	5744164	5479040	46.44	0.44	25.56	0.36
CHEMISTRY	IR	67	3262	2926	48.69	0.4	29.37	0.39
CHEMISTRY	XR	58	2655	2408	45.78	0.39	24.19	0.37
CHEMISTRY	cRef	259631	11676512	10663533	44.97	0.55	21.27	0.42
CLINICAL MEDICINE	IR	4	150	137	37.5		26.25	
CLINICAL MEDICINE	XR	0	0	0				
CLINICAL MEDICINE	cRef	5230	231138	220414	44.19	0.42	26.95	0.37
ENGINEERING	IR	1	50	43	50		16	
ENGINEERING	XR	0	0	0				
ENGINEERING	cRef	4169	135644	106290	32.54	0.5	11.84	0.54
ENVIRONMENT/ECOLOGY	IR	0	0	0				
ENVIRONMENT/ECOLOGY	XR	3	197	190	65.67		37	
ENVIRONMENT/ECOLOGY	cRef	2170	131787	121751	60.73	0.3	31.15	0.27
GEOSCIENCES	IR	6	442	376	73.67		22.5	
GEOSCIENCES	XR	0	0	0				
GEOSCIENCES	cRef	7874	587023	481479	74.55	0.44	23.74	0.4
IMMUNOLOGY	IR	0	0	0				
IMMUNOLOGY	XR	1	35	31	35		13	
IMMUNOLOGY	cRef	1322	43833	40878	33.16	0.6	17.36	0.48
MATERIALS SCIENCE	IR	3	84	69	28		15	
MATERIALS SCIENCE	XR	0	0	0				
MATERIALS SCIENCE	cRef	16686	603392	558443	36.16	0.38	19.76	0.38
MICROBIOLOGY	IR	1	12	11	12		11	

MICROBIOLOGY	XR	61	3495	3364	57.3	0.28	29.33	0.28
MICROBIOLOGY	cRef	35648	1865050	1766917	52.32	0.43	26.37	0.39
MOLECULAR BIOLOGYAND GENETICS	IR	1	44	42	44		35	
MOLECULAR BIOLOGYAND GENETICS	XR	113	4565	4400	40.4	0.4	21.52	0.32
MOLECULAR BIOLOGYAND GENETICS	cRef	20241	973645	943834	48.1	0.41	24.98	0.35
NEUROSCIENCEAND BEHAVIOR	IR	3	134	126	44.67		30.33	
NEUROSCIENCEAND BEHAVIOR	XR	1	17	16	17		11	
NEUROSCIENCEAND BEHAVIOR	cRef	7406	265431	242265	35.84	0.62	18.93	0.55
PHARMACOLOGYAND TOXICOLOGY	IR	2	44	37	22		15.5	
PHARMACOLOGYAND TOXICOLOGY	XR	9	450	430	50		27.44	
PHARMACOLOGYAND TOXICOLOGY	cRef	19793	844480	770818	42.67	0.66	24.43	0.59
PHYSICS	IR	20	591	502	29.55	0.49	14.75	0.51
PHYSICS	XR	4	138	124	34.5		12.75	
PHYSICS	cRef	161403	5251575	4752648	32.54	0.51	13.83	0.44
PLANTAND ANIMAL SCIENCE	IR	4	195	179	48.75		26	
PLANTAND ANIMAL SCIENCE	XR	6	403	386	67.17		33.33	
PLANTAND ANIMAL SCIENCE	cRef	14724	879002	828254	59.7	0.37	29.11	0.3
SPACE SCIENCE	IR	14	930	801	66.43		18.57	
SPACE SCIENCE	XR	0	0	0				
SPACE SCIENCE	cRef	65745	4040998	3620036	61.46	0.55	11.38	0.39
Multidisciplinary	IR	6	263	235	43.83		23.67	
Multidisciplinary	XR	83	4054	3899	48.84	0.25	26.02	0.25
Multidisciplinary	cRef	58236	2680743	2490038	46.03	0.36	23.71	0.36

General Conclusion

1 General Outlook

This thesis starts by raising the question on the role played by Large research Infrastructure on scientific creativity of its users. More specifically we ask, which is the role that it plays on the collective process of knowledge-creation. There are several motivations behind this choice of topic. RIs constitute an important instrument in European Science Policy with two main goals. The goals expressed by the European Commission for RIs consist first of the will to of conducting cutting-edge research and create a European Research Area that will be a world leader in addressing the global scientific challenges. Secondly, there is an explicit will to make RIs a place for collaboration[65] within Europe and with the rest of the world. There is, therefore, an explicit intention to perform creative research as well as doing it by means of collaboration, which is considered as an important factor for creativity. Yet, the impact evaluation that is done to monitor the performance of these RIs (and other research policies) does not take explicitly creativity into account. Additionally, the study of scientific creativity has recently gained relevance due to the growing concern of science becoming less creative and leading towards conformity.

In such a context we understandably ask whether and how RIs can contribute to scientific creativity and we focus particularly on the creativity of its users. In order to answer our research question, we first search for the organizational conditions and other favourable factors for scientific creativity. We also look for the mechanisms that interact with these favourable conditions, and we finally study the creative results that emerge from the use of RIs.

All along the thesis the main dimensions of the definition of scientific creativity are taken into consideration. Creativity is defined as "the ability to produce knowledge and ideas that are new, original, surprising and useful" (Hollingsworth, 2004; Simonton, 2004).

---

[65] See "*The networks of research infrastructures across Europe strengthen its human capital base by providing world-class training for a new generation of researchers and engineers and promoting interdisciplinary collaboration.*" Extracted from: http://ec.europa.eu/programmes/horizon2020/en/h2020-section/research-infrastructures-including-e-infrastructures and *"Research Infrastructures play a vital role in addressing these challenges. Developing global research infrastructures and reinforcing cooperation of EU research infrastructures at international level contribute to the Open to the World priority set by EU Research Commissioner Carlos Moedas."* Extracted from https://ec.europa.eu/research/infrastructures/index.cfm?pg=international_level

From this definition the ideas of originality and usefulness came up continuously. Usefulness is sometimes described as value or impact while originality is described as novelty. There is an important reason for continuously bringing back this definition. On the one side we find traditional evaluation methods as well as other rewarding systems focusing only on the idea of usefulness and value, which poses several problems to novel research. Because of this bias against novelty several researchers have focused solely on the novelty aspect of creativity. In this thesis we permanently stand on the two aspects when discussing creativity.

Because impact evaluation techniques for the assessment of creativity are missing and it is difficult to quantify creative output, we decided to start this investigation by exploratory research that focuses on "how" rather than "how much". The reason is a well that creativity is a complex concept that is involved at all the stages of the collective knowledge-creation process. This complexity comes with several relationships, dependencies, mechanisms and in general several dynamics. Studying the subject from a qualitative point of view allows to capture the richness of that complexity. We do this research, partly, with qualitative case study investigation. Chapter 2 and Chapter 3 study, each, a different kind of RI.

Chapter 2 focuses on a synchrotron, which is the one of most traditional kinds of RIs: a large instrument, that requires large investment and includes complex and advanced machines. Because of its cost and complexity and also because synchrotrons are needed in a large number of disciplines, most of its users are external users. They go to the synchrotron facilities to perform experiments and come from multiple organizations (private companies, universities, research institutes, etc). We analyse, in Chapter 2, which supportive factors for creativity the synchrotron offers to its users and the creative mechanisms that emerge. Additionally, we identify some creative outputs that occur because of the use of the synchrotron.

