

# GÉRER & COMPRENDRE

English language online selection

2022 - N° 7

YEAR 2022 - N° 7

A SERIES OF  
ANNALES  
DES MINES  
FOUNDED IN 1794

*Published with the support  
of the Institut Mines Télécom*

03

Managing Religion Situations in the Workplace:  
Roles of a Think Tank for Sharing and Implementing Management tools

**Géraldine Galindo**

ESCP Business School

and **Ewan Oiry**

ESG-UQAM

Original article in French published in *Gérer & comprendre*, March 2021, n° 143, pp. 37-48.

14

Transition “Niches” as Spaces for Re-envisioning the Energy System: The Case of Self-Consumption

**Élodie Gigout, Julie C. Mayer**

and **Hervé Dumez**

i3-CRG, École polytechnique, CNRS, IP Paris

Original article in French published in *Gérer & comprendre*, September 2021, n° 145, pp. 3-12

23

Exploring the Practice of Coaching

**Magali Ayache**

Thema, CY Cergy-Paris Université

and **Hervé Dumez**

CRG-i3, École Polytechnique, CNRS, IP Paris

Original article in French published in *Gérer & comprendre*, September 2021, n° 145, pp. 46-54.

31

Artificial intelligence: between science and the market

Some socio-historical elements to better understand a strange scientific experiment (1956-1990)

**Jean-Sébastien Vayre**

Senior lecturer at the University of Côte d’Azur

Original article in French published in *Gérer & comprendre*, September 2021, n° 145, pp. 55-69.

43

From Organisational Silence to Deontic Anger:  
When Whistleblowers Speak Up

**Philippe Jacquinot**

associate professor, Paris-Saclay University

and **Arnaud Pellissier-Tanon**

associate professor, University of Paris I Panthéon-Sorbonne2

Original article in French published in *Gérer & comprendre*, December 2021, n° 146, pp. 27-38.

# Managing Religion Situations in the Workplace: Roles of a Think Tank for Sharing and Implementing Management tools

By **Géraldine Galindo**  
ESCP Business School  
and **Ewan Oiry**  
ESG-UQAM

Original article in French published in *Gérer & comprendre*, March 2021, n° 143, pp. 37-48.

Religion is gaining prominence as a workplace issue in France. It is a sensitive question for human resources (HR) departments looking for ways to develop management tools for religion in the workplace. In this article, we first reply to this question by drawing on key concepts from the literature. It also offers an analysis of data collected during the regular meeting of a think tank focused on the religious diversity subjects in the workplace. This methodology proved to be essential for collecting new data from a large number of companies (26) on this subject, as French companies are often reluctant to talk about this issue publicly. Applied to these data, these concepts allow us to identify the specificities of the design process of the management systems of religious events in these companies. These tools are then being introduced gradually, with care, to managers who have previously been trained and strongly supported by management. This work allows to develop knowledge on the specificities of the different means to develop a management system to address religious facts in French companies.

## Introduction

In the workplace context, employees may choose either to reveal their religious identity, with varying degrees of expression and demands (Hicks, 2003; King et al., 2009), to hide it or even “passing” it (Clair et al., 2005), due to perceived risks (Exline & Bright, 2011; Gebert et al., 2014). Each employee negotiates with their own various identities (Kreiner, Hollensbe & Sheep, 2006). They may decide to not reveal their beliefs, something that may stigmatise them (Ahmad et al., 2018), and be identified as atheist or, at the opposite end of the spectrum, to exhibit radicalised behaviours (Honoré, 2016; de Maison Rouge, 2017). For employers, this means a range of different issues they may potentially have to manage, from personal requests (e.g. absences for religious holidays, prayer time) to transgressive behaviour (e.g. refusing to shake hands or take orders) (OFRE,<sup>(1)</sup> 2019). For a long time, such issues were not

disclosed as they were considered as taboo (Galindo & Surply, 2010), particularly in the context of France’s strict separation between Church and State (principle of *laïcité*<sup>(2)</sup>) and employers’ aversion to interfere in the private lives of their employees. Issues of religion in the workplace were therefore relegated to the part of the diversity iceberg that remains below the waterline, typically with little management interventions (Cui et al., 2015).

But workplaces have had to respond to these issues. According to the most recent (2019) survey by France’s Observatory on Religion in the Workplace (OFRE),<sup>(3)</sup> more than 70% of employers surveyed reported having had to manage religious issues either regularly or occasionally in the past year, compared to 44% in 2014. While still a sensitive area (Chan-Serafin, Brief & George, 2013), more than half of such situations involving religious expression by an employee required management intervention in 2019 (compared to a

<sup>(1)</sup> For the most recent version (in French): [www.institutmontaigne.org/publications/religion-au-travail-croire-au-dialogue-barometre-du-fait-religieux-en-entreprise-2019](http://www.institutmontaigne.org/publications/religion-au-travail-croire-au-dialogue-barometre-du-fait-religieux-en-entreprise-2019)

<sup>(2)</sup> We decided to keep in this text the principle of *laïcité*, a specific France state of secularism.

<sup>(3)</sup> *Observatoire du fait religieux en entreprise*

quarter in 2014, according to the OFRE). This topic has become a societal issue and, at the same time, is one of the most sensitive areas of HR management, where it can be tricky to assess the situation and to move towards action (Renzetti & Lee, 1993; Condomines & Hennequin, 2013).

The question addressed by this paper is: “How can management tools be designed to help manage religion in the workplace in France?” We answer this question by drawing on key concepts from the literature on management tools that has been developed in France since the early 1980s. We use this literature framework to analyse processes used to design management tools for religious issues as presented during the monthly meetings of a think tank, which for three years gathered representatives from large public and private employers operating in France. We played a central role in facilitating the group, which allowed us to collect data from numerous workplaces (26 organisations) on a sensitive topic that most employers are still reluctant to discuss publicly (Marinos, 2018). This approach is that of a generative case study (Siggelkow, 2007), producing knowledge both on how religious issues are managed and on the management tools used.

## How managing religion in the workplace has evolved in France?

The literature reveals the complexity and diversity of approaches to managing religion in the workplace. It is increasingly less concerned with analysing the question of “why (is this an issue)?” (Galindo & Surply, 2010) or “what shape does it take?” (Vickers, 2015), focusing more on the question of “how (to address the issue)?” (Syed et al., 2018). The challenge for organisations and researchers today is therefore learning how to handle these types of management situations.

### A significant shift in workplace practices and in the literature describing them

A literature review reveals that following an initial phase of shock, organisations gradually begin to form a structured response. We have identified four phases:

#### Phase 1 – Shock

Religion in the workplace is not a new phenomenon (Galindo & Surply, 2010). However, the first decade of the 2000s marked a turning point (Honoré, Galindo & Zannad, 2019). Increasingly, employees were seeking recognition of their overall identity, including practices related to their religious beliefs (King et al., 2009). Employers were initially unsettled by such expectations. While they encouraged their employees to be more open to their own identity, as part of growing work/life balance policies, they were now witnessing a new set of expectations, not only relating to their employees’ private lives (*forum internum*), but also visible through certain practices (*forum externum*). This meant more and more employees were no longer concealing their faith or religious beliefs (Guillet & Brasseur, 2019), as a way to balance or, at the extreme end of things, to merge their identities.

The initial reaction of employers is often to bypass or ignore such issues. Diversity representatives report being “powerless” or “paralysed” in their ability to respond to them (Galindo & Surply, 2010). Kirton & Greene (2015, p. 3) also point out that although preventive measures exist particularly in western countries to eliminate religious discrimination, organisations usually take a less proactive approach to addressing such issues. While managers are initially encouraged to resolve issues on a case-by-case basis (Banon & Chanlat, 2014), this eventually gives way to a need for a standardised response.

#### Phase 2 – Turning to legislation

After the shock phase, many employers (primarily large organisations) decide to search laws to help them address religion in the workplace. Since the principle of separation of Church and State does not concern the private sector, there are three legal principles on which private-sector companies can base their policies. The first is to guarantee their employees’ freedom to hold or not hold religious beliefs (according to the Declaration of the Rights of Man and of the Citizen [1789] and Article L.1121-1 of France’s Labour Code). Employers must also allow their employees to manifest their religion or beliefs (Article 9 of the European Convention on Human Rights). They must also ensure that employees receive equal treatment (Article 1 of the French Constitution [1958] and Directive 78/2000/EC) and are not discriminated against for their beliefs (Article L.1132-1 of the Labour Code). However, in searching for a legislative basis for their actions, employers find that there are grey areas. How, for example, can they guarantee the freedoms of some individuals while still ensuring equal treatment for all in the workplace?

#### Phase 3 – Producing best practice guides for employers

Faced with these legal uncertainties, organisations began seeking out other resources (Pastor, 2016). The early 2000s saw the development of corporate guidelines in the United States. Cash & Gray (2000), for example, list the factors that employers should consider for determining the most effective managerial response to requests for religious accommodations. In France, in a decision of 04/06/09, the HALDE (Equal Opportunities and Anti-Discrimination Commission<sup>(4)</sup>), ruled that an employee’s freedom of religion and belief ends at the point where it constitutes misuse of the right of expression, proselytising or an act of pressure toward other employees. It ensures the organisation is able to operate effectively (an expression of religion must not hinder the performance of work, how work is structured or the organisation’s business interests), preventing any kind of proselytising and protecting the health and safety of employees. There are also organisations that produce guides to address employers’ concerns<sup>(5)</sup>. Such guidance, not produced in-house by the employer but by non-profit or academic third parties, helps

<sup>(4)</sup> Now named *Défenseur des droits* “the Defender of Rights”

<sup>(5)</sup> An example: “Comment gérer la diversité religieuse en entreprise”, a guide on managing religious diversity in the workplace published by non-profit organisation IMS-Entreprendre pour la Cité, 15 May 2009.

clarify the issues surrounding religion in the workplace and contextualises the response to these issues. However, there is no way to ensure that all managers have access to these resources or the appropriate approaches to respond to their employees or colleagues.

#### Phase 4 – Producing in-house management rules

Faced with a wide range of questions and a need to disseminate best practices throughout the organisation, employers have increasingly begun formalising their practices and responses by designing management tools (Cintas et al., 2013; Galindo & Zannad, 2014). This results in each organisation producing its own in-house guidance, which is developed and structured according to its own criteria (Chan-Serafin et al., 2013). The objective of these new management rules is to influence how employees are governed, to attempt to manage the areas of freedom and autonomy they make for themselves (Reynaud, 1988, p. 10). Often, religion-specific guides are developed that include explanations on legal concepts and practical examples with FAQs on managing religious situations (Ludlum, 2016). Employers also use a range of other methods, such as training (Gaillard & Jolivet, 2019), to educate as many people as possible about such issues and how to respond to them.

The progress made by organisations on the sensitive issue of religion in the workplace has been a step by step process. Many large organisations have gone from a passive stance to an active one, deciding to produce their own management rules and introducing new management tools.

#### The literature on management tools

Building on the work of Girin (1981) and Berry (1983), a robust literature has been developed, particularly in France, on management tools. There are concepts from the literature that are particularly useful for offering solutions to employers looking to design management tools for religious issues.

The literature defines management tools as a formalisation of structured action (Moisdon, 1997, p. 7) and shows that they are social constructs (Gilbert, 1998; Akrich et al., 2006). This concept is important; due to their technical nature, management tools are often perceived by those who use them in workplaces as “given” (Lorino, 2005). Specifically, users assume these tools must be used as it is, that there is no other form they could possibly take and that they cannot be altered. This point is illustrated by Bayart (1995). In tracing the history of the concept of “quality” in industry, the author shows that what we intuitively believe to be given, immutable and unquestionable (a product is either of good quality or it is not) is actually the result of a real history, a social construct, something that takes time and involves developing tools and knowledge (including, specifically, a theory of statistical control). Research on management tools has shown this “representationalist” perception of management tools, in which they are considered to be an accurate reflection of an operational reality, is not an appropriate conceptualisation of management tools (Lorino, 2018). On the contrary, such tools are

social constructs, the result of complex interactions between stakeholders with different interests (Chiapello & Gilbert, 2013). Because of this, the final form of any tool is not given and could have ended up being very different. But the final form chosen for a management tool causes the swarm of social conditions influencing its design to disappear (Woolgar & Latour, 1988; Latour, 1992; Drevetton, 2010).

This concept is crucial for devising management tools for issues of religion, as it indicates that their design is not only based on the aforementioned legal framework, but is also influenced by the thought processes of individuals and by the specific context of organisations. The design of such tools should therefore be analysed as a true managerial act, through which different coalitions of stakeholders within an organisation collectively reach a compromise on how to manage religious issues.

The literature also shows that management tools can take a wide variety of forms (memos, guides, formal or informal reward/punishment systems, etc.) (Brivot & Gendron, 2011). This highlights the importance of not limiting analysis to a single type of management tools (Rabardel, 1995; Rabardel & Bourmaud, 2003). By looking at only the technical aspects of a single tool, the analysis will overlook what constitutes the essence of a management tool (Hatchuel & Weil, 1992; Labatut et al., 2011). The literature therefore shows that it is necessary to look at management tools, that some researchers call management “frameworks”, in other words the complex arrangements of multiple human and non-human components, with stakeholders and management tools given equal consideration (Boussard & Maugeri, 2003; Akrich et al., 2006). Management “frameworks” should therefore be thought of as being central to interactions between individuals, their work and the organisation (Rabardel & Bourmaud, 2003).

Lastly, the literature highlights the importance of analysing how management tools are used and how they evolve. In this respect, the concept of assimilation plays a central role (De Vaujany, 2006). Management tools are never employed in the exact way their designers envisioned (Aggeri & Labatut, 2010). While sometimes used as designed, in most cases these tools are not used as their designers initially planned (Grimand, 2012). Uses not considered by the original designers are therefore central to the assimilation process. Such uses are so common that, as a feedback effect, they often contribute to transforming the management tool itself (Oiry, 2011).

In summary, the literature on management tools that has been developed in France since the 1980s offers concepts that we consider to be particularly useful for identifying how to design management tools for religious issues in the workplace. Specifically, it proposes considering management tools:

- as complex arrangements (and not only isolated management tools)
- co-developed by multiple stakeholders in the workplace
- designed to address specific workplace challenges
- that undergo a transformation over the long term.

## Research study design

Our methodology is that of a generative case study (Siggelkow, 2007). The case study approach enabled us not only to analyse our data using the concepts found in the literature on management tools, but also to add to this literature, as well as to the literature on religion in the workplace. Although it is a highly topical issue, it remains a sensitive subject (Chan-Serafin, Brief & George, 2013) and employers are still reluctant to speak up about the challenges they face, and allow researchers to study their practices. To overcome this difficulty and collect a large number of case studies, we conducted an action research study (David, 2001), facilitating think tank meetings over a three-year period.

### Think tank as data collection method

A group of managers can be likened to a network of organisations: a group of stakeholders seeking to establish and maintain relationships with each other, but without any kind of official organising authority (Podolny & Page, 1998). There are numerous benefits to be gained from this type of forum for collaboration or discussion (Marinos, 2018). Tacit knowledge is developed as information and experiences are shared, which makes organisations more inclined to adopt innovative solutions (Bevort, 2006, in Marinos, 2018). It is an opportunity for group members to build up their social capital; the group allows them to solidify their relationships and occupy a position at the crossroads between their own organisation, other organisations and their broader environment. These group members belong to an intellectual community (Cucchi, 1999; Polge, 2009) organised around a mutual commitment (there is reciprocity in giving one's time for something in return), a joint undertaking (there is a structure) and shared resources (such as resources for communicating) (Marinos, 2018, p. 124).