Chapter 3 focuses on a completely different type of infrastructure: a database platform. This platform offers access to a very large amount of biological data. These data are needed in multiple areas of research in Life Science and Chemistry, however they are impossible to obtain by a single organisation: the time required for the collection, as well as its cost in terms of computing and storing capacity requires for it to be, as it is the case for synchrotrons, publicly funded. Additionally, because of its digital aspect multiple databases would mean redundancy on efforts and would not be efficient. Chapter 3 analyses the supportive factors for creativity that these databases offer to their users. More particularly it looks at which stages of the combinatory process of creativity benefit from the use of these data. We aim to understand in which ways the use of these databases brings situations that are favourable for creativity, as well as the underlying mechanisms.

Lastly, Chapter 4 contributes to the measurement issue, in order to assess whether an RI can help to produce more creative research? To do so we first analyse the definition of creativity and how this definition applies to science. We look at how other researchers have approached the measurement of creativity and we propose a methodology that is adapted to the specific case of RI and science by bringing together a series of indicators.

2 Main results

The results of this thesis consist of three types as they answer three questions. We first ask which supportive conditions for scientific creativity are found at RIs. This question is faced in chapters 2 and 3. Secondly, we ask which kind of creative output results from the use of RIs. chapters 2 and 3 give first insights on this matter and Chapter 4 uses a quantitative methodology to endorse the results of Chapter 2, for the case of the synchrotron. The third and last question examines whether and how can we approach the measurement of scientific output. More specifically we look for an impact evaluation method of creativity and we develop and propose a methodology.

2.1 Supportive conditions and mechanisms for creativity

When asking about the role of RIs in scientific creativity we split this question into two smaller ones. Literature has studied for years the favourable factors helping creativity to occur and has given large insight. Because of this our first sub question consists in identifying the supportive factors for creativity existing in the context of RI. This issue is analysed in both of the case studies of chapters 2 and 3. We know that communication and variety are crucial for creativity and we seek to observe how these two elements take place at RIs.

In Chapter 2 we see that synchrotrons are used with multiple research purposes and have users coming from multiple disciplines. These users communicate with one another and exchange knowledge. They communicate as well with the beamline scientists. This exchange is often intense and different scientists with different backgrounds discuss together on how to approach some research questions and technological challenges. This is an important factor enabling creativity, as discussions often turn into collaborations where individuals coming from different backgrounds and different organizations pursue together a common goal. Indeed, not only there is an informal exchange of knowledge, there are joint projects and even long-term partnerships as well.

Something very relevant to note here is that all of this is enabled by the beamline scientists. Indeed, all the interviewees expressed how, although the nature of the research infrastructure is helpful for encountering people, what really enables communication is the desire of the beamline scientist to encourage those encounters. Without the beamline scientists encouraging joint projects and including the users into the technological development discussions, most of the encounters, the discussions and specially the collaborations would not have taken place. This means that the existence of an RI is not enough and there is a risk of RIs management forgetting the importance of these human communication and exchanges.

All these factors: variety of backgrounds put together, exchanges of ideas and endorsements into joint projects, were present in the two beamlines we studied. We found however big differences among these beamlines, which suggests that, depending on the state of maturity of the technology proposed by the RI, the impact on creativity will vary. For instance, in the case of the X-ray beamline (which is a mature technology), variety is not always present (we do not systematically observe multidisciplinary teams, or variety of backgrounds for example). Communication comes often in the shape of technological advice and simple support of competences. By contrast, in the case of the Infrared beamline, the challenges are faced jointly through inter-actor collaboration: they work together on defining the objectives, on the ways to reach those objectives and on performing research. From this evidence we learn that, when a technology reaches its maturity, routines establish and it becomes rarer to solve new problems and search for new research questions that would push the knowledge frontier.

In Chapter 3 we look at a different kind of RIs, a digital platform containing access to biological databases and technological tools to exploit these databases. Here, contrary to what happens in the case of the synchrotron, users do not systematically go to physical facility and in general they access to the data from their own institutions[66]. Communication cannot, therefore, come in a traditional manner. We do, however, observe a certain form of exchange between communities and disciplines which has completely change the way knowledge is used in medical research.

Research in biology and medical sciences has been using data for a long time. However, it is only recently that these kinds of databases are used. Before the outstanding development of these large databases, pharmaceutical firms used in-house produced datasets (i.e. data on protein structure or gene expression). Now they use databases that collect the data produced in all the research centres across the world. The amount of data they have access to is thousands of times bigger to the one used previously. These data consist of information on some biological compounds as well as their behaviour and interactions with each other. But farther than the simple quantity, what is relevant here is that the databases include information coming from several different subdisciplines and communities (medical research in cancer, medical research on Parkinson, molecular biology, genetics, etc). In other words, when recombining knowledge, not only knowledge from one's discipline or community is used, there is a large variety of knowledge that is combined. This is typically a favourable condition for creativity. Additionally, what we observe is that, because of the large variety of data that are used, research teams are becoming more varied as well. The variety of data requires a variety of backgrounds to be able to properly exploit them and to make the most out of them.

One of the main contributions of this chapter is to explain what is making it possible to access to such a variety of data. We have found that is precisely the standardization of this data which makes it possible. Not only the data are offered, they are organized in such a way that users can easily find them and understand any information they want. This standardization consists not only in the way knowledge is expressed but also in the language that is used to name, for example, proteins. There was, for a long time, no

---

[66] There are training programmes as well as workshops and Partnership programmes that take place at EBI's facilities. Here, however, we do not focus on that kind of use and we analyse exclusively the remote use of the databases.

standard here, and different disciplines would name the same proteins differently. This made it very difficult for members of a discipline to find the research done in another discipline concerning a specific protein, because he could not know how this other discipline would name the protein. With the existence of standards this does not happen anymore. But it was not with standards alone that the problem was solved. There are language tools (ontologies) that allow researchers who do not know well the standards to query for information using the nomenclature that they know better. An important drawback concerning this access to a wide variety and quantity of knowledge is the potential emergence of lock in trajectories and self-reinforcement in searching processes, a phenomenon that could lead to vicious circles. For example, one can imagine that algorithms and artificial intelligence would continuously select the same kind of compounds or do the same kind of combinations if they are based on what is more popular or more promising. Additionally, some kinds of data might be more propense to find results and the fact that all researchers have access to the same data (and same search algorithms) might concentrate efforts on those that are the most productive, thus ignoring more risky paths of research.

Another relevant result consists on the observation of a mechanism for serendipity. When doing the literature review on serendipity we found that it is the accumulation of knowledge as well as the existence of networks that allow for serendipitous episodes. In a first place, simply by cumulating a large amount of knowledge, an individual can perceive the irregularities that lead to discoveries because only by knowing what we should expect we can identify an unexpected event. In a second place, the existence of networks of scientists allow for anomalous observations or results to be diffused and it is more likely that someone comes up with an idea. Or, in the same line of ideas, the existence of such a network is a prerequisite for an anomaly to be noticed, since the network allows for a general view on which results are normal. What databases offer follows a similar logic, although the mechanism in play is different. By having large amounts of data assembled together in a single database everyone has access to a whole picture of the knowledge produced in a certain domain. Because all the researchers have access to all the information instead of having the information split on smaller groups and each looking at one part of it, the chances of making connections that will lead to new and valuable knowledge are bigger. This idea, which emerged from the interviews, is not without criticism or drawbacks. One can easily imagine that when the amount of information faced is very large it might become difficult to make some sense out of it.

In summary, the potentially positive factors for creativity within RIs consist in their ability to enable communication between different disciplines and backgrounds. This is done at the synchrotron by means of multidisciplinary joint research as well as informal communication. For the database platform, this is achieved by giving access to knowledge coming from multiple disciplines. In the case of the Synchrotron an additional finding is that communities and leaders ensure the existence of these communication. Concerning the biological database platform, we found the existence of favourable conditions in terms of serendipity, which is an additional way for creative results to occur. An important contribution of this thesis is the creation of two conceptual models, one for each infrastructure, that represent all the factors that contribute to scientific creativity and how do they relate to each other.

2.2 Types of creative outcome

When asking which kinds of creative output emerge from the use of these RIs it is important to keep in mind that the output here is science production. For the case of the synchrotron there is evidence of two kinds of creative output. On the one side we have technological development. Technological development is an output that is happening continuously at the synchrotron as a result of joint projects of users with synchrotron scientists and/or projects with several synchrotrons working together. This is particularly true for the Infrared beamline where new applications and methodologies are continuously being developed together with the development of new machines to perform new kinds of experiments. Additionally, there is the creation of knowledge by itself. Several interviewed people said that the publications that came from the use of SOLEIL were among their best ones and that they often used SOLEIL to address the most complicated problems they had to address.

Another relevant result in terms of creative output of the synchrotron is the emergence of new communities. Communities play a role at every stage of the production process at the synchrotron as they enable some of the collaborations and communication channels. However, the most important discovery about communities in this chapter is the possibility for a new community to emerge around the use of the synchrotron. We observed several elements indicating the emergence of a new community of scientists: ongoing communication, trust, continuous joint projects and pursue of the same general scientific goals. This is probably the most relevant result of Chapter 2. This community is on an emerging phase, and it might not establish completely. We think, however, that it is important to follow up on this issue.

Concerning EBI, the interviewees agreed all on an important increase of identified targets, insisting more particularly on an increasing the rate between the hit generation and the target identification[67]. This result has to be qualified: at this upstream step in the R&D process, we are however still far from a final product. What is very relevant as a creative result for the case of EBI is the opening of new research paths. We observed that, because the databases exist and because some elements can now be put together, some new research questions are opened as well as research paths. Interviewees have particularly pointed at the endorsement into more difficult problems, ie. problems that had been neglected for years in favour of incremental research.

To sum up, the kind of creative results vary a lot from an RI to another. In the case of the synchrotron it consists mainly in novel good quality publications and technological development. In the case of EBI, it consists mainly in the discovery of new targets. When it comes to unexpected and very relevant outputs for our research question, the synchrotron case shows the emergence of a new community of scientists, which could mean the opening of a completely new research field. Similarly, EBI databases show the emergence of new research questions, new scientific challenges and in the end new research paths.

Finally, looking at Chapter 4, we found that concerning scientific publications and according to our indicator, the articles due to SOLEIL are more creative than the rest of the publications in the same field. The validity of this results, however, depends on the indicators used in our methodology, which are widely accepted indicators but not without criticism. This validity problem is discussed in next section. Indeed, we use indicators that are widely accepted and use but also very often criticized. The count of citations and

---

[67] The drug discovery process generally follows the following path that includes a hit to lead stage: target validation (TV), assay development, high-throughput screening, hit to lead (H2L), lead optimization (LO), preclinical drug development and clinical drug development

the journal impact factor are often criticized by not representing well enough value. The novelty indicator that we use is relatively young and although it has been well accepted it might have some flaws.

2.3 Methodology for evaluation of impact on creativity

The last chapter of this thesis offers a methodological contribution. Indeed we propose a methodology for the evaluation of creativity of research realised at the RI. This methodology allows us to determine how creative a sample of research articles is, compared to a universe consisting in the total of the research papers in the same scientific domain. This methodology stands on the idea of considering creativity as the co-existence of two attributes: value and novelty. Value has traditionally been evaluated through impact of publications. Novelty has only recently started to be evaluated and this is done by combinatory novelty of cited journals. We have adopted these notions and developed a methodology on two parts, one evaluating impact and the other evaluating novelty. For impact we considered the journal impact factor as well as the number of citations. For novelty we looked first at rarely cited journals and secondly at rare combinations of cited journals.

Let us look at the drawbacks of this methodology. In what concerns impact, although the indicators used are very common and accepted, they are also criticized. Several alternatives have been developed such as the “Eigenfactor” ranking and Article Influence Score, as explained in Chapter 4. In what concerns novelty, our indicator lies on several ideas or presuppositions. First, we considered that journals are representative of a topic and a community and different journals would represent different topics and communities. This assumption is not necessarily true for every single journal. Secondly we assumed that citing a certain journal is representative of using the knowledge coming from that journal in order to build new knowledge, which is again, not necessarily true for every single case. Lastly, when using the indicator for impact of RI, we observed a correlation between our indicator and the use of the RI, but correlation does not mean causality. Another explanation could be that the synchrotron accepts only very creative projects. The interviews conducted for the case study of Chapter 2 appear necessary to complement the quantitative approach and to confirm the impact of the large instrument on the creativity of tis users.

A solution to these drawbacks would be to follow the same steps but to use alternative indicators in order to check whether the results change or not. In what concerns value the Eigenfactor could be a good alternative. There are however other alternatives that should be explored. Considering the validity of the novelty indicator, one alternative would be to use text analysis tools to analyse novelty in texts. This is probably complicated to calculate and requires a lot of computing capacity. Another option that should probably be explored is the use of statistics on the combination of keywords, rather than the combination of cited journals. Another alternative option is to do survey studies among the users od the RIs and ask them directly about the papers that our indicators have considered as creative. We would ask them if they consider that that article is particularly creative or not. This, of course, leaves a big place to subjectivity. Finally, an important drawback of this methodology is the need to be able to identify scientific publications that have used an RI and that is often not possible. Sometimes RIs are mentioned in the acknowledgements section but this is still not something well-established and this information is often missing.

3 Policy and managerial implications

3.1 Managerial implications

Although managers have known for a long time the importance of mixing backgrounds and adding variety to teams, this is not always encouraged in science. Although multidisciplinary research and collaboration between different institutions is the more and more rewarded, research is still built up around disciplines rather than research questions and scientific challenges.

Scientific management should, therefore, overcome the barriers of disciplines and engage strongly in multidisciplinary research that is built around a research question and not a scientific discipline. This is something that is already encouraged officially by European Policy Makers as well as national ones (Horizon 2020 for example). However real transdisciplinary research, where questions are faced together and there is a joint work is still rare today. The ways to encourage a kind of interaction that consists on working together, rather than on a division of work based on competences and knowledge, should be studied.