Our data were collected during a think tank meetings, dedicated to sharing information and practices on the issue of religion. The objective of this group is consistent with that of other groups: the members work

in different fields in different organisations, and the knowledge that is produced gives rise to tools that can be operationalised as opposed to directly operational best practices (Polge, 2009, p. 229). The group met once a month over three – to five – month periods from 2016 to 2019, bringing together HR managers, diversity managers, legal executives and security executives (Table 1).

As shown in the table, the group's members were mainly large French organisations (public undertakings, private companies and EPICs<sup>(6)</sup>). Each meeting included presentations by researchers and the group's scientific lead, as well as presentations by managers (from both member and non-member organisations).

### Role of the researcher/facilitator

The challenge for the researcher was in serving both as a source of knowledge and perspective and as facilitator of the discussions. Over time, the interactions and documents collected from the meetings became sufficiently valuable to be useful for research purposes. The researcher guided interactions between group members without being forced to intervene for each of the organisations represented. Our role could therefore be described as participant observation, a research method involving situations where the researcher is a member of the social community he is observing (Platt, 1983, in Berger-Douce, 2010, p. 135). It involves both participating in discussions and observing the data that is being collected through the lens of a researcher. As a method, it was useful for gaining access to highly confidential information on the taboo subject of religion in the workplace, and for achieving a rare level of understanding and discussion afforded by a situation where the interactions are not originally designed to be a research opportunity (Crespo-Febvay & Loubès, 2019, p. 85). The quality of the data is a clear testament of the trust that the group members placed in the researcher, who self-identified as such.

<sup>(6)</sup> Établissements publics à caractère industriel et commercial (public organisations of an industrial and commercial nature)

Year	Number of meetings	Members	Number of presentations by researchers	Number of presentations by employer representatives
2016	5	ADP, Atlantic, Bouygues, CDC, Covea, EDF, Egide, Danone, MMA, Michelin, Orange, Pôle Emploi, RTE, SNCF, Société Générale, Veolia	9	5
2017	3	Atlantic, BPCE, CDC, Covea, Leroy Merlin, Orange, RATP, RTE, Safran, SEB, SNCF, Veolia	5	5
2017	1	Open meeting (group members and other organisations)	2	-
2019	3	BPCE, Enedis, Essilor, Française des Jeux, Pôle Emploi, Leroy Merlin, Orano, RATP, RTE, Safran, Total	4	5

Table 1: Group meetings and members

## Data collected

We used our privileged position as scientific lead and facilitator of the group's various sessions to collect a variety of data (**Table 2**). We conducted interviews with managers in advance of their presentations (ranging from 30 minutes to 2.5 hours). Detailed notes were taken during each session, supplemented by recordings of some discussions and minutes of the sessions taken by a third party. Documents (e.g. PowerPoint slides, guides, charters) were also collected at these sessions.

Using the step-by-step thematic analysis described by Braun & Clarke (2006), we first sought to familiarise ourselves with the range of research materials. To do this, we independently reviewed each type of data that had been collected, in order to begin identifying key concepts and ideas.

From this detailed review, we identified the main themes of our analysis: triggers, stakeholders, steps in the process, durations, tools introduced, barriers, levers, adjustments made to management tools. Next we conducted:

- Vertical thematic analyses, where each type of data was divided into different themes (primary, secondary, emerging),
- Horizontal thematic analyses, where data on each theme was reconciled, from which we were gradually able to triangulate the data.

We found this “bricolage” approach to data analysis, described by Dumez (2016), to be necessary to preserve the variety of the data and continue working toward reconciling the data. It was never a matter of trying to idealise the processes, but of recognising potential biases (Creswell, 2013) and identifying incidents or outliers (Bisel & Barge, 2011) that disrupt the homogeneity of the organisations' approaches.

## Results

The data collected from the think tank reveals both the diversity of issues being addressed by the organisations and their gradual progression towards designing more “systemic” management tools.

### Employers compelled to think about management tools

Confirming what has long been documented in the literature, the French organisations participating in the think tank adopted reactive attitudes with regards to the situations encountered on these religious issues. Some had joined the group as a way to take action (“*We’re looking for a firm position on the matter*”).<sup>(7)</sup> However, most organisations had already undertaken a process of developing rules in response to three types of events:

- **Alignment with legislation:** For public employers, there is sometimes the need to clarify the scope and application of the principle of laïcité. The drafting of a neutrality clause can therefore open the door to broader conversations on how to enforce such a clause, instigating a process of designing applicable tools.
- **Integration as part of a broader inclusivity initiative:** Many employers also have more wide-ranging diversity policies, in which religious diversity is just one component. For example, one organisation has a diversity and inclusion policy with five priorities, one of them is on the origins divided into sub-topics: minorities, interculturality and religion. The issue of religion is therefore introduced as part of other rules put in place, and becomes a subject addressed by management rules.

<sup>(7)</sup> All quotations from participating organisations have been translated from French. For confidentiality reasons, we are unable to attribute quotations to specific organisations (refer to the methodology table for the full list of participants).

Year	Type of data collected
2016	5 semi-structured interviews in preparation for group sessions Preparatory questionnaire on actions implemented to manage religion in the workplace (11 responses) 5 sets of minutes from work sessions 4 guides/charters 3 PowerPoint presentations on measures introduced by organisations Audio recordings of discussion sessions Notes taken during every session
2017	5 semi-structured interviews in preparation for group sessions 5 PowerPoint presentations on measures introduced by organisations 2 guides/charters Audio recordings of discussion sessions 4 sets of minutes from work sessions Notes taken during every session
2019	5 semi-structured interviews in preparation for group sessions 4 PowerPoint presentations on measures introduced by organisations 2 guides/charters 3 sets of minutes from work sessions Notes taken during every session

Table 2: Data collected for the research study

- **In response to an event:** For some employers, the instigator is a significant event. At one organisation, a secret prayer room was found at the head office and HR management wanted to “come on strong”. At another, an employee was photographed in his uniform praying next to their vehicle. Sometimes an employer’s reaction is the result of a succession of events, none necessarily significant in themselves, but that cause religion to be identified as a management issue. Another reason cited for developing a management tool is feedback from the field, most often in the form of questions. It is therefore a matter of shifting away from making calls for “common sense” or “peaceful coexistence” and towards managing this type of organisational situation. The challenge is to “provide managers with keys to understanding” that are shared by all.

Management tools for religious issues are therefore often introduced either in response to a context that is conducive to this type of discussion, or where a swift response is expected. For others, participating in the group was also a way to learn more about the topic and start to think about tools they might want to introduce if they were to initiate such a process.

### Tools that include training and a structured rollout plan

For these organisations, there are three complementary dimensions that went into the design of their management tool.

#### Tools, the central components of the process

In searching for a way to manage the situation, some employers opt to produce their own guidance. The aim is to equip managers to handle situations (“to give managers the key to understanding”). Depending on the situation, they may decide to draft a charter or guide or adopt an existing guide (“during the training, we handed out packages including a guide developed by the Ministry of Labour”). In all cases, these management resources include explanations of legal concepts and practical examples for direct managers. A formalised tool is therefore deemed necessary to outline the policies and practices currently in effect and those to be adopted in the future. These tools are seen as a way to ensure a standardised and consistent response, and to engage in discussions around shared concepts, particularly legal guidance (“It’s a tool for discussions with employees”). These tools are an internal indication that rules and responses exist and that everyone is expected to be familiar with them. They are central to all of the processes undertaken on this issue. However, they are non-binding and do not have any legal value (it is about “producing a best practices guide in response to requests from managers”). In rare cases, charters are presented as more binding on employees (“Everyone has to sign it”).

#### Training perceived as essential

All organisations of the group made training central to their process. In-person sessions were used to help managers understand the context (namely the legal context) of these workplace situations, to conduct role-playing exercises or to provide information about

sensitive topics such as radicalisation (“It creates meaning, provides a frame of reference”). Some training was also provided online, via educational games or short videos on specific topics (e.g. *laïcité*). During the training, direct stakeholders are able to speak openly about this sensitive subject (“People start opening up as the day goes on”), to discover and familiarise themselves with the tools, and to sometimes bring up cases that are not known to management. Training sessions are also spaces for managers to talk to other managers, to help them feel less alone in handling these situations.

#### Structured rollout plan

The methods used to disseminate the tools vary from organisation to organisation. At one, physical copies are handed out at the end of a training session (“You have to complete the day of training to get the guide”), another publishes them on the workplace intranet (“Buried 15 clicks deep”) and another limits distribution to diversity and HR managers (“Just in case”). To help users grasp the content of the tools, some employers use a question/answer game or multiple-choice quiz. In most organisations, there are phases to the rollout plan. For example, one organisation planned to roll out the tools to managers and HR in an initial phase, with a second rollout (still to be determined) “Possibly organisation-wide or broader”. In some organisations, distributing the tools is also optional (“Guides are distributed as managers see fit; there’s no obligation”). The group members seemed to be more in favour of distributing the tools “naturally” rather than on a “forced schedule” (“People are talking about the training... word is spreading”).

Using the example of Organisation A, we can see how these three dimensions can be combined in practice.

#### BOX 1: EXAMPLE OF A MANAGEMENT TOOL INTRODUCED BY ORGANISATION A

1. A guide on “coexisting in diversity” is put together by a working group involving field workers. The guide includes case studies and a managerial decision-making tool.
2. The guide is distributed during rollout meetings to all on-site employees.
3. Training is delivered to all stakeholders, along with a card game to help employees understand the different types of stakeholders and reactions in the organisation.
4. Diversity managers are appointed and receive training (“The representatives pass on messages and, in the event of an incident, encourage the person to speak with the other person involved instead of leaving the situation unresolved”).

#### Features of the tools

The approach taken to develop these tools can be described as cautious and “inclusive”:



### Long implementation period

The organisations took their time (between 1 and 2.5 years) between taking the first step and rolling out the management tool. It was a process, often described as a succession of key steps (e.g. presenting to management, putting together working groups, delivering the first training workshops). Notably, in most cases, the tools that ended up being introduced were tweaked or even redesigned after a few months or years.

As shown in the timeline above, Organisation E took several years to roll out its framework.

This is a prime example of a management tool that was introduced gradually over time involving different versions. Other organisations also reworked their tools (mainly updating guides), but above all they reviewed their entire process by involving additional stakeholders and making a communication plan for the rollout and update.

### Caution taken by senior management

Throughout the process, some group members reported encountering resistance, or even outright rejection,

from management at first (from a diversity manager: “*The president initially vetoed it, saying: ‘We’re not going to do it that way’, and then they forced us to start by interviewing the organisation’s 100 managers in France and he told us: ‘Start at the top’*”). It is clear that even if requests are coming from the field, the process is initiated by management, in a top-down approach, with the objective of ensuring consistency but also protecting the organisation’s reputation, since some see it as a strategic issue (“*In our organisation, it’s a way of affirming that we have values*”). This sense of caution results in having to “*weigh every word*” and accept that measures will be carried out “*step by step*”. Managers position themselves as the “owners” of the process and requires approval over every step of the development process. Reactions to management making it a strategic issue are mixed: sometimes it speeds up the process (“*The issue is led by senior management in order to attach importance to it*”); other times it slows things down due to requests for additional clarification.

### Co-development by a wide range of stakeholders

In all the organisations, the process of introducing a management tool involved a variety of stakeholders, which on the one hand ensures a diversity of viewpoints

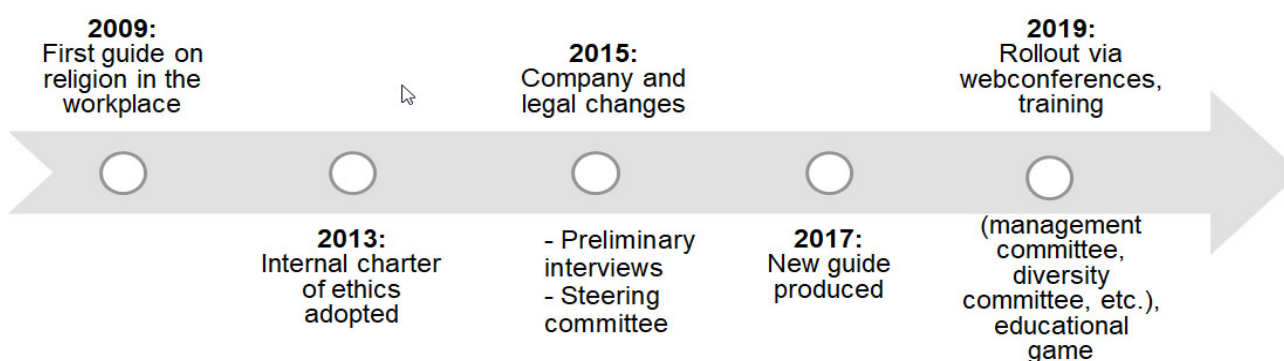


FIGURE 1: A TOOL INTRODUCED OVER TIME IN ORGANISATION E

### Box 2: KEY ELEMENTS OF THE MANAGEMENT TOOL INTRODUCED BY ORGANISATION E

Although the first version was considered to be “*rather weak compared to the organisation’s position*”, it was a way to begin the process and plan for future changes.

1. To fine-tune the new approach, interviews were conducted with business line managers to identify situations, two field visits were organised, meetings were held with researchers, etc.
2. A steering committee (comprising innovation, diversity and job performance units, managers from the business lines and the ethics, security and business intelligence functions) then tackled the more detail-oriented task of reworking the content for the new version of the guide, which was more focused on “*case studies and managerial decision-making tools*”. This enabled the organisation to transition from a stance where it is “*focused solely on the employee making the request*” to one where the organisation is asking questions: “*What are the employer’s rights vis-à-vis the employee? And what are the employee’s obligations to the employer?*” (diversity managers).
3. Unlike with the first version, a series of activities were planned: before being rolled out, the guide was reviewed and presented to various stakeholders (selected managers, top management, HR, legal affairs, unions, etc.).
4. The rollout strategy included training, educational games, etc. to allow stakeholders to “*assimilate*” the tools.
5. Long-term monitoring and steering activities were planned, via a network of diversity and business line managers and the ethics and compliance department, to ensure consistency in the organisation’s responses and to identify any potential issues that arise.

on the issue, but on the other hand inevitably slows down the process. Some stakeholders are considered “obvious” drivers of the process. Diversity managers are often on the front line, as initiators, sponsors and owners of these tools. They work in collaboration with other functions in the organisation (“*The process was initiated by the group’s HR management in conjunction with managers from diversity, legal, ethics and security*”). A variety of stakeholders, with different practices and timelines, are therefore involved in the process, sometimes in the form of a project group devoted to producing tools or at least responses to the most common situations (“*We put together a working group on the issue of how to respond to a female employee wearing a headscarf*”). Unions also play different roles: as a driving force in one organisation (“*You didn’t go far enough, you could have even been more strict*”), as collaborators in some (“*The charter was co-written with the unions*”, “*The unions were receptive to finding solutions together*”), or followers in others (“*The unions have kept a low profile*”), but very rarely do they interfere.

**BOX 3: EXAMPLE OF A CO-DEVELOPMENT APPROACH BETWEEN STAKEHOLDERS IN ORGANISATION C**

1. The state of play of the situations needing to be addressed : by diversity managers.
2. A survey was sent to all employees and managers.
3. A project group (14 people) was put together to conduct awareness-raising activities and share the results of the survey.
4. The project group produced a set of recommendations.
5. The project group produced a guide.