Another important managerial implication would involve the recruitment and career evolution incentives. Literature shows, and this is somewhat supported by our quantitative empirical analysis, that creative science is not always rewarded with publications. Because creative science is very new, it has very little to build up on and often does not fill in the requirements in order to be published in a good journal.

We have observed as well, for the case of the synchrotron, that although the favourable conditions for creativity exist by themselves they only act positively on creativity because of the strong involvement of the management. Leaders should not content themselves with building a space that is shared by disciplines and communities. They should be continuously involved in order to assure encounters, build trust and enable collaboration and joint projects. Decision makers (i.e. university or lab directors or firm managers) should always keep this in mind and assure the right organizational structure that allows team and department leaders to be invested on the research that is taking place.

3.2 Policy Implications

RIs are built with the main goal of advancing research in difficult but essential areas and being able to face challenging scientific problems. This requires creativity and we have observed in this thesis that RI programmes are going in the right directions. A small problem relates to the fact that mature communities tend to acquire routines and their way of producing knowledge may become slightly standardized. This discourages engaging into new problems and trying to develop new technologies. For this reason, it would be positive, not only to encourage new disciplines but as well to promote creative behaviours in mature ones. Some solutions are, for instance, the promotion of innovative and risky projects within this mature communities, or the encouragement to make them participate in interdisciplinary projects or alliances.

When it comes to the evaluation methodology proposed by this thesis, it depends strongly on the possibility to link scientific publications to an RI that has been used for that particular scientific publications. This does not always happen systematically. Policy makers should, if they desire to be able to monitor the impact of Large RIs on creativity, ensure that RIs are properly acknowledged in the publications that used them. Additionally, they should be able to identify and follow up on some other creative results and find alternatives to publications for those contexts in which publications are not an option (such as the case of the pharmaceutical industry)

4 Limits and perspectives

The aim of this thesis was to find whether and how RIs contributed to scientific creativity. We could expect, by doing a literature review, that variety would play a crucial role as RIs are places that can often be used with multiple proposes and by multiple disciplines, and therefore people with a variety of backgrounds. Additionally, performing creative research is the reason why RIs are created, to be able to face the most challenging scientific questions. We have found that RIs offer several positive conditions for creativity as well as some creative results.

In what concerns the general limits of this thesis, the main one consists on its external validity. To what extent our results can be generalised? Is our measurement methodology relevant in other RI cases or more generally other research institutions? RIs have several elements in common. They have, however, big differences as well. There is a very large variety of RIs, each one holding particular interactions with the users. They might have, as well, different supportive conditions for creativity. When it comes to Chapter 4 there are 2 main limitations. On one side the indicators used have alternatives that should be tested in order to confirm the results. Additionally, surveys or other kinds of investigation should be performed to confirm the results. An idea would be to ask scientists themselves which are their most creative articles, and check whether it fits the “score” given by our indicator.

Further research should focus on the study of creativity in different kinds of RIs. RIs take a large variety of forms and because of this the proper situations for creativity that can appear may vary a lot from one another. For this reason, a large study of creativity across different kinds of RI would allow for a mapping of types of RIs and types of positive factors for creativity. This would allow managers and policy makers to understand which their organisational assets are, where is creativity taking place and if it is not, how to improve that situation.

Additionally, to the extension of our quantitative study to other RIs should be done. To do this, there is the need to be able to identify the publications associated to the use of an IR, this is, today, rarely possible so the way to manage it should be investigated as well. A problem with this quantitative methodology is that it concerns four different indicators. We have, in our analysis, had results that went mostly on the same direction. A question that remains open is, what do we do when the different indicators point in different directions? How do we approach this situation? This is a question that must be addressed in the future of we want that indicator to be

5 Concluding Remarks

The objective of this thesis has been to answer the question: what is the impact of large research infrastructure on scientific creativity? Which is the impact specifically in the creativity of its users? This question is relevant for many reasons. In first place, scientific advance requires creativity in order to face big societal challenges. Incremental research, by itself, is not enough. This need for creativity is particularly true in the case of Large RIs because doing creative science is the very reason they have been built for: to solve the big scientific challenges and to push the knowledge

This thesis has provided, by means of case study research, with some relevant insights on the kinds of supportive factors for creativity that are present in RIs. RIs provide with the cutting-edge technology and additionally with a place where an heterogeneity of disciplines and backgrounds encounter. This heterogeneity acts through different mechanisms in order to enable creativity. RIs are a favourable place for creative research and how this creativity occurs should continue to be studied in order to be able to preserve it and maintain the mechanisms that make it possible. Creativity is a precious resource and scientific advancement (and as a consequence societal and technological advancement) depend on it.

One last lesson that we can learn from this thesis is that the existence and access to a variety of backgrounds can bring back some of the original hazardous component that science. In the synchrotron we go with the idea of performing an experiment and we might end up endorsing into a collaboration for a topic that we never thought we would be working on. This is related to the idea of serendipity found in Chapter 3. Facing big quantities of varied knowledge and being able to understand it can lead to surprising hazardous discoveries. For this to happen science has to keep going towards a logic of openness and sharing as it is the availability of this big variety of knowledge and the possibility to see it that allows people to come up with ideas and make connections.

References

Adams, J., Black, G.C., Clemmons, J.R., Stephan, P., 2005. Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999. Research Policy 34, 259–285.

Akrich, M., Callon, M., Latour, B., 1988. A quoi tient le succès des innovations? 1: L'art de l'intéressement; 2: Le choix des porte-parole 43.

Alves, J., Marques, M.J., Saur, I., Marques, P., 2007. Creativity and Innovation through Multidisciplinary and Multisectoral Cooperation. Creativity and Innovation Management 16, 27–34. https://doi.org/10.1111/j.1467-8691.2007.00417.x

Amabile, T.M., 2011. The Progress Principle: Using Small Wins to Ignite Joy, Engagement, and Creativity at Work. Harvard Business Review Press.

Amabile, T.M., 1997. Motivating Creativity in Organizations: On Doing What You Love and Loving What You Do.

Amabile, T.M., 1996. Creativity and Innovation in Organizations.

Amabile, T.M., 1988. A Model of Creativity and Innovation in Organizations.

Amabile, T.M., 1983. The Social Psychology of Creativity: A Componential Conceptualization.

Amabile, T.M., 1982. The Social Psychology of Creativity: A Consensual Assessment Technique.

André, F., 2014. Chapitre 5. Déluge des données de la recherche ? Information et strategie 77–95.

Aragon, C.R., Poon, S.S., Monroy-Hernández, A., Aragon, D., 2009. A Tale of Two Online Communities: Fostering Collaboration and Creativity in Scientists and Children, in: Proceedings of the Seventh ACM Conference on Creativity and Cognition, C&C '09. ACM, New York, NY, USA, pp. 9–18. https://doi.org/10.1145/1640233.1640239

Arrow, K., 1962. Economic Welfare and the Allocation of Resources for Invention (NBER Chapters). National Bureau of Economic Research, Inc.

Arthur, W.B., 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. The Economic Journal 99, 116–131. https://doi.org/10.2307/2234208

Attwood, T.K., Gisel, A., Eriksson, N.-E., Bongcam-Rudloff, E., 2011. Concepts, Historical Milestones and the Central Place of Bioinformatics in Modern Biology: A European Perspective, in: A. Mahdavi, M. (Ed.), Bioinformatics - Trends and Methodologies. InTech. https://doi.org/10.5772/23535

Audretsch, D.B., Feldman, M.P., 1996. R&D Spillovers and the Geography of Innovation and Production. American Economic Review 86, 630–640.