Taking caution and involving a variety of stakeholders at different stages slows down the process, sometimes even impeding it or resulting in a change of plans. In all cases, the organisations proceeded cautiously in the face of these issues and did not hesitate to slow down the project to make it as secure as possible. The time taken to develop the tool allowed them to consider how practices might work together as part of the process.

**Country-specific tools**

These management tools are often presented as being specific to France. Country-level differences are used to justify limiting tools to national level (“*We’re being extremely cautious in terms of adopting a global perspective*”, “*There is no way to have a one-size-fits-all policy; France is a special case in this area*”). The management tools therefore do not concern and are not used in other contexts, for example in the United States, where “*Religion is everywhere*”, or in countries where “*Women wear headscarves and there is no issue with it*”. However, many of these organisations also mentioned how France is seen as an “example” (“*France’s position on the issue has attracted international attention*”).

**Tools that address managers’ challenges**

“Before, managers were on their own. Now they have a tool for having discussions” or “For managing discomfort”; “Information is being provided”; “The act of clarifying a rule has a big impact”. These frameworks help employers produce clear and consistent responses (“There’s consistency in our responses”) and extricate themselves from contentious and/or risky situations for the organisation’s image (“It helped us defuse the issue”; “We’ve put up firewalls”). In this context, the ultimate objective is “getting along”, “working well together” or “peaceful coexistence”, expressions widely used in the think tank to consider the issues at stake in the ongoing process. However, it should be emphasised that the aim is not to arrive at an ideal response, but rather a managerial response that is in line with legal principles and the culture of the organisation.

**Discussion**

Our analysis of the data using the concepts provided by the literature on management frameworks has yielded a number of theoretical and managerial insights.

**Theoretical insights**

While many studies point to the pervasiveness of “management-itis”, i.e. the tendency to constantly and rapidly develop new tools, which are not always appropriate in relation to managers’ practices (Detchessahart & Journé, 2007), our data shows that, on the issue of religion in the workplace, organisations have demonstrated what might be called a “model” approach to designing their management tools. Our data show that organisations take their time, make adjustments to their tools, develop sophisticated process combining different tools, and think through training and communication plans. These are all the features of polyphonic change management (Pichault, 2013). Organisations then complain of resistance to change (Bareil, 2009) or tools that do not produce the anticipated effects and are quickly abandoned (Chiapello & Gilbert, 2013). Faced with sensitive HR management issues (Renzetti & Lee, 1993; Condomines & Hennequin, 2013), it appears that organisations are rediscovering the importance of these aspects of polyphonic management: as part of a long-term process, they take a cautious approach and attach importance to their tools by involving as many stakeholders as possible. The caution taken by organisations in developing policies on sensitive HR topics appears to be a best practice that should possibly be followed in all areas of HR management.

Our data also show a number of cases of “successful” incentive-based (as opposed to restriction-based) management tools. These results show that management tools, developed with the abovementioned features in mind, do not need to adopt a reward/punishment model to be used and yield results in organisations. To draw a parallel with research on variable compensation (Landry et al., 2017), it appears that management tools for religious issues should be “informational” rather than “controlling”. The methods expected by the organisation for managing religion in the workplace should therefore be further developed after the fact as opposed to being imposed as-is from the outset.

Our work therefore proposes an alternative perspective: a positive perspective that showcases “model management tools” that have been introduced by organisations, in contrast to the majority of the literature on managing religion in the workplace, which focuses on the difficulties that are encountered, as recently pointed out by Miller (2020).<sup>(8)</sup> Like Miller, without denying the “conflictual forces” on the issue, we demonstrate how management frameworks can support a “faith-friendly” approach (Miller & Ewest, 2015) or an accommodation-based approach (Galindo & Zannad, 2014), focusing on common guidelines for action, even though the topic of religion is considered to be highly contextualised (Honoré et al., 2019).

This research also repositions management tools as central to diversity initiatives. In this way, it differs from many studies centred on policies introduced by organisations and their related issues, or on individual expressions and specific features of diversity (Héliot et al., 2020). It falls somewhere in between, shedding light on practices and processes effectively introduced in organisations. It places these actions within an emerging approach for managing diversity identified by Thomas and Ely (1996) which they call the “learning-and-effectiveness” paradigm. The challenge faced by organisations does indeed correspond to this new approach: wanting to both recognise and value employees’ differences, as part of a shared learning effort. Our research therefore departs from the diversity management approaches traditionally used in the first two paradigms identified by Thomas and Ely, which are based on a normative perspective (where all individuals are held to a common standard) and a differentiating perspective (where individuals are recognised for their differences). Our results, which focus on articulation of several management tools (as opposed to isolated tools) and the caution taken in designing them, present an opportunity to provide real substance for this new and inclusive approach to diversity.

### Managerial insights

Our research offers pragmatic guidance for organisations looking to introduce or expand measures for managing religious diversity in the workplace. It encourages thinking systemically about the approach, expanding on a perspective that is often focused on a single tool (a guide) and driven by a copy-paste impulse. It also highlights the need to involve stakeholders from across the organisation: not only senior management, to help attach importance to the initiative, but also all managers and employees, to ensure they fully assimilate the tool. While diversity initiatives often follow a top-down approach (Thomas & Ely, 1996), our research highlights the role of a bottom-up approach, in order to understand the expectations of managers, and employees more broadly, and to get their feedback on the tool after it is introduced. It also confirms that processes should be guided by caution, even when that means delays if it seems the tool is not meeting the needs of all stakeholders. Lastly, our research shows

the value of a think tank for employers, as a source of finding and sharing improvements, both formal and informal, and of encouraging action between, over and above the information they set out looking for.

### Future research directions

Like all research studies, our analysis has its limitations, which also present avenues for future research. Specifically, this study is based on data collected during group meetings from discussions between members of organisations that had already introduced tools or were beginning to do so. These members therefore had a benchmark and benefited from each other’s lessons learned, which led them to adopt the systemic and cautious approach described above. For a future study, it would be interesting to analyse management frameworks introduced by other organisations, not belonging to this group, with different features, such as small businesses or mid-sized companies. Would they take the same type of approach to this sensitive issue? The same question could be put to organisations operating in other countries, namely English-speaking countries, where historical, legal and cultural differences likely influence how religion is managed in the workplace (Honoré et al., 2019).

It would also be interesting to study the application of these methods to other aspects of diversity. Policies and practices have been introduced for visible aspects of diversity (gender, age, disability) and then applied to other invisible aspects (Cui et al., 2015). Could management of religious diversity initiate a reverse process, providing a renewed perspective of approaches that have already been used for other types of diversity?

Finally, the future of these processes is a subject for investigation. The organisations in our sample went beyond the “emerging learning” described by Galindo & Surply (2013). With these management frameworks, they entered into an integration phase, which allowed them to move towards a shared understanding of the subject internally and towards coordination through mutual adjustment. It would be interesting to see how they could reach the final step of the learning process: institutionalisation (Crossan et al., 1999), or the implementation of routines and repeated actions. That said, the question is whether the organisations want to systematise their responses on an issue as sensitive and unpredictable as religion in the workplace.

### References

- AGGERI F. & LABATUT J. (2010), “La gestion au prisme de ses instruments : Une analyse généalogique des approches par les instruments en gestion”, *Finance, Contrôle, Stratégie*, 13(3), pp. 5-38.
- AHMAD A., SABAT I. & KING E. (2018), “The Upsides of Disclosing Your Religion, Sexual Orientation, or Parental Status at Work”, *Harvard Business Review*, 12, March, pp. 2-12.
- AKRICH M., CALLON M. & LATOUR B. (2006), *Sociologie de la traduction. Textes fondateurs*, Paris, Presses des Mines de Paris.
- BANON P. & CHANLAT J. (2014), “La diversité religieuse et culturelle dans les organisations françaises contemporaines”, in PERRET V., NOGATCHEWSKY G. (dir.), *L’état des entreprises*, Paris, La Découverte, Coll. Repères, pp. 31-44.

<sup>(8)</sup> According to Miller (2020), religion in the workplace is divisive and leads to harassment, proselytising and quid pro quos, and accommodations are disruptive to the work environment.

- BAREIL C. (2009), "Décoder les préoccupations et les résistances à l'égard des changements", *Gestion*, 34(4), pp. 32-38.
- BAYART D. (1995), "Des objets qui solidifient une théorie : l'histoire du contrôle statistique de fabrication", in CHARUE-DUBOC F., *Des savoirs en action*, Paris, L'Harmattan, 1995, pp. 139-173. <https://hal.archives-ouvertes.fr/hal-00263204>
- BERGER-DOUCE S. (2010), "La fabrique d'une stratégie collective entrepreneuriale orientée environnement : l'expérience d'un club de dirigeants des TIC", *Revue internationale P.M.E.*, 23 (3-4), pp. 123-154.
- BERRY M. (1983), *Une technologie invisible : l'impact des systèmes de gestion sur les comportements humains*, Paris, Éditions de l'École polytechnique. <https://hal.archives-ouvertes.fr/hal-00263141>
- BOUSSARD V. & MAUGERI S. (2003), *Du politique dans les organisations*, Paris, L'Harmattan.
- BRIVOT M. & GENDRON Y. (2011), "Beyond Panopticism: On the Ramifications of Surveillance in a Contemporary Professional Setting", *Accounting Organizations and Society*, 36(3), pp. 135-155.
- CASH K. C. & GRAY G. R. (2000), "A Framework for Accommodating Religion and Spirituality in the Workplace", *Academy of Management Perspectives*, 14(3), pp. 124-133.
- CHAN-SERAFIN S., BRIEF A. & GEORGE J. (2013), "How Does Religion Matter and Why? Religion and the Organizational Sciences", *Organization Science*, 24(5), pp. 1585-1600.
- CHIAPELLO E. & GILBERT P. (2013), *Sociologie des outils de gestion*, Paris, La Découverte.
- CINTAS C., GOSSE B. & VATTEVILLE E. (2013), "Religious Identity: A New Dimension of HRM? A French View", *Employee Relations*, 35(6), pp. 576-592.
- CLAIR J. A., BEATTY J. E. & MACLEAN T. L. (2005), "Out of Sight but Not Out of Mind: Managing Invisible Social Identities in the Workplace", *Academy of Management Review*, 30(1), pp. 78-95.
- CONDOMINES B. & HENNEQUIN E. (2013), "Étudier des sujets sensibles : les apports de l'approche mixte", *Revue Interdisciplinaire Management, Homme(s) & Entreprise*, 5(1), pp. 12-27.
- CRESPO-FEBVAY A. & LOUBÈS A. (2019), "Le dialogue interentreprises comme levier d'innovation responsable", *Innovations*, 59(2), pp. 75-102.
- CROSSAN M. M., LANE H. W. & WHITE R. E. (1999), "An Organizational Learning Framework: From Intuition to Institution", *Academy of Management Review*, 24(3), pp. 522-537.
- CUI J., JO H., NA H. & VELASQUEZ M. G. (2015), "Workforce Diversity and Religiosity", *Journal of Business Ethics*, 128(4), pp. 743-767.
- DAVID A. (2001), "Logique, épistémologie et méthodologie en sciences de gestion: trois hypothèses revisitées", in DAVID A., HATCHUEL A. & LAUFER R. (coord.), *Les nouvelles fondations des sciences de gestion*, Paris, Vuibert.
- DETCHESSAHART M. & JOURNÉ B. (2007), "Une approche narrative des outils de gestion", *Revue Française de Gestion*, 174, pp. 77-92.
- DREVEYON B. & ROCHER S. (2010), "'Lost in translation', étude de la construction d'un outil de gestion dans une région française", *Comptabilité-Contrôle-Audit*, 16(1), pp. 83-100.
- EXLINE J. & BRIGHT D. S. (2011), "Spiritual and Religious Struggles in the Workplace", *Journal of Management, Spirituality & Religion*, 8(2), pp. 123-142.
- GAILLARD H. & JOLIVET T. (2019), "L'entreprise émancipatrice : Dépasser la "régulation" du fait religieux au travail par le disputatio", *Revue des Sciences de Gestion*, 3, pp. 89-101.
- GALINDO G. & SURPLY J. (2010), "Quelles régulations du fait religieux en entreprise ?", *Revue internationale de Psychosociologie*, novembre, pp. 29-55.
- GALINDO G. & SURPLY J. (2013), "Quel processus d'apprentissage de la gestion du fait religieux dans les entreprises françaises ?", *Management International*, 17, pp. 37-49.
- GALINDO G. & ZANNAD H. (2014), "Les grandes entreprises françaises et la religion. Proposition d'une grille d'analyse pour décrypter les postures adoptées", *RIMHE*, 13, pp. 40-53.
- GEBERT D., BOERNER S., KEARNEY E., KING J. E., ZHANG K. & SONG L. J. (2014), "Expressing Religious Identities in the Workplace: Analyzing a Neglected Diversity Dimension", *Human Relations*, 67(5), pp. 543-563.
- GILBERT P. (1998), *L'instrumentation de gestion*, Paris, Economica.
- GIRIN J. (1981), *Les machines de gestion*, Paris, Éditions de l'École polytechnique.
- GRIMAND A. (2012), "L'appropriation des outils de gestion et ses effets sur les dynamiques organisationnelles : le cas du déploiement d'un référentiel des emplois et des compétences", *Revue Management Avenir*, 4, pp. 237-257.
- GUILLET O. & BRASSEUR M. (2019), "Le comportement des managers face au fait religieux. Apports de la théorie du comportement planifié", *Revue des Sciences de Gestion*, n°297/298, pp. 11-17.
- HATCHUEL A. & WEIL B. (1992), *L'expert et le système, suivi de quatre histoires de systèmes-experts*, Paris, Economica.
- HÉLIOT Y., GLEIBS I. H., COYLE A., ROUSSEAU D. & ROJON C. (2020), "Religious Identity in the Workplace: A Systematic Review, Research Agenda, and Practical Implications", *Human Resource Management*, 59(2), pp. 153-173.
- HICKS D. A. (2003), *Religion and the Workplace: Pluralism, Spirituality, Leadership*, Cambridge, Cambridge University Press.
- KING J. E., BELL M. P. & LAWRENCE E. (2009), "Religion as an Aspect of Workplace Diversity: An Examination of the US Context and a Call for International Research", *Journal of Management, Spirituality & Religion*, 6(1), pp. 43-57.
- KIRTON G. & GREENE A.-M. (2015), *The Dynamics of Managing Diversity: A Critical Approach*, Fourth edition, New York, Routledge.
- KREINER G. E., HOLLENSBE E. C. & SHEEP M. L. (2006), "On the Edge of Identity: Boundary Dynamics at the Interface of Individual and Organizational Identities", *Human Relations*, 59, pp. 1315-1341.
- LABATUT J., AGGERI F. & GIRARD N. (2012), "Discipline and Change: How Technologies and Organizational Routines Interact in New Practice Creation", *Organization Studies*, 33(1), pp. 39-69.
- LANDRY A.T., GAGNÉ M., FOREST J., GUERRERO S., SEGUIN M. & PAPACHRISTOPOULOS K. (2017), "The Relation Between Financial Incentives, Motivation, and Performance", *Journal of Personnel Psychology*, 16(2), pp. 61-76.
- LATOURE B. (1992), *Aramis ou l'amour des techniques*, Paris, La Découverte.
- LORINO P. (2018), *Pragmatism and organization studies*, Oxford, Oxford University Press.
- LORINO P. (2005), "Théorie des organisations, sens et action : le cheminement historique, du rationalisme à la genèse instrumentale des organisations", in TEULIER R. & LORINO P., *Entre connaissance et organisation : l'activité collective. L'entreprise au défi de la connaissance*, Paris, La Découverte, pp. 54-71.
- MARINOS C. (2018), "Espaces collaboratifs de travail et clubs d'entreprises : des réseaux au cœur des dynamiques collaboratives d'innovation", *Innovations*, 55, pp. 119-141.
- MILLER D. W. (2020), "Faith and Work: Brave New World?", in AKRAM T., RASHID S. (eds) *Faith, Finance, and Economy*, Cham., Palgrave Macmillan.
- MILLER D. W. & EWEST T. (2015), "A New Framework for Analyzing Organizational Workplace Religion and Spirituality", *Journal of Management, Spirituality & Religion* 12(4), pp. 305-328.
- MOISDON J. C. (1997), *Du mode d'existence des outils de gestion*, Paris, Seli Arslan.
- OIRY E. (2011), "Usages imprévus et dynamique des instruments de gestion. Réflexions à partir du cas d'un instrument de gestion des compétences", *Management International*, 15(2), pp. 11-22.

- PASTOR E. R. (2016), "Towards Substantive Equality for Religious Believers in the Workplace? Two Supranational European Courts, Two Different Approaches", *Oxford Journal of Law and Religion*, 5, pp. 255-279.
- PICHAULT F. (2013), *Gestion du changement. Vers un management polyphonique*, Bruxelles, De Boeck.
- PODOLNY J. & PAGE K. (1998), "Network Forms of Organizations", *Annual Review of Sociology*, 24, pp. 57-76.
- POLGE M. (2009), "La recherche "actée" en TPE : l'exemple d'un club de dirigeants", *Management & Avenir*, 30(10), pp. 227-241.
- RABARDEL P. & BOURMAUD G. (2003), "From Computer to Instrument System: A Developmental Perspective", *Interacting with Computers*, 15(5), pp. 665-691.
- RABARDEL P. (1995), *Les hommes et les technologies, approche cognitive des instruments contemporains*, Paris, Armand Colin.
- RENZETTI C. M. & LEE R. M. (1993), *Researching Sensitive Topics*, London, Sage Focus Edition.
- REYNAUD J.-D. (1988), "Les régulations dans les organisations : régulation de contrôle et régulation autonome", *Revue française de sociologie*, 29(1), pp. 5-18.
- SIGGELKOW N. (2007), "Persuasion with Case Studies", *Academy of Management Journal*, 50(1), pp. 20-24.
- SYED J., KLARSFELD A., NGUNJIRI F. W. & HÄRTEL C. (2018), "Introduction", in SYED J., KLARSFELD A., NGUNJIRI F. W. & HÄRTEL C., *Religious Diversity in the Workplace*, Cambridge, Cambridge University Press, pp. 1-37.
- THOMAS D. & ELY D. (1996), "Making Differences Matter: A New Paradigm for Managing Diversity", *Harvard Business Review*, 74(5), pp. 79-90.
- VAUJANY F. X. de (2006), "Pour une théorie de l'appropriation des outils de gestion: vers un dépassement de l'opposition conception-usage", *Revue Management & Avenir*, 9, pp. 159-179.
- VICKERS L. (2016), *Religious Freedom, Religious Discrimination and the Workplace*, Oxford and Portland (Oregon), Hart Publishing.
- WOOLGAR S. & LATOUR B. (1988), *La Vie de laboratoire : la production des faits scientifiques*, Paris, La Découverte.

# Transition “Niches” as Spaces for Re-envisioning the Energy System: The Case of Self-Consumption

By [Élodie Gigout](#), [Julie C. Mayer](#) and [Hervé Dumez](#)  
i3-CRG, École polytechnique, CNRS, IP Paris

Original article in French published in *Gérer & comprendre*,  
September 2021, n° 145, pp. 3-12.

The emerging practice of self-consumption is seen as a potential contributor to the energy transition. But the idea of expanding it on a large scale is surprisingly controversial. Defined as producing and consuming one's own electricity, self-consumption is still struggling to catch on and is the subject of much industry debate in France. In this article, we aim to shed light on this new practice, looking at it through the concept of a transition “niche” (Schot & Geels, 2007), a space for experimentation that, under the right conditions, can help to radically transform an established system. It is easy to identify a niche that led to a systemic transformation after the fact. But while a niche is still a niche, it is likely to be the subject of considerable debate. This is the angle from which we propose studying self-consumption: in the discourse shaping their practices, what are the differing positions of stakeholders, between those advocating controlled expansion of the niche and those looking to transform the system?

In the midst of the climate crisis, France is frequently criticised for its “inertia” in making the energy and green transition. While expanding renewable energy sources is a key plank in energy transition roadmaps, it must go hand in hand with bolder targets for distributed energy generation and new consumption practices. One such practice is self-consumption, which is where an electricity consumer generates their own power. But although France has had a legal framework for it in place since 2017 (Act of 24 February 2017 ratifying the Order of 27 July 2016), the practice has yet to really catch on, with the country's energy industry debating how to scale it up and whether it should be. To understand the issues surrounding the expansion of self-consumption, we propose studying it using the concept of a “transition niche”: an emerging practice that can have an impact on a system, the transition of which is based on changes to multiple interconnected levels and dimensions.

Beginning with an overview of the main challenges involved in transitioning France's electric power system, we will identify self-consumption as a “transition niche” and demonstrate the interest of studying it via the discourse of electricity stakeholders. We will then present our analysis, by identifying the different representations of self-consumption as a niche that can have an effect, whether positive or negative, on different dimensions of France's electric power system (specifically, the regulatory/political, economic/commercial, technological and social dimensions).

We will show that, in addition to the usual factors (technological, economic and regulatory) used to identify a niche (Turnheim & Geels, 2019), collective spaces of commentary and debate also influence its expansion and its integration in the transition process.

## The “socio-technical” transition of France's electric power system

Although self-consumption has only recently been included in transition roadmaps, France's electric power system has a long history of change (Dunsky, 2004; Raineau, 2011), undergoing numerous transformations since it was built in the late 19<sup>th</sup> century (Beltran & Carré, 2017). The result is a centralised power system that, until it was opened up to competition in the 2000s, was run by a single public corporation, EDF, in charge of electricity generation, supply, transmission and distribution. With the development of France's substantial nuclear power programme, undertaken to ensure the country's energy independence and to control costs, the electric power system was progressively built up around large plants, supplying the entire country via a power transmission and distribution system at a single rate (the French principle of *péréquation tarifaire*, or tariff equalisation).

As electricity cannot be stored, operating the power system requires a constant balancing of supply and demand. France has a robust system in that

consumption is aggregated, which means demand can be smoothed and generation can be more easily adapted – an argument in favour of a centralised system. It is also a highly reliable system with a very low failure rate.

That said, the transformations France is currently faced with appear to be unprecedented in scope (Rüdinger et al., 2017). There are two inherent limitations of the dominant power system: the scarcity of conventional fossil and fissile fuels, and climate change. As demand continues to steadily rise, these limitations suggest that the system is facing a new energy transition. While there is little remaining debate about the facts of the climate emergency, the 2015 Energy Transition and Green Growth Act set out a roadmap with targets for reducing greenhouse gas emissions and increasing the share of renewables in the energy mix, putting the electric system front and centre in the energy transition.

These targets involve moving away from a centralised power generation system, managed by a small number of dominant players with clearly defined roles, to a decentralised system that is more difficult to manage due to the intermittent nature of renewable energies, involving numerous stakeholders with redistributed roles and a heavy focus on digital technology. Faced with this combination of uncertainty and complexity, the “inertia” preventing the power system from making the transition is the subject of great political debate in France. Some see it as a problem of an entrenched centralised system (Boutaud, 2013), where the dominant players are holding back its transformation (Evrard, 2014). But according to a fact-finding mission on the obstacles to the energy transition launched by France’s National Assembly in summer 2018, the issue appears to be more complex than that.

To study the issue, we propose considering that the electric power system is undergoing a “socio-technical” transition, as conceptualised by Schot and Geels (2007). A transition is a process by which a system, i.e. an organised set of structures and actors, shifts from state A to state B via change processes on multiple levels (Geels, 2010). A transition is “socio-technical” when there are dual dimensions of technological change and structural change to the rules, beliefs and routines underlying the system (Geels, 2011). At macro level, the “landscape” in which the socio-technical regime is embedded can undergo change: structural transformations of the environment over the long term force the regime to adapt. At micro level, “niches”, i.e. spaces for technological or social innovation, can develop, initially on the margins of the system before potentially disrupting or transforming the existing regime. According to Geels’s model, the transformations of a regime in transition take place on multiple dimensions: technological, economic/commercial, regulatory/political and social (Geels, 2002). The technological and regulatory dimensions are often identified as foundational to the transformation of the energy system.

But Geels’s approach is retrospective. Although it is able to identify niches in hindsight, after the transition has occurred, it is much more difficult to determine mid-transition whether a given practice is a niche

that will lead to a major systemic change, or whether it will remain relatively minor. In our view, the case of self-consumption appears to illustrate the fact that, during the transition process, a niche is a space for debate that cannot be used to predict the outcome of the transition.

## Self-consumption: a controversial transition niche

Self-consumption is when an individual, entity or community consumes electricity that they generate themselves, typically via a photovoltaic (solar power) system. While France has an established photovoltaic industry (ADEME, 2016), self-consumption is a relatively new practice. Until 2016, existing laws only allowed users to sell any excess solar energy they produced to EDF under a specific buyback agreement. The option to practise self-consumption was formalised with the Act of 24 February 2017, which sets out two scenarios: “individual” self-consumption, where power is produced and consumed by the same individual, and “collective” self-consumption, where power is shared between one or more producers and one or more consumers. (See Table 3 on page 18 for a breakdown of the differences between the two). While this legislation officially introduced the practice of self-consumption in France, it remains in its early stages, with only 16 “collective” self-consumption setups registered so far.

Despite its fledgling status, self-consumption has had strong proponents from the start, as well as its share of sceptics and doubters. France’s Energy Regulation Commission (CRE) organised a wide-ranging consultation exercise between 2017 and 2018 in an attempt to more clearly define the technical and legal scope of the practice as well as principles for pricing and subsidies. Everyone was able to express their views on the nature and extent of the potential “disruptions” posed by self-consumption.

Given its status as an emerging and experimental practice, self-consumption can be considered a “niche”: it meets the criteria of a practice that departs from the existing system (or “regime”), is disrupting the system and, under certain conditions, is working to transform it (Geels, 2011). At present, however, it is more the subject of debate as to what it could be as opposed to something concrete. This calls into question the traditional conception of a niche. First, while the literature suggests that the experimental nature of a niche makes it essentially technological in nature (Schot & Geels, 2007), self-consumption is experimenting with solutions that encompass more than just technology, such as uses and behaviour associated with the practice, its pricing principles and the governance of the regime. Second, the current definitions of a niche conceive of it as a “protected” space (Smith & Raven, 2012), relatively sheltered from the institutional pressures of the existing regime (Turnheim & Geels, 2019). Yet experimentation with self-consumption has sparked a host of questions about how it might disrupt the existing system. We propose studying this discourse to understand how this “transition niche” is perceived, and in fact socially constructed, by stakeholders.

## Methodology

We conducted an analysis of the discourse on self-consumption in the electric power industry. There were three phases.

First, we conducted a documentary analysis on the concept of self-consumption (legislation, reports, press review) in early 2018 to understand the context of the subject matter and put together a list of interview subjects. We therefore focused on industry stakeholders that had explicitly and publicly spoken out on the matter during the consultation exercise led by the CRE.

Second, we conducted a series of interviews (16 interviews lasting an average of 80 minutes, of which 14 were recorded and transcribed) in mid-2018, after the consultation exercise, when the debate was at its most intense and stakeholders were taking clear positions on the matter. We used this data to identify the structural dimensions of the discourse around self-consumption. In light of the definition of a niche as a space for technological and social experimentation (Schot & Geels, 2007), we analysed what appear to be topics of experimentation, definition and debate. A first attempt at open coding revealed a large majority of topics that were more systemic in nature. In other words, in response to the question “In your opinion, what is self-consumption and how do you see it potentially expanding?”, interview subjects instead discussed “What would the future electricity system look like if self-consumption were to expand?”. After several attempts at coding, we used multi-thematic coding to group the different points of tension in the discourse (Ayache & Dumez, 2011), reflecting the different dimensions of the existing regime (here

the current electric power system) mentioned by the interview subjects.

Third, we monitored developments in the legislation and the official discourse of the interview subjects (press releases, media appearances) between 2018 and 2020, in order to identify their position over the entire period. This enabled us to triangulate the data collected during the interviews.

## Industry discourse on self-consumption

One outcome of self-consumption would be a change in the roles of electricity industry stakeholders, who as a result may or may not be in favour of its expansion. However, there are still many areas of uncertainty as to the impact it would have on these stakeholders. The industry discourse has created a “theatre” of discussion and debate, gradually building and shaping this still experimental niche.

### Three positions

The scale of the debates that followed the CRE consultation in 2018 illustrates to what extent the issue of the expansion of self-consumption, particularly the “collective” category, has divided the electric power industry. When a consumer pays their electricity bill, a portion goes towards funding the operation of the system, but also toward the profits of a certain number of participants in the value chain: the electricity producers, the transmitter, the distributor, the suppliers, etc. Our study involves a sample of these stakeholders, presented in Table 1.

Stakeholder	Role in the electricity system
Commission de régulation de l'énergie (CRE)	Regulator. Independent administrative authority.
Réseau de transport d'électricité (RTE)	Transmission system operator (high and ultra-high voltage lines). Monopoly, subsidiary of EDF (50.1%).
Enedis	Distribution system operator (low and medium voltage lines). Near monopoly, subsidiary of EDF (100%).
Enercoop	100% renewable energy supplier. Cooperative, contracts directly with energy producers.
EDF Énergies nouvelles (EDF EN)	Renewable energy electricity producer. Subsidiary of EDF (100%).
Total Solar	Solar power project developer. Subsidiary of Total.
Enerplan	France-wide multi-sector employers' association for the solar power industry. Members belong to the manufacturing, construction, trade and service industries.
Groupement des particuliers producteurs d'électricité photovoltaïque (GPPEP)	Association of individual photovoltaic electricity producers (more than 9,000 members).
Fédération nationale des collectivités concédantes et régies (FNCCR)	Federation of local authorities and their government-funded institutions responsible for organising and/or operating certain utilities (more than 800 members).
Région Occitanie	Local authority.
Bouygues Immobilier	Property developer.
Embix	Start-up specialising in smart grid solutions.
Schneider Electric	Provider of digital power and automation solutions.

Table 1. Stakeholders in the electric power industry



Many stakeholders depend on the end user's electricity bill, but self-consumption would disrupt the underlying formula, with self-consumers still continuing to use electricity from the power grid when they are not getting enough from their solar panels. Faced with the uncertainties associated with these disruptions, stakeholders in the existing system have adopted and defend different positions, based on their interpretation of the situation (Reverdy, 2010). From the way they express these interpretations, their positions can be divided into three categories: committed, ambivalent and hesitant. These categories reflect the content of stakeholders' explicit discourse, as well as their interests and roles in the electric power system.

At one end, the "committed" camp includes stakeholders with a direct interest in the expansion of solar power and self-consumption products and services, as a potential growth vector of the photovoltaic industry. Solar energy employers' association Enerplan is the leading stakeholder having demonstrated a strong commitment in favour of expanding self-consumption and supporting it through various tax and economic incentives.

At the other end, the "hesitant" group includes those who, at national level, oversee the operation of the electric power system, such as the RTE and the CRE. It should be noted that these stakeholders are not against self-consumption, but advocate controlled expansion of the practice.