Avadikyan, A., Bach, L., Lambert, G., Lerch, C., Wolff, S., 2014. Dynamique des modèles d'affaires et écosystème : le cas des synchrotrons. Revue d'économie industrielle 153–186. https://doi.org/10.4000/rei.5822

Avadikyan, A., Müller, M., 2017. AVADIKYAN Arman, MÜLLER Moritz (2017) "Management of creativity in a large scale research facility" in The Global Management of Creativity (Eds) WAGNER Marcus, VALLS-PASOLA Jaume, BURGER-HELMCHEN Thierry, pp. 140-158, Routledge, NY.

Bach, L., Matt, M., Wolff, S., 2014. How do firms perceive policy rationales behind the variety of instruments supporting collaborative R&D? Lessons from the European Framework Programs. Technovation 34, 327–337. https://doi.org/10.1016/j.technovation.2014.02.008

Bach, L., Wolff, S., 2017. The BETA-EvaRIO impact evaluation method: towards a bridging approach? The Journal of Technology Transfer. https://doi.org/10.1007/s10961-017-9603-y

Barber, B., 1961. Resistance by Scientists to Scientific Discovery. Science 134, 596–602.

Barron, F., 2012. Creativity And Psychological Health: Origins Of Personal Vitality And Creative Freedom. Literary Licensing, LLC.

Batallas, D.A., Yassine, A.A., 2006. Information Leaders in Product Development Organizational Networks: Social Network Analysis of the Design Structure Matrix. IEEE Transactions on Engineering Management 53, 570–582. https://doi.org/10.1109/TEM.2006.883706

Bernward, J., Terry, S., 2001. CHAPETR 1: A fresh outlook at instrumentation: An introduction, in: Instrumentation Between Science, State and Industry. Kluwer Academic Publishers, Dordrecht, Boston,London, pp. 1–13.

Besnier, J.-M., 2016. Les théories de la connaissance: « Que sais-je ? ». Presses Universitaires de France.

Bolser, D.M., Chibon, P.-Y., Palopoli, N., Gong, S., Jacob, D., Del Angel, V.D., Swan, D., Bassi, S., González, V., Suravajhala, P., Hwang, S., Romano, P., Edwards, R., Bishop, B., Eargle, J., Shtatland, T., Provart, N.J., Clements, D., Renfro, D.P., Bhak, D., Bhak, J., 2012. MetaBase--the wiki-database of biological databases. Nucleic Acids Res. 40, D1250-1254. https://doi.org/10.1093/nar/gkr1099

Bozeman, B., Boardman, C., 2014. Research Collaboration and Team Science: A State-of-the-Art Review and Agenda. Springer.

Brown, J.S., Duguid, P., 1991. Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation. Organization Science 2, 40–57.

Burt, R.S., 2004. Structural Holes and Good Ideas. American Journal of Sociology 110, 349–399. https://doi.org/10.1086/421787

Carayol, N., Llopis, O., Lahate, A., 2018. The Right Job and the Job Right: Novelty, Impact and Journal Stratification in Science. Paper to be presented at DRUID18 Copenhagen Business School, Copenhagen, Denmark June 11-13, 2018.

Chamorro-Premuzic, T., 2017. Does Diversity Actually Increase Creativity? [WWW Document]. Harvard Business Review. URL https://hbr.org/2017/06/does-diversity-actually-increase-creativity (accessed 10.6.18).

Charriere Petit, S., Durieux, F., 2007. Explorer et tester: les deux voies de la recherche., in: Méthodes de Recherche En Management. Thiétart R.-A. et Al.

Cohen, W.M., Levinthal, D.A., 1990a. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35, 128–152. https://doi.org/10.2307/2393553

Cohen, W.M., Levinthal, D.A., 1990b. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35, 128. https://doi.org/10.2307/2393553

Cohendet, P., 2016. Arthur Koestler - Aux origines de l'acte créatif : la bisociation. in Les grands auteurs en management de l'innovation et de la créativité.

Cohendet, P., Diani, M., 2003. L'organisation comme une communauté de communautés croyances collectives et culture d'entreprise. Revue d'économie politique 113, 697–720. https://doi.org/10.3917/redp.135.0697

Cohendet, P., Harvey, J.-F., Simon, L., Harvey, J.-F., Simon, L., 2013. Managing creativity in the fi rm: the fuzzy front end of innovation and dynamic capabilities PAT RICK COHEN DET, J EAN-FR ANÇOIS H A RV EY : The fuzzy front end of innovation and dynamic capabilities [WWW Document]. The Economics of Creativity. https://doi.org/10.4324/9780203073414-17

Collis, J., Hussey, R., 2003. Business Research. Palgrave Macmillan.

Cowan, R., David, P.A., Foray, D., n.d. The Explicit Economics of Knowledge Codification and Tacitness 32.

Crane, D., 1969. Social Structure in a Group of Scientists: A Test of the "Invisible College" Hypothesis on JSTOR. Americal Sociological Review 34, 335–352.

Cronin, B., 1982. Invisible colleges and information transfer a review and commentary with particular reference to the social sciences. Journal of Documentation 38, 212–236. https://doi.org/10.1108/eb026730

Csikszentmihalyi, M., 2014. Society, Culture, and Person: A Systems View of Creativity, in: Csikszentmihalyi, M. (Ed.), The Systems Model of Creativity: The Collected Works of Mihaly Csikszentmihalyi. Springer Netherlands, Dordrecht, pp. 47–61. https://doi.org/10.1007/978-94-017-9085-7_4

Csikszentmihalyi, M., montuori, A., Purser, R.E., 1995. Deconstructing the Lone Genius Myth: Toward a Contextual View of Creativity 35, 69–112.

David, P., 1985. Clio and the Economics of QWERTY. American Economic Review 75, 332–37.

De Solla Price, D.J., Beaver, D., 1966. Collaboration in an invisible college. American Psychologist 21, 1011–1018. https://doi.org/10.1037/h0024051

Donovan, C., 2011. State of the art in assessing research impact: introduction to a special issue. Res Eval 20, 175–179. https://doi.org/10.3152/095820211X13118583635918

Donovan, C., 2008. The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research - Donovan - 2008 - New Directions for Evaluation - Wiley Online Library [WWW Document]. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ev.260 (accessed 8.30.18).

Eisenhardt, K.M., 1989. Building Theories from Case Study Research.

Fleming, L., 2001. Recombinant Uncertainty in Technological Search. Management Science 47, 117–132. https://doi.org/10.1287/mnsc.47.1.117.10671

Ford, C.M., 1996. A theory of individual creative action in multiple social domains. The Academy of Management Review 21, 1112–1142. https://doi.org/10.2307/259166

Freeman, C., 1987. Technology Policy and Economic Performance: Lessons from Japan. Printer Publishers.

Freeman, C., Clark, J., Soete, L., 1982. Unemployment and technical innovation: a study of long waves and economic development. Frances Pinter.

Garfield, E., 1977. Essays of an information scientist. ISI Press.