In between these two positions, those categorised as "ambivalent" appear to be partially on board with some aspects of these changes but are mindful of the consequences, considering there to be still too much uncertainty to take a firm position.

Table 2 categorises the electricity stakeholders based on the position expressed in their discourse as part of this study.

These positions, which are relative, reflect our own analysis of the discourse collected during the study and do not necessarily imply that these stakeholders are entrenched in their position, that their strategy is limited to that position, or that there are not other positions held within the organisation.

### The discourse around the prospect of an expansion of self-consumption

We will illustrate these three positions by studying the stakeholders' discourse around four dimensions of the system that are currently in transition and that the expansion of self-consumption would disrupt: regulatory/political, economic/commercial, technological and social.

Position	Stakeholder
Hesitant	CRE, RTE
Ambivalent	Enedis, Enercoop, EDF EN, FNCCR, Embix
Committed	Enerplan, GPPEP, Bouygues Immobilier, Schneider Electric, Total Solar, Région Occitanie

Table 2: Summary of positions reflected in the discourse of electricity stakeholders vis-à-vis the expansion of self-consumption

### The regulatory/political dimension

Traditionally, France's electricity pricing formula has been based on the principle of *péréquation tarifaire* (tariff equalisation), which ensures "solidarity", or a fair distribution of the cost among individuals and regions. Accounting for roughly a third of the price per kWh, the public transmission system access tariff (TURPE) goes toward funding the system operators to ensure access for all users. A self-consumer who generates their own electricity will either not use the system at all or use it very little, only when they generate more than they consume or consume more than they generate. When they are not using the system, should they have to pay this contribution toward the its management? And the same issue applies to taxes on electricity, which account for another third of the price of a kWh consumed from the system, in terms of a smaller tax base and lower contributions from self-consumers. Deciding whether self-consumers should be treated the same as other consumers opens the door to reconsidering the principles of electricity pricing, a hotly debated subject.

Indeed, pricing is the focus of much debate. It is a major factor in the decision to expand or limit the practice of self-consumption.

These mechanisms reveal a two-tier political and regulatory transition. At one level, initiatives are being introduced to encourage the expansion of self-consumption and related technologies (tax exemptions, investment incentives, calls for tender, etc.). At another level, there is a high degree of hesitancy from the regulator in the face of the uncertainties that we will outline in the following sections. However, the speed at which France makes the transition on the policy and regulatory fronts is not a trivial concern. To understand this, it is necessary to contextualise France as one actor among others with varying degrees of power and influence in the transition race. On the one hand, France wants to win the competition against powers such as China, which will require agility and speed. On the other hand, regulations tend to tightly control the expansion of self-consumption. This is where stakeholders' positions diverge in opposite directions.

The "hesitant" group advocates controlled expansion, so that any real impacts on the system can be gradually integrated:

"As for support mechanisms, they must be adapted to the wide range of situations and allow for an optimal and controlled expansion of self-consumption. [...] However, support for self-consumption must not impede the development of large ground-mounted

solar power plants, which are a major, and affordable, contributor to the objectives of renewable energy expansion”<sup>(1)</sup> (CRE report, 2018).

“Committed” stakeholders would like to see a faster expansion of self-consumption in order to develop solutions to the issues it raises, which implies protections and incentives:

“To ensure it doesn’t thrive right away, it will be prevented from expanding. That’s the revolutionary recommendation that has emerged from the CRE. [...] It’s as if they’re wearing bifocals: magnifying anything that might pose a problem, and minimising the rest. [...] Our recommendation is to say: ‘we’re still early days on this’. If the only thing we need is to not be taxed and to not to receive any subsidies in return, it really is a free-market system. And then: ‘at first, there will be no macroeconomic effect, let’s have tax exemptions for the CSPE,<sup>(2)</sup> the TICFE,<sup>(3)</sup> for 15 years” (interview with Enerplan).

In between these two positions are the “ambivalent” stakeholders, who see the change as presenting opportunities in other areas, but also risks.

### The economic/commercial dimension

As a practice, there are two main facets to self-consumption: new manners of production (local, decentralised) and new manners of consumption. The trend toward decentralisation and the arrival of activist consumers<sup>(4)</sup> (Cochet, 2000) is driving a transformation of the economic regime that France’s electricity system is built on. Self-consumption changes the value-creation mechanisms around electricity. First, it is forcing

electricity suppliers to reconsider their economic model:

“Historically, we have been buyers and sellers of electricity, but all electricity suppliers are asking themselves: ‘How do we break out of this single-product model? How can we diversify? And how can we offer new services?’ [...] What’s at stake for us as suppliers, and for all suppliers, is that these new services are in almost direct opposition to the core of our business” (interview with Enercoop).

It also concerns the transmission and distribution of electricity, the first managed by RTE and the second mainly by Enedis. For these stakeholders, there is an additional layer of uncertainty: What will the system of the future look like? How much of it will be made up of renewables? Where does self-consumption fit in? Or the move to moderate energy consumption? How will these new forms of electricity generation and consumption be distributed geographically? How will regional integration work (smart cities, positive energy areas, etc.)? The answers to all of these questions have impacts on both the infrastructure and the operation of electricity systems, as well as their funding model:

“Wide-scale expansion of self-consumption must be planned for, monitored and controlled so as not to jeopardise how the system is funded and operated. We also need to be careful about the pricing of self-consumption so as not to generate costly deadweight effect for the community. Electric power systems are based on the principle of solidarity, i.e. fair distribution of cost for the community. As an energy supply method, self-consumption must be an option that works alongside others, but it can’t disrupt the overall balance of solidarity” (interview with RTE).

The issue of the coverage area for collective self-consumption (see Table 3) is particularly significant, since it means limiting the arrangement to users of a same electrical substation, which is often a single building. The “committed” group say that this limits the size of the setup and minimises economies of scale, whereas the “hesitant” group say they are considering the reality of electricity flows, which always travel to the nearest exit point.

<sup>(1)</sup> Translator’s note (TN): All citations attributed to stakeholders have been translated from French.

<sup>(2)</sup> TN: Contribution to the public electricity service (*contribution au service public de l’électricité*).

<sup>(3)</sup> TN: Domestic consumption tax on electricity for end-users (*taxe intérieure sur la consommation finale d’électricité*).

<sup>(4)</sup> TN: Loose translation of “*consomm’acteur*”, a portmanteau of “consumer” and “actor” (in the sense of “participant” or “activist”).

		Individual	Collective
	Producer	An individual person	One or more persons
	Consumer	The same individual person	One or more persons
	Structure	N/A	Producers and consumers grouped together under an organising entity to distribute the self-produced electricity
	Coverage area	N/A	Participants sharing the same electrical substation
Indirect support	Exempt from TURPE	Yes	No
	Exempt from CSPE	Yes	No
Direct support	Investment incentive	Yes	

Table 3: Comparison of individual and collective self-consumption

Although self-consumption is challenging the traditional electricity market transaction mechanisms, it is also a source of new monetisable needs. We are seeing the emergence of new types of service providers, such as “aggregators”, which balance electrical capacity from decentralised production sites to ensure more flexibility between supply and demand. Energy producers are beginning to expand their commercial offerings for self-consumers, as well as support services to help consumers optimise their bills.

There is also the ability of a niche to gain strength through connections with other niches:

“Because the issue for the industry, with home automation, electric vehicles, smart charging and vehicle-to-grid and vehicle-to-building charging, is to make the best possible use of local flexibility. That’s where the real challenges are. As long as we stick with a super-centralised system, without any incentives for system operators or distributors to optimise how they work, and covering all their costs, well... we’ll stay stuck in 1946. We won’t join the 21<sup>st</sup> century, like other countries are doing” (interview with Enerplan).

At regional level, there are also challenges in terms of supporting the growth of SMEs, new entrants in the energy sector, and issues of savings on their own (often high) energy bills, but as owners of the systems they are not indifferent to the potential additional costs. The “ambivalent” group remain moderate:

On the issue of changes in electricity pricing: “It’s another area we’re keeping an eye on, to see signs that there will be a certain level of equalisation, and that we don’t end up completely overhauling the system. [...] The overall vision of elected officials is to maintain a certain level of service quality [...] so that we maintain an optimum level, and so that we see a return of small and medium-sized enterprises and industries to the regions” (interview with the FNCCR).

In view of these uncertainty factors, the current picture of the economic consequences at national level remains unclear.

### The technological dimension

Large-scale expansion of self-consumption would involve spikes in production at times of day and periods of the year of off-peak consumption. While self-consumption represents a “grow your own” option for electricity consumption and a way to lower household electricity bills – France’s environmental and energy control agency, ADEME, estimates these savings to be between 15% and 25% (ADEME, 2018) – there are also other possibilities. For consumers who want to do more than just cut costs and sell electricity back to the grid, then an energy storage solution may become necessary:

“Storage would allow holding a surplus of electricity over a relatively short period, and this relatively short period is what is called a ‘power peak’. If you absorb the peak, there is no need to adapt the grid or the nuclear plants to handle it” (interview with GPPEP).

Storage is therefore the second major technological innovation underlying an expansion of self-consumption. It should be noted that storage could also help avoid having to make power grid reinforcements (and the associated costs) and help secure the supply of electricity in edge-of-grid areas. Although a range

of solutions are under development (batteries, use of electric vehicles, etc.), storage remains a major uncertainty variable due to cost. While it would seem to be an essential innovation, the CRE has noted its absence in existing projects:

“The consultation led by the CRE revealed that storage is still rare in self-consumption setups” (CRE report).

For some private stakeholders, the regulatory framework is to blame:

“Right now in France, the regulatory framework penalises – or rather does not encourage, to put it more tactfully – the installation of batteries. Not at all. Economically, it doesn’t make sense to have a self-consumption setup with a battery. Because basically, if you produce electricity locally but don’t consume it yourself, you get compensation for selling it back to the grid” (interview with EDF EN).

The third technological innovation is digital. Eventually, technology would enable self-consumption to be not just a way to earn extra income for a handful of households, but rather a broader restructuring of the electric power system. To this end, “smart” technologies (smart grids, micro-grids, smart meters, blockchain) would allow energy consumption and production to be managed in real time, by distributing locally produced electricity in response to needs and, crucially, by adapting needs to production.

The degree to which self-consumption would disrupt the existing system depends on whether it is used in conjunction with storage or with smart technologies. If there is a massive expansion of renewable energies, including via self-consumption, the production of this energy would be intermittent by nature (sun, wind) and a new method would be needed for balancing supply and demand: without storage, consumption would need to adapt to the constraints of intermittent power. This ability to manage electricity demand is known as “demand-side management”. But once again, there are not many projects factoring in this variable:

On the topic of calls for tender in the region: “In our system for scoring bids, there is a criterion for demand-side management [...] but very few projects take it into account, and if they do, it’s with thermal energy storage” (interview with Région Occitanie).

### The social dimension

While there is mounting debate on the technological, regulatory and commercial aspects of self-consumption, little is being said about the social transformations that it could lead to. We have therefore identified some unexplored social transition aspects.

First, the expansion of self-consumption raises the question of social acceptance: Would everyone want to become a self-consumer? Intuitively, the “grow your own” approach to electricity would seem to have its appeal. However, the issue of underestimating social acceptance has more than once taken France’s electricity stakeholders by surprise (Chamaret et al., 2020): Linky smart meters, resistance to high-voltage lines, etc. There is also the recurring argument that self-consumption could lead to individuals taking control of their production and consumption by making them visible. Consumers would become activist consumers or consumer-producers (Cochet, 2000):

“Simply by looking at the graph, with your electricity production on one side and your consumption on the other, and when you notice that you’re consuming more than you’re producing, you think, ‘Can’t I just try to consume less?’ And so the effect is to spur people to use less energy” (interview with GPPEP).

If this kind of consumer awareness occurs, it raises the question of whether it will affect how people use energy from the grid:

“So the relatively stable portion, that’s handled internally, it’s the consumer who manages that. And then what they give us, or what they consume, is only what we call ‘la dentelle’,<sup>(5)</sup> in other words, anything over and above that. And that’s what’s the most complicated for us to manage. First, because it’s a smaller volume but the same amount of management [...]. And so when we collect little bits of “dentelle” here and there, based on cloud cover, etc., and we no longer have the base, that has a major impact for us. [...] It doesn’t mean it’s not worthwhile, but it does mean we really need to think about how to do it” (interview with Enercoop).

Lastly, this new method of energy consumption requires considering the effects that its expansion could have in terms of inequality, in two respects. The first is that self-consumption may not be an accessible option for lower-income households, causing them to contribute more to funding the system than self-consumers, who, in the case of the “individual” category of user, contribute nothing at all for the self-generated electricity they consume (TURPE and tax exemptions, see Table 3). The second is regional inequality in terms of disparities in the number of sunlight hours. In both cases, it is the principle of tariff equalisation and solidarity between individuals and regions that is at issue. Maintaining equalisation is pitted against recreating solidarity through other mechanisms:

“We were told, ‘Yes, but nationwide solidarity...’ Agreed! But what if we came up with new models? That’s what Enerplan is now proposing, to open up contracts for selling surplus electricity to community solar programmes, for example. So that when there is extra electricity, it can be redirected to low-income consumers” (interview with Enerplan).

Our analysis of the discourse around self-consumption reveals that the experimentation taking place within this niche is not in relation to the technology itself, or the associated practices, but the vision of the entire system and how it might evolve. However, there is nowhere near a consensus on this vision.

### The role of stakeholders in gradually expanding the definition of self-consumption

The power struggle between electricity stakeholders following the CRE consultations in 2018 has persisted, particularly with regard to the rules around collective self-consumption, with the individual category benefiting from enough support measures and tax exemptions to be economically viable – and these measures have not been challenged (see Table 3). Changes to the definition

have gradually been fuelled by the discourse behind the various positions described above. Beginning in 2018, the discourse of the “committed” camp found a policy position in the solar power plan of the Ministry for the Ecological Transition, which uses a wider coverage area for collective self-consumption, does not limit support measures for facilities (< 500 kWp) and opens up the possibility of third-party investors. Similar developments are underway at European level with the EU Directive of 11 December 2018. In 2019, the PACTE<sup>(6)</sup> Act and the associated ministerial order permanently broadened the scope of collective self-consumption and increased the cumulative power limit for facilities. The category is still not exempt from TURPE or CSPE, to the dismay of the solar power industry. To address the remaining financial constraints, the regions have begun to play a major role by granting subsidies to collective self-consumption projects.

At each of these stages, interventions by stakeholders to influence politicians and lawmakers have been the subject of controversy. For example, the most ardent of the “committed” camp, such as Enerplan, complain that the CRE has been holding things back, calling for limits every time new measures are proposed. Other stakeholders, like this manager of a solar power design office who has published numerous articles on the subject, have also spoken about their own lobbying efforts:

“It’s taken time, planning, DGECC<sup>(7)</sup> consultations and meetings with MPs and senior DGECC<sup>(7)</sup> officials to get here. That’s a fact” (A. Joffre, TECSOL, Vertsun, 26 September 2019).

This debate between the “committed” and “hesitant” sides has also involved other, more direct means of action, such as court challenges. Enerplan has twice petitioned the *Conseil d’État*, France’s highest administrative court: once in 2017, against a circular from the economy and finance ministries (Bercy) on CSPE exemptions, and a second time concerning the CRE’s 2018 decision on optional TURPE pricing specific to collective self-consumption setups, a measure considered to be “punitive”.

Furthermore, as details relating to collective setups have been clarified, there has been increasing engagement on the matter from stakeholders whose position was initially less clear or more reluctant. This has resulted in the development of commercial offerings and communication campaigns designed to raise consumer awareness about collective self-consumption. At least two of the main suppliers, EDF EN and Total Direct Énergie, now have full-scale self-consumption offerings, from project assessment to implementation, including support services and smart management solutions. EDF was also involved in one of the 20 collective setups that have materialised so far. Enercoop produces educational webinars on the subject. Enedis was involved in the first collective setup and is supporting five more projects at national level.

<sup>(5)</sup> TN: Literally, “lace”.

<sup>(6)</sup> TN: Business Growth and Transformation Action Plan.

<sup>(7)</sup> TN: Directorate General for Energy and Climate.

Although self-consumption is yet to be practised on a large scale, it is continuing to expand through a space of discussion and debate. In the following section, we will discuss the theoretical and practical implications of these results.

## Discussion and conclusion

The persistent debates on self-consumption demonstrate why “transition niches” should be understood as spaces for commentary and discourse, as well as spaces for technological experimentation. Within these spaces, stakeholders observe the uncertain development of a still-emerging practice, leading them to come together to deliberate on how the system could be reconfigured. While Geels’s approach cannot be used to identify whether, in principle, an activity is a niche or not, it can nevertheless serve as a guide for discussion and analysis: it is an “orienting theory” as defined by Whyte (1984). As such, we believe it could be used for future research on the possible disruptions of other identified niches, such as smart technologies (smart meters, connected homes, etc.) or new forms of storage or mobility. Our analysis of self-consumption as a transition niche has allowed us to draw several theoretical and practical insights on how transitions take place.

First, our study of the discourse surrounding self-consumption has helped to clarify the nature of the controversy over its expansion. Our analysis illustrates the degree to which the transition to which a niche belongs requires deconstructing all of the existing structures – technological, political and regulatory structures, as well as social, economic and commercial ones (Schot & Geels, 2007). While there is consensus among stakeholders as to the disruptive potential of self-consumption, where viewpoints diverge or falter tends to be on the issue of how to find a new equilibrium: whereas the positions of stakeholders are relatively clear in the regulatory and economic dimensions, they are less certain or remain unspoken in the social and technological dimensions. This conclusion suggests further exploring the mechanisms at work in the reconstruction of a shared representation of a system in transition, taking into account the uncertainties present in the different dimensions of the system (Reverdy, 2010).

Our analysis of the discourse on self-consumption as a niche, in terms of its effect on the dimensions of the regime, reveals the existence of particularly strong uncertainty in the social dimension. This is indicative of the difficulty of taking practices into account in navigating the energy transition: despite experience from past transformations of large systems, analysis is more focused on the technological aspects, with little attention paid to practices. Yet practices are key to analysing the trajectory of self-consumption. Studying energy consumption from a theory of practice approach (Warde, 2005) seems particularly promising, in that it offers a way of moving away from an analysis of “having” toward an analysis of “doing” (Dubuisson-Quellier & Plessz, 2013). In the matter at hand,

from a consumer’s perspective, the consumption of electricity would appear to have more to do with “doing” something (through their use of the electricity) than “having” something (i.e., owning a number of kilowatt hours). Reckwitz (2002) defines a social practice as a routinised type of behaviour consisting of several interconnected elements: bodily activities, mental activities, knowledge, know-how, and things and their use. An energy consumer remains relatively unaware of the practices associated with energy use, which are based on both habit and the structures in which the uses are established (Gram-Hanssen, 2011; 2014).

From a management perspective, a better understanding of electricity consumption practices would help identify catalysts for change. The stakeholder discourse we have analysed contains practically no mention of the practices of self-consumers. As it stands, the discourse is based on relatively vague depictions, even inventions, of the self-consumer. Are there specific obstacles and opportunities based on energy consumption practices? Answering such questions would help to better envision how to structure, initiate and manage the trajectory for the expansion of self-consumption (Dubuisson-Quellier, 2016; Dumez & Renou, 2018).

In conclusion, the concept of a “niche” has helped to better define the disruptive nature of the phenomenon of self-consumption. It has revealed the difficulty in building a shared representation of the new equilibrium that needs to be found, in light of the uncertainties associated with each dimension and how they interact. Furthermore, the case of self-consumption suggests the need to take a theory of practice approach to the study of socio-technical transitions, in order to better understand how a technical innovation turns into a social transformation, and to understand the role that such transformations could play in the transition of the system as a whole.

---

The authors would like to thank the participants in the AEGIS writing workshop who helped to improve this text, as well as the two anonymous reviewers.

## Bibliography

- ADEME (2016), “Mix électrique 100 % renouvelable ? Analyses et optimisations”, étude.
- ADEME (2018), “L’autoconsommation d’électricité d’origine photovoltaïque”, avis, février.
- AYACHE M. & DUMEZ H. (2011), “Le codage dans la recherche qualitative une nouvelle perspective ?”, *Le Libellio d’Aegis*, 7, n°2, pp. 33-46.
- BELTRAN A. & CARRE P. (2017), *La vie électrique. Histoire et imaginaire (XVIII<sup>e</sup>-XX<sup>e</sup> siècle)*, Paris, Belin.
- BOUTAUD B. (2013), “Les énergies renouvelables, énergies des collectivités territoriales?”, *Annuaire des Collectivités Locales*, 33(1), pp. 195-204.
- CHAMARET C., STEYER V. & MAYER J. C. (2020), “‘Hands off my meter!’ When municipalities resist smart meters: Linking arguments and degrees of resistance”, *Energy Policy*, vol. 144, September, 11156.

- COCHET Y. (2000), *Stratégie et moyens de développement de l'efficacité énergétique et des sources d'énergie renouvelables en France : rapport au Premier ministre*, Paris, La Documentation française.
- DUBUISSON-QUELLIER S. & PLESSZ M. (2013), "La théorie des pratiques. Quels apports pour l'étude sociologique de la consommation ?", *Sociologie*, 4(4), pp. 451-469.
- DUBUISSON-QUELLIER S. (2016), *Gouverner les conduites*, Paris, Les Presses de Sciences Po.
- DUMEZ H. & RENOU S. (2018), "Les énergies renouvelables existent-elles et peut-on piloter la transition énergétique ?", *Gérer et Comprendre*, n°134, décembre, pp. 3-13.
- DUNSKY P. U. (2004), "La révolution électrique en cours : portrait de l'émergence d'une nouvelle architecture dans les pays industrialisés", *Vertigo-la revue électronique en sciences de l'environnement*, 5(1), pp. 1-13.
- EVARD A. (2014), "Les énergies renouvelables et l'électricité", *Écologie & politique*, n°2, pp. 67-80.
- GEELS F. W. (2002), "Technological Transitions as Evolutionary Reconfiguration Processes: A Multi-Level Perspective and a Case-Study", *Research Policy*, 31(8-9), pp. 1257-1274.
- GEELS F. W. (2010), "Ontologies, Socio-Technical Transitions (to Sustainability), and the Multi-Level Perspective", *Research Policy*, 39(4), pp. 495-510.
- GEELS F. W. (2011), "The Multi-Level Perspective on Sustainability Transitions: Responses to Seven Criticisms", *Environmental Innovation and Societal Transitions*, 1(1), pp. 24-40.
- GRAM-HANSEN K. (2011), "Understanding Change and Continuity in Residential Energy Consumption", *Journal of Consumer Culture*, 11(1), pp. 61-78.
- GRAM-HANSEN K. (2014), "New Needs for Better Understanding of Household's Energy Consumption – Behaviour, Lifestyle or Practices?", *Architectural Engineering and Design Management*, 10(1-2), pp. 91-107.
- RAINEAU L. (2011), "Dossier "Adaptation aux changements climatiques" - Vers une transition énergétique ?", *Natures Sciences Sociétés*, 19(2), pp. 133-143.
- RECKWITZ A. (2002), "Toward a Theory of Social Practices: A Development in Culturalist Theorizing", *European Journal of Social Theory*, 5(2), pp. 243-263.
- REVERDY T. (2010), "Assumer les incertitudes dans un marché en transition", *Revue française de Gestion*, 2010/4, n°203, pp. 101-117.
- RÜDINGER A., COLOMBIER M., BERGHMANS N., CRIQUI P. & MENANTEAU P. (2017), "La transition du système électrique français à l'horizon 2030 : Une analyse exploratoire des enjeux et des trajectoires", *Institut du développement durable et des relations internationales*, pp. 1-38.
- SCHOT J. & GEELS F. W. (2007), "Niches in Evolutionary Theories of Technical Change", *Journal of Evolutionary Economics*, 17(5), pp. 605-622.
- SMITH A. & RAVEN R. (2012), "What is Protective Space? Reconsidering Niches in Transitions to Sustainability", *Research Policy*, 41(6), pp. 1025-1036.
- TURNHEIM B. & GEELS F. W. (2019), "Incumbent Actors, Guided Search Paths, and Landmark Projects in Infra-System Transitions: Re-thinking Strategic Niche Management With a Case Study of French Tramway Diffusion (1971–2016)", *Research Policy*, 48(6), pp. 1412-1428.
- WARDE A. (2005), "Consumption and Theories of Practice", *Journal of Consumer Culture*, 5(2), pp. 131-153.
- WHYTE W. F. (1984), *Learning From the Field: A Guide From Experience*, Thousand Oaks (CA), Sage.

# Exploring the Practice of Coaching

By Magali Ayache

Thema, CY Cergy-Paris Université

and Hervé Dumez

CRG-i3, École Polytechnique, CNRS, IP Paris

Original article in French published in *Gérer & comprendre*, September 2021, n° 145, pp. 46-54.

The coaching market is booming. Drawing on a series of interviews with freelance coaches, coaches employed by private coaching firms and internal corporate coaches, as well as on a non-participating observation of two group mentoring sessions, this paper will examine the paradox of a practice that is highly structured – i.e. generally unfolding over ten sessions, including two tripartite meetings involving the coach, the client and a sponsor from the client's company – yet lacking a theoretical grounding (while pulling from a multitude of fields, such as psychology and neuroscience). We posit that the origin of this paradox lies in the commercial nature of the coaching relationship, which would explain why the rigorously structured practice is compatible with a multitude of approaches that coaches are free to incorporate into their work.

In the space of about ten years, the practice of coaching has become widespread in the corporate world. The International Coaching Federation (ICF), the oldest international association of professional coaches, had 1,500 members in 1999, 16,000 in 2011 and 42,700 in 2020, spread across 140 countries. Practitioners produce a vast body of literature to market their coaching services to businesses and potential clients. There is an equally extensive body of critical literature, some of which argues that capitalism increased performance pressure so much that it caused workplace stress to skyrocket, and then put forth coaching as a solution to the very problems it created (Guilhaume, 2009; Fatien & Nizet, 2011; Fatien Diochon & Nizet, 2012). Furthermore, there is a body of scientific literature that tries to find a middle ground, exploring how coaching can be used effectively (Hackman & Wageman, 2005) or to evaluate the impact of the practice (Passmore & Fillery-Travis, 2011; Theeboom et al., 2014), which provides a functionalist analysis of coaching (Nizet, 2012), or that seeks to establish a solid theoretical grounding for it (Arnaud, 2003; Vanheule & Arnaud, 2016).

The practice itself seems as diverse as the discourse surrounding it (Fatien, 2008), as if there is no consensus on how to define coaching. During an interview conducted as part of our research, one coach described it, rather surprisingly, like negative theology:<sup>(1)</sup>

“Coaching is mostly defined by what it isn't. It's not therapy, training or advice. But at the same time, it's kind of all those things.”<sup>(2)</sup>

How, then, to explore the practice as an intermingling of different discourses and actions (Schatzki, 2008; Gherardi, 2019) through its practitioners and their clients? It is a challenge, as the profession's codes of ethics dictate that coaching sessions remain strictly confidential. This means that researchers are only allowed to observe the practice indirectly. Due to these limitations, we conducted a series of interviews with coaches and attended two group mentoring sessions as non-participating observers in an attempt to define the practice. (It should be noted that while mentoring is similar to coaching, it differs somewhat in that mentors tap into their professional and personal experience to guide their mentees, whereas coaches do not dispense advice.)

We analysed the interview summaries and observation notes using the evenly suspended attention technique, and subsequently used a type of thematic coding grounded in this technique (Dumez, 2021).

Our research produced three key findings: (i) Coaching is not a practice with any kind of theoretical basis. Nevertheless, (ii) it is extremely standardised or regulated in that it is structured by a framework, ethics and supervision. Lastly, (iii) analysing coaching as a commercial relationship provides us with a better understanding of it. These three key findings have ultimately allowed us to form a firm picture of the practice.

<sup>(1)</sup> Negative theology asserts that God can only be described by what he is not, as opposed to by what he is.

<sup>(2)</sup> All interview excerpts have been translated from French.

## Research methodology

Our research is based on ten interviews conducted between March 2017 and September 2019 and three additional interviews conducted in September and November 2020. We sought to explore a wide variety of professional situations (e.g. freelance coaches, coaches employed by private coaching firms, internal corporate coaches and mentors) and coaching associations (e.g. SF Coach, the ICF, the European Mentoring and Coaching Council [EMCC], In'Coach and the Professional Supervisors Federation [PSF]). A study of ten interviews may sound insufficient, but we reached a saturation point – which is characteristic of qualitative research – after ten interviews, a phenomenon related to one of coaching's central features: Because it is so regulated or standardised, all of its actors describe it in much the same way. Although the profession's codes of ethics make coaching sessions difficult (and usually impossible) to observe, we attended two two-hour group mentoring sessions, with the permission of the mentors and mentees, as non-participating observers.

We analysed this interview material using the evenly suspended attention technique, followed by thematic coding (Ayache & Dumez, 2011; Dumez, 2021).

We then compared the findings of our analysis of the interview material to the real-world experience of a coach (who we interviewed twice) and a mentor in September and November 2020, based on an approach recommended by Piore (2006).

We decided to quote liberally from the interviews we conducted in this paper in order to give a clear picture of the practice.

## An atheoretical practice (due to an overabundance of theories)

In the space of just ten interviews, we were surprised by the diversity of theories used in relation to coaching, as well as the wide range of standing of these “theories”. It all starts with philosophy, and naturally Socrates. But Montaigne also frequently comes up, followed by eminent psychoanalysts: Freud, of course, as well as Jung and Erikson, and Balint and Lacan (Arnaud, 2003). We noted, for one, that coaches often mention the phenomena of transference and countertransference. Also referred to are systems analysis, transactional analysis, process communication, psycholinguistic analysis, *Gestalt* psychology and the Palo Alto School, in addition to Leonard Laskow (Laskow & Chertier, 2015) and clean language (David Grove, see Wilson, 2017), the work of W. Timothy Gallwey (2000), Enneagram, neuro-linguistic programming (NLP), nonviolent communication, singular mediation (or *médiation singulière* in French, a practice created by Dominique Lecocq, a professor and psychoanalyst who teaches at the *Conservatoire national des arts et métiers* [CNAM]) and transformance. Neuroscience is also cited, as is the work of Joseph Campbell on comparative mythology and that of Jean-Pascal Debailleul (2010) on the narrative structure of stories.

One of our interview subjects offered the following summary of the various approaches on which the coaching relationship is based:

“None of the key aspects of coaching came out of thin air. The practice draws heavily on Socrates and the Bible; for instance, “God helps those who help themselves”. Other ideas are taken from the Koran or Buddha. Yin and yang. Freud himself was influenced by Eastern philosophy. Jung's archetypes were drawn from Buddhism. We didn't make anything up. Relationship manuals have existed for thousands

of years; *The Knight in Rusty Armor*, *La Voie de l'amoureux* [by French author Arouna Lipschitz]. Some things come from Kabbalah, gnosis and Sufism. You see what works. If you take out the religious aspects, there are still worthwhile things. Also, if you look into the Palo Alto School, that's also completely focused on relationships.”