Garud, R., Tuertscher, P., Van de Ven, A.H., 2013. Perspectives on Innovation Processes. The Academy of Management Annals 7, 775–819. https://doi.org/10.1080/19416520.2013.791066

Gavard-Perret, M., Gotteland, D., Haon, C., Jolibert, A., 2012. Methodologie de la recherche - Marie-Laure Gavard-Perret, David Gotteland, Christophe Haon, Alain Jolibert - Livres. Pearson.

Gibbons, M., 2000. Mode 2 society and the emergence of context-sensitive science. Sci Public Policy 27, 159–163. https://doi.org/10.3152/147154300781782011

Gioia, D.A., Corley, K.G., Hamilton, A.L., 2013. Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. Organizational Research Methods 16, 15–31. https://doi.org/10.1177/1094428112452151

Giordano, Y., 2003. Conduire un projet de recherche. Une perspective qualitative (No. halshs-00440011), Post-Print. HAL.

Gong, S., Worth, C.L., Cheng, T.M.K., Blundell, T.L., 2011. Meet me halfway: when genomics meets structural bioinformatics. J Cardiovasc Transl Res 4, 281–303. https://doi.org/10.1007/s12265-011-9259-1

Gruber, H.E., Barrett, P.H., 1974. Darwin on man: A psychological study of scientific creativity, Darwin on man: A psychological study of scientific creativity. E. P. Dutton, New York, NY, England.

Guilford, J.P., 1967. The Nature of Human Intelligence. McGraw-Hill, NY.

Guilford, J.P., 1950. Creativity. American Psychologist 5, 44–54.

Haas, P.M., 1992. Introduction: Epistemic Communities and International Policy Coordination. International Organization 46, 1–35.

Hage, J., Mote, J., 2010. Transformational Organizations and a Burst of Scientific Breakthroughs: The Institut Pasteur and Biomedicine, 1889-1919. Social Science History 34, 13–46.

Hagtvedt, L.P., Dossinger, K., Harrison, S.H., Huang, L., 2019. Curiosity made the cat more creative: Specific curiosity as a driver of creativity. Organizational Behavior and Human Decision Processes 150, 1–13. https://doi.org/10.1016/j.obhdp.2018.10.007

Hanney, S.R., González-Block, M.A., 2011. Yes, research can inform health policy; but can we bridge the "Do-Knowing It's Been Done" gap? Health Res Policy Syst 9, 23. https://doi.org/10.1186/1478-4505-9-23

Harvey, J.-F., Cohendet, P., Simon, L., Borzillo, S., 2015. Knowing Communities in the Front End of Innovation. Research-Technology Management 58, 46–54. https://doi.org/10.5437/08956308X5801198

Hayes, J.R., 1989. Cognitive Processes in Creativity, in: Glover, J.A., Ronning, R.R., Reynolds, C.R. (Eds.), Handbook of Creativity, Perspectives on Individual Differences. Springer US, Boston, MA, pp. 135–145. https://doi.org/10.1007/978-1-4757-5356-1_7

He, Z.-L., Wong, P.-K., 2004. Exploration vs. Exploitation: An Empirical Test of the Ambidexterity Hypothesis. Organization Science 15, 481–494. https://doi.org/10.1287/orsc.1040.0078

Heinze, T., Bauer, G., 2007. Characterizing creative scientists in nano-S&T: Productivity, multidisciplinarity, and network brokerage in a longitudinal perspective. Scientometrics 70, 811–830. https://doi.org/10.1007/s11192-007-0313-3

Heinze, T., Shapira, P., Rogers, J.D., Senker, J.M., 2009. Organizational and institutional influences on creativity in scientific research. Research Policy, Special Issue: Emerging Challenges for Science, Technology and Innovation Policy Research: A Reflexive Overview 38, 610–623. https://doi.org/10.1016/j.respol.2009.01.014

Hey, T., Tansley, S., Tolle, K., 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery.

Hienerth, C., von Hippel, E., Berg Jensen, M., 2014. User community vs. producer innovation development efficiency: A first empirical study. Research Policy 43, 190–201. https://doi.org/10.1016/j.respol.2013.07.010

Hollingsworth, J.R., 2002. Research organizations and major discoveries in twentieth-century science: a case study of excellence in biomedical research.

Hollingsworth, R, 2004. .Institutionalizing excellence in biomedical research: the case of Rockefeller University., in: Creating a Tradition of Biomedical Research. Contributions to the History of the Rockefeller University. Rockefeller University Press,.

Ilbeigi, A., 2017. Public R&D Policy Impact Evaluation:Propensity Score Matching and Structural Modeling Estimations (phd). University of Trento.

Karwowski, M., Dul, J., Jauk, E., Jankowska, D.M., Gaddja, A., 2016. M. Karwowski, J. Dul, J. Gralewski, E. Jauk, D.M. Jankowska, A. Gajda, et al. Is creativity without intelligence possible? A necessary condition analysis. Intelligence.

Klahr, D., Simon, H.A., 1999. Studies of scientific discovery: Complementary approaches and convergent findings. Psychological Bulletin 125, 524–543. https://doi.org/10.1037/0033-2909.125.5.524

Koestler, A., 1964. The Act of Creation. Arkana.

Kogut, B., Zander, U., 1996. What Firms Do? Coordination, Identity, and Learning. Organization Science 7, 502–518. https://doi.org/10.1287/orsc.7.5.502

Koskinen, K.U., Pihlanto, P., Vanharanta, H., 2003. Tacit knowledge acquisition and sharing in a project work context. International Journal of Project Management 21, 281–290. https://doi.org/10.1016/S0263-7863(02)00030-3

Krammer, S., 2015. Do good institutions enhance the effect of technological spillovers on productivity? Comparative evidence from developed and transition economies. Technol. Forecast. Soc. Chang. 94, 133–154.

Kuhn, T.S., 1970. The structure of scientific revolutions, [2d ed., enl. ed, International encyclopedia of unified science. Foundations of the unity of science, v. 2, no. 2. University of Chicago Press, Chicago.

Landau, R., Rosenberg, N., Engineering, N.A. of, 1986. The Positive Sum Strategy: Harnessing Technology for Economic Growth. National Academies.

Latour, B., 1999. On Recalling Ant. The Sociological Review 47, 15–25. https://doi.org/10.1111/j.1467-954X.1999.tb03480.x

Latour, B., Woolgar, S., Salk, J., 1979. Laboratory Life: The Construction of Scientific Facts., 2nd edition. ed. Princeton University Press, Princeton, N.J.

Lee, Y.-N., Walsh, J.P., Wang, J., 2015. Creativity in scientific teams: Unpacking novelty and impact. Research Policy 44, 684–697.

Lesser, E.L., Storck, J., 2001. Communities of practice and organizational performance. IBM Systems Journal 40, 831–841. https://doi.org/10.1147/sj.404.0831

Li, E.Y., Liao, C.H., Yen, H.R., 2013. Co-authorship networks and research impact: A social capital perspective. Research Policy 42, 1515–1530. https://doi.org/10.1016/j.respol.2013.06.012

Li, H., Li, F., Chen, T., 2018. A motivational–cognitive model of creativity and the role of autonomy. Journal of Business Research 92, 179–188. https://doi.org/10.1016/j.jbusres.2018.07.025

Li, Q., Maggitti, P.G., Smith, K.G., Tesluk, P.E., Katila, R., 2013. Top Management Attention to Innovation: The Role of Search Selection and Intensity in New Product Introductions. Academy of Management Journal 56, 893–916. https://doi.org/10.5465/amj.2010.0844

Lundvall, B.-Å., 1992. National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. Pinter Publishers.