The practice of coaching thus pulls from a plethora of philosophical, psychological and esoteric sources (in addition to existential coaching, ontological coaching, etc., all approaches presented in Cox et al., 2018). Costa and Garmston (2016, p. 4) describe coaching rather colourfully as “[a] blend of the psychological orientations of cognitive theorists and the interpersonal bonding of humanists”. Some experts extol the virtues of managed eclecticism, an approach which is liberating in the sense that it does not limit coaches to a single approach (Clutterbuck, 2010).

The coaching relationship can be described as follows: A client has a job-related problem (for example, they struggle to delegate tasks) or is starting a new position (an engineer becomes the manager of a team of engineers in their own technical field – i.e. they are looking to continue their career in their field – or, instead, they are working outside their field and thus feel like an imposter). They work on their problem with a coach who is not there to tell them what they should do, but rather to help them find a solution to their problem. This involves working on the unconscious (if the client was perfectly aware of the origin of their problem, then they would be able to resolve it easily on their own), but it is not therapy; the work is focused on changing certain behaviours. As one of the coaches we interviewed put it, coaching is not therapy, advice or training.



It helps people change their professional conduct, but instead of the support coming from their superiors, as is generally the case in the working world, it comes from a third party, the coach:

“Initially, coaching was created to help executives who felt isolated. Who do they have to talk to? An executive can’t really talk to their employees. Some might be able to confide in a romantic partner, but that isn’t always the case. So who can they turn to? Coaching became a way for executives to deal with their isolation.”

Given that this distinctive practice – which is much easier to define by what it *is not* than by what it *is* – is not grounded in an established theoretical framework, on a most basic level we could expect coaching to be a nebulous concept, taking on a variety of different forms depending on whether its practitioners are influenced by Jung, Gallwey, Erikson or singular mediation. But this is not at all the case, as coaching is highly structured.

## How the coaching relationship is structured

The practice of coaching is structured by a framework, ethics and supervision.

### Framework

The practice of coaching is thoroughly structured: The client begins by choosing a coach, a contract is drawn up specifying the number of sessions, the sessions take place and the coaching engagement concludes with a wrap-up session.

The client must choose to be coached. Typically, the future coaching client meets with several potential coaches and selects the one who suits them best. If the client is being coached within a company they work for, the coaching manager must first make sure that the request for coaching is made by the client themselves, not one of their superiors. Then, a potential coach is put forward and the future coaching client is informed that after their initial meeting with the coach, they can be introduced to a different coach if they have any reservations. In any event, the coaching manager must be sure that the employee chose to be coached, i.e. that it was not forced on them by someone else, even if only indirectly or if they were pressured into it, and that the choice of coach was an informed decision, made on the basis of trust.

The contract is the first, and most foundational, part of the coaching framework:

“What’s the coaching relationship, you ask? It’s structured by a contract; it’s not a hierarchical relationship but a contractual one. From the outset, it’s established what the client needs out of the relationship and what they seek to accomplish through our sessions. I make the client define their expectations. A mutual commitment must be established around a contract and trust. So there’s a contract, a commitment and trust.”

“A lot comes down to the initial meeting. That’s when I show the client the contract and we go over it together. The contract really lays the foundation.”

To establish the contract, a tripartite meeting takes place:

“The coaching relationship exists because the client and the sponsoring company have objectives that they want to achieve with the help of a coach. It serves the intentions and objectives of the coach, who has their own set of intentions, wishes and aims, as well as the interests of the company and the client. A tripartite meeting forms the basis of the coaching contract. It allows us to establish the objectives we’ll be working towards, and the client is the focus of this work.”

The contract is drawn up and filed. The client identifies their issue and sets (themselves) the objective of resolving it by changing their professional conduct. It is made clear that all future face-to-face sessions between the client and their coach will be kept strictly confidential.

The second part of the coaching framework is the end of the coaching engagement: A coaching relationship must come to an end (Freud himself was absorbed by the question of whether or not a psychoanalysis comes to an end, 1985/1937). The coaching sequence is rather standardised, though it does leave some room for flexibility: The coach holds a first meeting with the future coaching client, which is then followed by a tripartite meeting with the coach, the client and the representative of the sponsoring company (e.g. a manager or HR manager), six to seven face-to-face sessions and a tripartite progress review to determine how the objective(s) outlined in the contract were achieved. There are usually a total of ten sessions, which typically last from an hour and a half to two hours each and are scheduled over three to six weeks:

“What works well for me is about ten face-to-face coaching sessions of two hours, plus two tripartite meetings (one at the beginning and another at the end). The client is the one who does the progress review at the end. I think it’s important to take your time. The client must be able to put things into perspective, so two hours isn’t excessive. With just one hour, you feel rushed. As for remote sessions, an hour to an hour and a half minimum, it’s more tiring, they’re more frequent (every two weeks).”

Any materials used, such as flip boards, are photographed before being destroyed. Since coaching is not about guiding the client, the coach intervenes as little as possible, only asking questions:

“The coaching is done by the client alone. That’s important. The more I do, the less successful it is. There’s the masculine energy, which is the active role. The feminine energy is more about making the person feel welcome, creating a sense of security. When the client gets tired, I take over. That’s when I play the masculine role and the client plays the feminine role. If we’re both playing the masculine role at the same time, we’re going to talk over each other and things will heat up. If the client is tired, I’ll give more of myself. If they have the energy for it, they lead the session.”

Clients are assigned homework and exercises to do in between sessions:

“I ask them to keep a journal, which helps them learn to hold a mirror up to themselves. They get into the habit of it and are always supported along the way. These little exercises are very practical.”

“Just a few exercises: When you meet with your co-workers, ask them, ‘What is it you need?’ And you let them speak. You listen. Create a situation where the person comes up against themselves by how they operate.”

These assignments help the coach monitor the client’s progress: If the client does not do the planned exercises in between sessions, the coach knows where things stand. Typically, sessions take place less frequently as the engagement draws to a close:

“There are two phases: The relationship-building phase, during which sessions take place more often, every two weeks, followed by the empowerment phase, which readies the client for the end of the coaching relationship, during which sessions become less frequent.”

A final session then concludes the coaching relationship:

“Ending the coaching relationship is a big deal. Well, maybe not a ‘big deal’... But it’s really important. We take stock of the coaching sessions. Before the last session, I send the client a progress report. During the last session, we discuss the report and assess the client’s progress. I’m not meant to be part of peoples’ lives forever, though I do love it when they give me updates. A healthy coach-client relationship means being there for a certain period of time and then exiting their lives. This assumes that the client now has all the resources they need.”

“Another aspect that makes coaching different from therapy is that we’re not treating a person’s troubles; we’re working on objectives that they set for themselves, for a limited period of time.”

The practice of coaching seems particularly structured – and perhaps even standardised in the way it is performed – given that it (i) is based on a written contract established by a coach and a client, along with a representative of the sponsoring company (who is often from the human resources department), (ii) plans a set number of sessions and homework in between sessions, and (iii) ends with a wrap-up session. Mentoring is also highly structured and sometimes adds another step. Midway through the mentoring process, the mentor brings in a sponsor who is neither a manager of the mentee nor the requesting entity within the mentee’s company, but a person at the organisation who knows the mentee and can talk about how they are perceived in the context of their job and what areas they need to improve. The sponsor takes part in the assessment phase, noting in what ways the mentee has improved over the course of the mentorship.

## Ethics

The coaching relationship is also structured by ethics. Coaching associations have drawn up specific, stringent codes of ethics. That means there is a common set of rules shared by all coaching associations and professionals, the first of which being that coaching sessions are confidential:

“What goes on over the course of a coaching engagement is only the business of the coach and the client. Coaches, much like doctors, must respect client confidentiality. Companies don’t always cooperate, but coaches must fully maintain confidentiality. Some HR reps try to get around this by asking you out to lunch, and try to extract information from you.”

The second rule states that coaches have a duty to protect their clients:

“Because they’re the client, and my duty is to protect them. That’s non-negotiable.”

The third rule is that coaches should never agree to an engagement that involves coaching a manager and someone from their team. In addition to having to follow these rules, coaches must at times navigate delicate situations that pose ethical dilemmas (Fatien Diochon & Nizet, 2015):

“A client told me, ‘I want to leave my job, I want to change professions.’ But for 10 or 15 minutes he talked non-stop about his wife and what she thought. I said to him, ‘Coaching isn’t what you need.’ Recognising when to say that... that’s what ethics means.”

“A client said, ‘My objective is for my employees to like me.’ I told them I couldn’t do that for them.”

“While drafting the contract, a client explained one of their objectives: ‘I want people to do what I ask them to do.’ I said no, that’s manipulative. He reconsidered and ultimately modified his objective. We ended up working together.”

The most delicate types of situation are when a company has reached out to a coach as a last resort or when it quickly becomes apparent that the coaching client’s manager is the one who could really use the coaching:

“Coaches are extremely wary of what we like to call ‘last chance’ coaching. That kind of engagement is awful. If it isn’t successful, the client is going to be fired. It’s not always made explicit, but you catch on. Engagements like that are really difficult, because that’s not what coaching is about. And then there are always ways to spoil a good thing, to use it for bad designs. There’s coaching that takes place under false pretences, and situations where it isn’t the client, but the client’s bosses who should be the ones receiving coaching.”

Managing the coach-client relationship, in which both parties often form a strong bond, also raises ethical problems when the partnership goes outside the bounds of a normal coaching relationship:

“There have been rare occasions where I thought I could’ve become friends with a client, with the relationship turning into more of a friendship. But because every coaching engagement has an end date, I don’t try to see my clients once our time is up. I’ve never become friends with my clients, though I’ve had people ask if we could become friends.”

“Becoming friends is a possibility, but that’s not the goal. You might invite a former client out for a drink to check in with them. It’s sort of like after-sales service. Some coaches maintain that becoming friends with a client isn’t allowed and that it’s unethical. If I was a therapist, I wouldn’t do it. But I don’t mind calling someone to get an update on how they’re doing.”

Internal corporate coaching raises a particular set of ethical problems:

“We’re not allowed to talk about our coaching sessions, though we can anonymously pass on things that are said. The company I work for fully complies with these rules. The HR department doesn’t want to know who I’m coaching, but they do want to be informed of early warning signs, such as burnout or lack of well-being. When I was just starting out and I heard about internal coaching, I wondered how you could be an employee *and* independent. You’re being paid by an organisation, there’s a duty of loyalty, so how do you comply with the code of ethics? External coaches

had the worst possible things to say about internal coaching. I discovered the profession and found it to be much more demanding than external coaching. It's so easy to get it wrong."

In addition to a framework, the practice of coaching is also structured by codes of ethics drawn up by professional associations. While these codes may vary somewhat from one association to another, they all share a common core of ideas.

### Supervision

Owing to the ethical problems, difficulties and predicaments that can arise in the coaching relationship, coaches must generally be supervised. This is the case for all coaches who have been certified by a professional association, as supervision is a requirement for certification (supervisors may also be certified by coaching associations or an association of coaching supervisors). It is more difficult to say how many uncertified coaches are supervised:

"What makes evaluating coaches so hard is that sometimes what they say is different from what they actually do. People sign the coaching code of ethics with the sincerest intentions. They don't necessarily realise when they don't follow it. For example, when coaches make judgments, they are acting in an advisory capacity, mixing their issue with that of the client's, which is why supervision is needed."

"Supervisors help us when we're struggling with or not making progress on an engagement, and they even give us insight into why an engagement went well. Supervisors are there to point out my blind spots."

Supervisors work either one-to-one with coaches or with a group of coaches. In the latter case, a supervisor meets with the coaches about once per month:

"The coach lays out their problem. Then they sit back. The other coaches talk about what they would have done in their position, while the first coach listens. All of them offer their perspective as coaches and explain how they would've handled the situation."

Internal corporate coaches are supervised by professionals outside their company.

In summary, coaching is a highly structured practice involving a framework, ethics and mandatory supervision. And yet it draws on a tremendous variety of theoretical approaches. Why is this so? We suggest that the answer lies in the commercial nature of the coaching relationship.

### Coaching as a commercial relationship

Although reliable figures are not available, it appears that companies initiate 90% to 95% of coaching engagements. But companies are faced with two uncertainties in the coaching relationship. The first is that of the quality of the service provided, as described by the lemons problem theory (Akerlof, 1970). Corporate coaching differs markedly from sports coaching on this point: The quality of the latter is evaluated in a direct, transparent way, based on the performance of the team being coached. The team either wins its games and the coach is praised, or it loses and the coach is fired. In a corporate setting, assessing the outcome of a coaching

engagement is more problematic because it is not as directly apparent, even when performance indicators have been established:

"One day, I was coaching someone to help them improve their delegation skills. This person had a great team and a senior position, but he wasn't delegating enough. I asked the HR manager, 'How will you know at the end of the engagement that Mr So-and-So has made progress?' To which she replied, laughing, 'We'll know that he has learned how to delegate after he uses up all 38 of his banked vacation days'."

Uncertainty over outcomes is heightened by coaching's most fundamental rule of ethics, i.e. that coaching sessions are confidential. A company must not and cannot have any knowledge of what goes on in the context of a coaching engagement. This means that a company can neither evaluate the quality of the service provided nor the manner in which it was provided, even though it must cover the relatively expensive cost of the service, as coaching is not covered by government-funded training programmes:

"Companies include coaching engagements in their training budget, but they must bear the cost since the government doesn't consider them to be a form of training."

And yet there is an overabundance of supply in the coaching market. Upon completing what may be rather ill-defined training in the practice, anyone with a certain amount of business experience can call themselves a coach and sell their services.

It would seem impossible for supply to meet demand in such an environment. The stakes of coaching can be high for companies, particularly where executives are concerned; if they are not able to judge the quality of a given coach, how can they buy their services with any confidence?

Three key elements make this commercial relationship possible: Training, certification (or accreditation) and a structured coaching relationship.

Coaching training programmes have grown rapidly. Many elite institutions of higher learning (including HEC) offer such programmes, along with a number of universities (Paris 8, for example, has a postgraduate degree, or DESU [*diplôme d'études supérieures universitaires*] in the discipline) and private educational institutions. But educational establishments cannot regulate the commercial relationship on their own: Company managers can complete a course of study in coaching in one to two years, but some institutions offer programmes that take just six months. Assessing the quality of such a profusion of programmes proves difficult:

"It sustains the illusion that if you know a little bit about transactional analysis, process communication, psycholinguistic analysis... get some training on how to apply tools and write a short thesis, and you're a coach."

Hence the need for certification. Given the wide-ranging intellectual origins of the practice, it seems virtually impossible to implement a uniform certification programme. The creation of a diverse range of certification mechanisms to ensure quality has filled the void. SF Coach, founded in 1996, was the first coaching

association ever created in France, followed by the French branch of the International Coaching Federation (ICF) in 1999 (the global federation dates back to 1995) and the European Mentoring and Coaching Council (EMCC) in 2002 (its predecessor organisation, the European Mentoring Centre, was founded in 1992). The ICF represents the American style of coaching, whereas SF Coach is more steeped in the psychoanalytic tradition. Coaches can thus join whichever organisation they identify most closely with. There are also a number of other, less influential professional associations:

“The goal is to put in place qualification processes to avoid an ‘anything goes’ situation with things like ‘self-coaching’ and all kinds of other nonsense. Not to mention cults. We’re here to prevent unsavoury practices.”