March, J.G., 1995. The Future, Disposable Organizations and the Rigidities of Imagination.

March, J.G., 1991. March, J.G.: Exploration and exploitation in organizational learning. Organ. Sci. 2, 71-87. Organisational Science 2, 71–87.

March, J.G., Levitt, barbara, 1988. Organizational Learning. Annual Review of Sociology 14, 19–40.

Martin, B.R., 2012. The evolution of science policy and innovation studies. Research Policy, Exploring the Emerging Knowledge Base of “The Knowledge Society” 41, 1219–1239. https://doi.org/10.1016/j.respol.2012.03.012

McLeod, P.L., Lobel, S.A., Cox, T.H., 1996. Ethnic Diversity and Creativity in Small Groups. Small Group Research 27, 248–264. https://doi.org/10.1177/1046496496272003

McLure Wasko, M., Faraj, S., 2000. “It is what one does”: why people participate and help others in electronic communities of practice. The Journal of Strategic Information Systems 9, 155–173. https://doi.org/10.1016/S0963-8687(00)00045-7

Mednick, S., 1962. The associative basis of the creative process. Psychological Review 69, 220–232. https://doi.org/10.1037/h0048850

Merton, R., 1975. The Sociology of Science. Theoretical and empirical investigations. University of Chicago Press.

Merton, R., Barber, B., 2004. The Travels and Adventures of Serendipity.

Miles, M.B., Huberman, A.M., 2003. Analyse des données qualitatives. De Boeck Supérieur.

Minbaeva, D., Pedersen, T., Björkman, I., Fey, C.F., Park, H.J., 2003. MNC knowledge transfer, subsidiary absorptive capacity, and HRM. J Int Bus Stud 34, 586–599. https://doi.org/10.1057/palgrave.jibs.8400056

Mumford, M.D., Scott, G.M., Gaddis, B., Strange, J.M., 2002. Leading creative people: Orchestrating expertise and relationships. The Leadership Quarterly 13, 705–750. https://doi.org/10.1016/S1048-9843(02)00158-3

Munos, B., 2009. Lessons from 60 years of pharmaceutical innovation. Nature Reviews Drug Discovery 8, 959–968. https://doi.org/10.1038/nrd2961

Murayama, K., Nirei, M., Shimizu, H., 2015. Management of science, serendipity, and research performance: Evidence from a survey of scientists in Japan and the U.S. Research Policy 44, 862–873. https://doi.org/10.1016/j.respol.2015.01.018

Nelson, R.R., 1993. National Innovation Systems: a comparative analysis. Oxford University Press.

Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Cambridge MA and London: Harvard University Press.

Neumann, C.J., 2007. Fostering creativity. A model for developing a culture of collective creativity in science. EMBO Rep 8, 202–206. https://doi.org/10.1038/sj.embor.7400913

Nonaka, I., 2000. A firm as a knowledge-creating entity: a new perspective on the theory of the firm. Industrial and Corporate Change 9, 1–20. https://doi.org/10.1093/icc/9.1.1

Nooteboom, B., 2000. Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. Journal of Management & Governance 4, 69–92.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., van den Oord, A., 2007. Optimal cognitive distance and absorptive capacity. Research Policy 36, 1016–1034. https://doi.org/10.1016/j.respol.2007.04.003

O’Mahony, S., Lakhani, K.R., 2011. Organizations in the Shadow of Communities, in: Communities and Organizations, Research in the Sociology of Organizations. Emerald Group Publishing Limited, pp. 3–36. https://doi.org/10.1108/S0733-558X(2011)0000033004

Orsenigo, L., Malerba, F., 2015. The evolution of the pharmaceutical industry: Business History. Business History 57, 664–687.

Patel, P., Pavitt, K., 1991. Large Firms in the Production of the World’s Technology: An Important Case of “Non-Globalisation.” J Int Bus Stud 22, 1–21. https://doi.org/10.1057/palgrave.jibs.8490289

Paulus, P.B., Nijstad, B.A., 2003. Group Creativity: Innovation through Collaboration. Oxford University Press.

Pavitt, K., 1998. Technologies, Products and Organization in the Innovating Firm: What Adam Smith Tells Us and Joseph Schumpeter Doesn't. Ind Corp Change 7, 433–452. https://doi.org/10.1093/icc/7.3.433

Pelz, D.C., Andrews, F.M., 1976. Scientists in organizations: productive climates for research and development. Institute for Social Research, University of Michigan.

Pelz, D.C., Andrews, F.M., 1966. Scientists in Organizations: Productive Climates for Research and Development. Wiley.

Powell, W.W., Koput, K.W., Smith-Doerr, L., 1996. Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology 41, 116–145.

Rispal, M.H., 2002. La méthode des cas. De Boeck Supérieur. https://doi.org/10.3917/dbu.hlady.2002.01

Rodan, S., Galunic, C., 2004. More than Network Structure: How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness. Strategic Management Journal 25, 541–562.

Romer, P.M., 1994. The Origins of Endogenous Growth. Journal of Economic Perspectives 8, 3–22. https://doi.org/10.1257/jep.8.1.3

Rosenberg, N., 2009. Endogenous Forces in Twentieth-Century America (World Scientific Book Chapters). World Scientific Publishing Co. Pte. Ltd.

Rosenberg, N., 1992. Scientific instrumentation and university research. Research Policy 21, 381–390. https://doi.org/10.1016/0048-7333(92)90035-3

Rosenberg, N., 1982. Chapter 7: How exogeneous is Science ?, in: Inside the Black Box: Technology and Economcis. Cambridge University Press.

Rothwell, R., Freeman, C., Horlsey, A., Jervis, V.T.P., Robertson, A.B., Townsend, J., 1974. SAPPHO updated - project SAPPHO phase II. Research Policy 3, 258–291.

Schilling, M., Green, E., 2011. Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. Research Policy 40, 1321–1331. https://doi.org/10.1016/j.respol.2011.06.009

Schumpeter, J.., 1934. The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle, New Brunswick (U.S.A) and London (U.K.): Transaction Publishers.

Shi, B., Wang, L., Yang, J., Zhang, M., Xu, L., 2017. Relationship between Divergent Thinking and Intelligence: An Empirical Study of the Threshold Hypothesis with Chinese Children. Front Psychol 8. https://doi.org/10.3389/fpsyg.2017.00254

Shin, S.Y., 2016. Recent update in HIV vaccine development. Clin Exp Vaccine Res 5, 6–11. https://doi.org/10.7774/cevr.2016.5.1.6

Simon, H.A., 1977. Scientific Discovery and the Psychology of Problem Solving, in: Simon, H.A. (Ed.), Models of Discovery: And Other Topics in the Methods of Science, Boston Studies in the Philosophy of Science. Springer Netherlands, Dordrecht, pp. 286–303. https://doi.org/10.1007/978-94-010-9521-1_16

Simon, H.A., Langley, P.W., Bradshaw, G.L., 1981. Scientific Discovery as Problem Solving. Synthese 47, 1–27.

Simonton, D.K., 2004. Creativity in Science: Chance, Logic, Genius, and Zeitgeist. Cambridge University Press.

Smith, E.A., 2001. The role of tacit and explicit knowledge in the workplace. J of Knowledge Management 5, 311–321. https://doi.org/10.1108/13673270110411733

Soler, L., 2000. Introduction à l'épistémologie. Ellipses.

Spier, R., 2002. The history of the peer-review process. Trends in Biotechnology 20, 357–358. https://doi.org/10.1016/S0167-7799(02)01985-6

Stegmann, J., 1997. How to evaluate journal impact factors [WWW Document]. Nature. https://doi.org/10.1038/37463

Stent, G.S., 1973. Prematurity and uniqueness in scientific discovery. -. Scientific American 84–93.

Stephan, P., 2012. How Economics Shapes Science — Paula Stephan | Harvard University Press [WWW Document]. URL http://www.hup.harvard.edu/catalog.php?isbn=9780674088160&content=reviews (accessed 8.5.18).

Takeuchi, H., Nonaka, I., 1995. The New Dynamism of the Knowledge-Creating Company 10.

Tobias, P.V., 2009. Homo habilis—A Premature Discovery: Remembered by One of Its Founding Fathers, 42 Years Later, in: Grine, F.E., Fleagle, J.G., Leakey, R.E. (Eds.), The First Humans – Origin and Early Evolution of the Genus Homo: Contributions from the Third Stony Brook Human Evolution Symposium and Workshop October 3 – October 7, 2006, Vertebrate Paleobiology and Paleoanthropology. Springer Netherlands, Dordrecht, pp. 7–15. https://doi.org/10.1007/978-1-4020-9980-9_2

Todorova, G., Durisin, B., 2007. Absorptive capacity: Valuing a reconceptualization. AMR 32, 774–786. https://doi.org/10.5465/amr.2007.25275513

Torrance, E.., 1966. The Torrance tests of creative thinking-norms-technical manual research edition-verbal tests, forms a and B-figural tests, forms a and B. Personnel Press, Princeton, NJ.

Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical Combinations and Scientific Impact. Science 342, 468–472. https://doi.org/10.1126/science.1240474

Van de Ven, A., 2007. Engaged scholarship: A guide for organizational and social research. Oxford University Press.

Voight, M.L., Hoogenboom, B.J., 2012. PUBLISHING YOUR WORK IN A JOURNAL: UNDERSTANDING THE PEER REVIEW PROCESS. Int J Sports Phys Ther 7, 452–460.

von Hippel, E., 1976. The dominant role of users in the scientific instrument innovation process. Research Policy 5, 212–239.

Wacheux, F., 1996. Méthodes qualitatives et recherche en gestion.

Wagner, C.S., Leydesdorff, L., 2005. Network structure, self-organization, and the growth of international collaboration in science. Research Policy 34, 1608–1618. https://doi.org/10.1016/j.respol.2005.08.002

Wagner, M., Valls-Pasola, J., Burger-Helmchen, T., 2016. The Global Management of Creativity. Routledge.

Wallach, M.A., Kogan, N., 1965. A new look at the creativity-intelligence distinction. Journal of Personality 33, 348–369. https://doi.org/10.1111/j.1467-6494.1965.tb01391.x

Wang, J., Veugelers, R., Stephan, P., 2017. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. Research Policy 46, 1416–1436. https://doi.org/10.1016/j.respol.2017.06.006

Wechsler, S.M., Saiz, C., Rivas, S.F., Vendramini, C.M.M., Almeida, L.S., Mundim, M.C., Franco, A., 2018. Creative and critical thinking: Independent or overlapping components? Thinking Skills and Creativity 27, 114–122. https://doi.org/10.1016/j.tsc.2017.12.003

Weitzman, M.L., 1998. Recombinant Growth. The Quarterly Journal of Economics 113, 331–360.

Wenger, E., McDermott, R.A., Snyder, W., 2002. Cultivating Communities of Practice: A Guide to Managing Knowledge. Harvard Business Press.

WOODMAN, R.W., 1993. TOWARD A THEORY OF ORGANIZATIONAL CREATIVITY. Academy of Management Review 33.

Yaqub, O., 2016. Serendipity: Towards a Taxonomy and a Theory (SSRN Scholarly Paper No. ID 2841236). Social Science Research Network, Rochester, NY.

Yin, R.K., 2003. Case Study Research: Design and Methods. SAGE.

Yong, K., Sauer, S.J., Mannix, E.A., 2014. Conflict and creativity in interdisciplinary teams. Small Group Research 45, 266–289. https://doi.org/10.1177/1046496414530789

Zou, D., Ma, L., Yu, J., Zhang, Z., 2015. Biological Databases for Human Research. Genomics, Proteomics & Bioinformatics 13, 55–63. https://doi.org/10.1016/j.gpb.2015.01.006

Zuccala, A., 2006. Modeling the invisible college. Journal of the American Society for Information Science and Technology 57, 152–168. https://doi.org/10.1002/asi.20256

Zuckerman, H., 1967. Nobel Laureates in Science: Patterns of Productivity, Collaboration, and Authorship. American Sociological Review Vol. 32, No. 3 (Jun., 1967), pp. 391-403 [WWW Document]. URL https://www.jstor.org/stable/2091086 (accessed 8.19.18).

Zuijdam, F., Deuten, J., Meijer, I., Vermeulen, N., 2011. The role and added value of large-scale research facilities. Technopolis Group.

Table of figures

Table of tables

Université de Strasbourg

Augustin Cournot Doctoral School

Nuria MORATAL-FERRANDO

The Role of Large Research Infrastructures in Scientific Creativity

A l'origine de cette thèse il y a le constat d’une science en changement. Ce changement se caractérise par deux grandes tendances globales : la dépendance croissante à des grands équipements coûteux et partagés et la production de données de masse qui sont également très coûteuses à stocker et gérer. Dans les deux cas ces ressources sont financées par des programmes publics et proposés à la communauté scientifique selon un principe d’ouverture à des utilisateurs extérieurs sous forme de Infrastructures de recherche (IR). Plusieurs facteurs peuvent nous amener à penser que les IR sont des lieux favorables à la créativité. Cependant les moyens par lesquels les IR favorisent la créativité n’ont pas été étudiés. L’objectif de cette thèse est de répondre à cette question. La problématique se décline en deux sous-questions de recherche. D’abord nous nous demandons, comment les IR peuvent-elles contribuer à la créativité scientifique de leurs utilisateurs ? Puis nous nous interrogeons sur : comment mesurer cet impact ?

At the origin of this thesis there is the observation of a changing science. This change is characterized by two major global trends: the growing reliance on large expensive and shared equipment and the production of mass data which are also very expensive to store and manage. In both cases these resources are financed by public programs and proposed to the scientific community according to a principle of openness to external users in the form of Research Infrastructures (RIs). Several factors may lead us to believe that RIs are favourable places for creativity. However, the means by which RIs promote creativity have not been studied. The purpose of this thesis is to answer this question. The research question is divided into two sub-questions of research. First, we wonder how IRs can contribute to the scientific creativity of their users. Then we ask ourselves: how to measure this impact