These professional associations have taken different routes: While the ICF has significantly expanded its membership, SF Coach has a very strict, selective membership policy and considerably limits its number of members. Each association has its own code of ethics, although they share the same body of basic rules. Likewise, the EMCC and the ICF have established competence frameworks; once again, they differ but are built on similar core ideas:

“The EMCC has eight competence categories in its framework and the ICF has 11. If you look at them closely, you see that they’re the same, just framed differently.”

In the end, the profession has become structured:

“Coaching has become very well structured. It’s been quite a success.”

When companies work with coaches who have been certified or accredited by an association, they are reassured that they are buying a high-quality service. But above all else, what establishes the commercial nature of the coaching relationship is the fact that it is structured. Companies need to know what they are buying before they purchase a coach’s services:

“It’s a tripartite relationship: the company that is paying for the service, the coach and the client. For everything to go smoothly, you have to establish a detailed contract when you get a coaching request, since it’s the company that pays, so that they’ll be reassured, as under confidentiality rules the company won’t know anything more. They’re buying a service.”

As previously established, coaching has, in this way, become a relatively standardised practice. The nature of the service being sold, if not the content of the service itself, needed to be clearly defined:

“As an external coach, you’re selling six coaching sessions, plus a meeting with the client and two tripartite meetings.”

Interestingly, the coaching contract fulfils two obligations: First, the relationship is a commercial one, and the company must know what they are buying (it is inconceivable for a company to buy a service, for instance, which has no set end date, such as psychotherapy). Second, coaches, who draw on a vast range of intellectual foundations, must be able to carry out their engagements as they see fit, relying on whatever approaches they choose. This is allowed under the contract: The company knows that they are buying nine to ten coaching sessions and that they will be able to attend two of them (one during which their

voice will be heard and another during which the coach will take stock of their experience with the client). With the engagement structured in this way, the coach then has six to eight sessions that they can conduct as they feel appropriate, in complete confidentiality. These two aspects – the non-hierarchical support that coaching provides and the commercial nature of the relationship – have become solidified in the practice’s characteristic *modus operandi*.

## Future outlook and conclusion

In this exploration of the practice of coaching we have demonstrated that it should be considered from two angles: as a form of support in the workplace, uncommon in that it is non-hierarchical, and as a commercial relationship, wherein a company is buying a service. The intellectual foundations of coaching are surprisingly wide-ranging, as many authors have remarked, but the way the coaching relationship has been structured gives coaches the ability to use their own approach, in line with the sponsoring company’s requirements, and offers potential clients an array of coaching methods to choose from.

The coaching market has entered a phase of major growth which is likely to disrupt the very nature of the coaching relationship and coaching services. On the demand side, companies began by exposing their senior executives to coaching. After they were satisfied with the experience, coaching spread from the top down, particularly due to the flattening of management structures, which has put greater pressure on middle management (Littler et al., 2003; Hales, 2006):

“Back in 2002 and 2003, I remember giving talks on coaching as ‘punishment or reward’. Coaching was this secret thing and you weren’t supposed to tell anyone that you were being coached. If you were seeing a coach, that meant there was a problem with you or something. Coaching was practically a shameful thing at first, but that’s no longer the case today. Stuff like that might still go on, but it isn’t the norm. Coaching has become one of a number of support structures. In the early days, it was for senior executives and managers. Nowadays, employees are being coached too. The cost varies depending on the client’s position in the company.”

Other types of coaching followed, such as team coaching, project coaching and organisational coaching. The practice of mentoring took the same trajectory, with one-to-one mentoring leading to the emergence of group mentoring (in which several managers from either the same firm or different firms are mentored). One of the companies we reached out to shared that their coaching practice breaks down into one-third one-to-one coaching, one-third team coaching and one-third project coaching. In fact, this expansion occurred rather organically:

“I started coaching teams very early on because I noticed that one-to-one coaching had some drawbacks in certain situations. It put too much emphasis on changing a particular person when in reality it was either the team that needed to behave differently or the manager and their team that needed to change. I felt that one-to-one coaching had limits and that team coaching could sometimes be more relevant.”

Supply has grown to accommodate the rise in demand for coaching. Late-career managers find new meaning in helping others, especially their younger colleagues. And many coaching clients end up wanting to become coaches themselves. Higher education has kept up with the trend, creating a myriad of training programmes that teach the fundamentals of coaching to those wanting to join the profession.

To meet the needs of the rapidly growing coaching market, the industry had to institutionalise, ultimately becoming stratified. At the upper tier of the market, demand for coaching services comes from large corporations for their executives. Demand is met by a supply of highly regarded coaches, working either as freelancers or at coaching firms, who are certified by major coaching associations. The next tier of the market concerns middle management at large corporations. Recent years have seen the emergence of tenders directed exclusively at coaching firms which offer to introduce potential clients to two or three possible coaches so that the client is free to select among several options, for some 20 coaching sessions per year over two or three years, with six months of coaching costing roughly €15,000 excluding VAT.<sup>(3)</sup> Mentoring rates are in a similar ballpark.

In the midst of the COVID-19 crisis, a mentoring firm told us that it had set a minimum rate of €12,000 for mentoring engagements of six months. Large corporations often use a mix of external and internal coaches, with internal ones being mid-level employees who devote a portion of their time to coaching (20%). As such employees' working hours are difficult to gauge and frequently open-ended, the cost for corporations is negligible, with training representing the largest burden.

One of the lower tiers of the coaching market caters to small- and medium-sized enterprises (SMEs), among others, and offers a wider range of services which can be provided by less reputable coaches who are not always certified. The bottom-most tier is increasingly occupied by low-cost services, such as three-session coaching packages and online coaching platforms. On the whole, only a small number of coaches make a living from it. Most work as coaches part-time alongside another job, as instructor for example.

Two shifts underway have the potential to upend the market as it currently stands, as well as the practice itself. The first regards the rapid growth the market has seen, as mentioned in our analysis. It is reasonable to wonder whether coaching will become a victim of its own success. As we pointed out, mainstream companies have democratised coaching, making it available to executives, middle managers, teams, projects and entire organisations. Today we are witnessing the emergence of agile coaches, who, more often than not, are coaches in name only. Furthermore, a new form of organisation called the liberated company (Gilbert et al., 2017), or holacracy (Battistelli, 2019), does away with hierarchy altogether:

<sup>(3)</sup> This cost is based on information obtained during interviews. Fatien Diochon and Nizet (2012, p. 28), however, indicate lower costs ranging between €5,000 and €12,000.

“The biggest trend nowadays is that everyone wants to become a coaching manager, as though the organisational hierarchy has been flattened.”

But this would barely classify as coaching and is essentially a contradiction in terms, in that coaching is by definition a non-hierarchical relationship. How, then, could such a relationship exist alongside a managerial relationship? How is it possible to introduce something as foreign as a non-hierarchical relationship into the management culture of the corporate world? Are we bearing witness to the “coachification” of companies, at the same time as a form of “coachification” of society? Like every other management fad, coaching may well be destined to decline in a few years after becoming a victim of its own success (Midler, 1986; Abrahamson & Fairchild, 1999).

One of our interviews raised a second shift – one both dreaded and eagerly anticipated – that of artificial intelligence (AI), at a time when phone and video coaching have become more widespread, largely owing to the COVID-19 crisis, and perhaps one day holograms, making it possible for coaches and their clients to meet remotely in the same “room”:

“Ten years from now, it will be the turn of AI. It's going to change things. We're already seeing more and more coaching via Skype and over the phone. AI will take over, for good or ill.”

We did not ask any questions on the topic of AI in our interviews, as it was not part of our initial orienting theory, and the above response produced an isolated occurrence, or hapax, during the thematic coding process. We then ran a search on Google Scholar using the keywords “coaching” and “artificial intelligence”, and some results came up, including a recent seminal paper published by Clutterbuck (2020). His essay reviews the current state of AI research at the University of Southern California, making reference to the development of AI therapists, a technology that enables real-time analysis of micro-expressions, physical signs of stress and responses expressed by patients during therapeutic conversations. In addition, the technology keeps record of previous sessions and is being tested for use in coaching. The AI's ability to provide “real-time information about what is going on in the conversation”, suggest questions, check coaches' intuitions and help them review the way they run coaching sessions could have a profound impact on the practice of coaching.

## Bibliography

- ABRAHAMSON E. & FAIRCHILD G. (1999), “Management Fashion: Lifecycles, Triggers, and Collective Learning Processes”, *Administrative Science Quarterly*, 44(4), pp. 708-740.
- AKERLOF G. A. (1970), “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism”, *The Quarterly Journal of Economics*, 84(3), pp. 488-500.
- ARNAUD G. (2003), “A Coach or a Couch? A Lacanian Perspective on Executive Coaching and Consulting”, *Human Relations*, 56(9), pp. 1131-1154.
- AYACHE M. & DUMEZ H. (2011), “Le codage dans la recherche qualitative : une nouvelle perspective ?”, *Le Libellio d'Aegis*, 7(2), été, pp. 33-46.

- BATTISTELLI M. (2019), "Les apports contrastés de l'holocratie à la démocratie délibérative en entreprise. Une étude ethnographique dans une PME de l'Yonne", *RIMHE*, 2019/2, n°35, pp. 3-23.
- CLUTTERBUCK D. (2010), "Coaching Reflection: The Liberated Coach", *Coaching: An International Journal of Theory, Research and Practice*, 3(1), pp. 73-81.
- CLUTTERBUCK D. (2020), "The Challenges of Coaching and Mentoring in a Digitally Connected World" in WEGENER R., ACKERMANN S., AMSTUTZ J., DEPLAZES S., KÜNZLI H. & RYTER A. (éd.), *Coaching im Digitalen Wandel*, Göttingen, Vandenhoeck & Ruprecht (Open-Access-Publikation), pp. 19-29.
- COSTA A. L. & GARMSTON R. J. (2016), *Cognitive Coaching: Developing Self-directed Leaders and Learners*, Lanham (Maryland), Rowman & Littlefield Publishers, 3<sup>rd</sup> ed.
- COX E., BACHKIROVA T. & CLUTTERBUCK D. (2018), *The Complete Handbook of Coaching*, Thousand Oaks, Sage, 3<sup>rd</sup> ed..
- DEBAILLEUL J.-P. (2010), *Manuel de thérapie par les contes de fées : Du symptôme à l'accomplissement de soi, Quinze parcours de rêve éveillé à partir de contes merveilleux*, Gap, Le souffle d'or.
- DUMEZ H. (2021), *Méthodologie de la recherche qualitative*, Paris, Vuibert.
- FATIEN P. (2008), "Des ambiguïtés des mots/maux du coaching", *Nouvelle Revue de Psychosociologie*, 2(6), pp. 193-211.
- FATIEN P. & NIZET J. (2011), "Le coaching, une pratique critique ?", in TASKIN L. & NANTEUIL M. de (éd.), *Perspectives critiques en management*, Bruxelles, de Boeck.
- FATIEN DIOCHON P. & NIZET J. (2012), *Le coaching dans les organisations*, Paris, La Découverte, Repères.
- FATIEN DIOCHON P. & NIZET J. (2015), "Ethical Codes and Executive Coaches: One Size Does Not Fit All", *The Journal of Applied Behavioral Science*, 51(2), pp. 277-301.
- FREUD S. (1985/1937), "L'analyse avec fin et l'analyse sans fin", in FREUD S., *Résultats, Idées, Problèmes*, Paris, PUF, tome II, pp. 231-268.
- GALLWEY W. T. (2000), *The Inner Game of Work: Focus, Learning, Pleasure, and Mobility in the Workplace*, New York, Random House.
- GHERARDI S. (2019), *How to Conduct a Practice-Based Study? Problems and Methods*, Cheltenham, Edward Elgar, 2<sup>nd</sup> ed.
- GILBERT P., TEGLBORG A.-C. & RAULET-CROSET N. (2017), "L'entreprise libérée, innovation radicale ou simple avatar du management participatif ?", *Gérer et Comprendre*, n°127, mars, pp. 38-49.
- GUILHAUME G. (2009), *L'ère du coaching. Critique d'une violence euphémisée*, Paris, Syllepse.
- HACKMAN R. J. & WAGEMAN R. (2005), "A Theory of Team Coaching", *Academy of Management Review*, 30(2), pp. 269-287.
- HALES C. (2006), "Moving Down the Line? The Shifting Boundary Between Middle and First-line Management", *Journal of General Management*, 32(2), pp. 31-55.
- LASKOW L. & CHERTIER G. (2015), *Guérir par l'Amour : Favoriser les processus naturels de guérison*, Archamps, Jouvence.
- LITTLER C. R., WIESNER R. & DUNFORD R. (2003), "The Dynamics of Delaying: Changing Management Structures in Three Countries", *The Journal of Management Studies*, 40(2), pp. 225-256.
- MIDLER C. (1986), "Logique de la mode managériale", *Gérer et Comprendre*, n°3, juin, pp. 74-85.
- NIZET J. & FATIEN DIOCHON P. (2012), "Comprendre les ambiguïtés du coaching, à l'éclairage du fonctionnalisme", *Gérer et Comprendre*, n°110, décembre, pp. 24-33.
- PASSMORE J. & FILLERY-TRAVIS A. (2011), "A Critical Review of Executive Coaching Research: A Decade of Progress and What's to Come", *Coaching: An International Journal of Theory, Research and Practice*, 4(2), pp. 70-88.
- PIORE M. J. (2006), "Qualitative Research: Does It Fit in Economics?", *European Management Review*, 3(1), pp. 17-23.
- SCHATZKI T. R. (2008), *Social Practices. A Wittgensteinian Approach to Human Activity and the Social*, Cambridge, Cambridge University Press.
- THEEBOOM T., BEERSMA B. & VAN VIANEN A. (2014), "Does Coaching Work? A Meta-analysis on the Effects of Coaching on Individual Level Outcomes in an Organizational Context", *The Journal of Positive Psychology*, 9(1), pp. 1-18.
- VANHEULE S. & ARNAUD G. (2016), "Working with Symbolic Transference: A Lacanian Perspective on Executive Coaching", *The Journal of Applied Behavioral Science*, 52(3), pp. 296-319.
- WILSON C. (2017), *The Work and Life of David Grove: Clean Language and Emergent Knowledge*, Kibworth Beauchamp, Matador.

# Artificial intelligence: between science and the market

## Some socio-historical elements to better understand a strange scientific experiment (1956-1990)

By Jean-Sébastien Vayre

Senior lecturer at the University of Côte d'Azur

Original article in French published in *Gérer & comprendre*,  
September 2021, n° 145, pp. 55-69.

Artificial intelligence (AI) is all the rage nowadays, with it being enthusiastically promoted by leading political and economic stakeholders involved in the development of digital technology. However, it is also a source of controversy, with some even claiming that it does not exist. It is a truly muddled affair. Is artificial intelligence but a mere illusion? This paper will explain why this is not the case. We will more precisely detail how this confusion surrounding artificial intelligence - very much a reality - has come about. To do this, we will posit that artificial intelligence is a scientific discipline that from its very origins was intertwined with an economic practice, resulting in an imbalance between basic and applied research. We will also build on this by concluding that it is precisely this imbalance that causes a lack of clarity surrounding the scientific discipline and, more generally, the instability of its development.

### Foreword

There is much debate today among economic and political stakeholders concerning the notion of artificial intelligence: should we not instead talk about augmented intelligence, biological intelligence or remote intelligence? Some experts go even further by arguing that it would be best to no longer talk about artificial intelligence. From a scientific standpoint, this kind of discussion is interesting because it has been around since the emergence of AI. For example, Herbert A. Simon and Allen Newell were not particularly taken by the expression “artificial intelligence” coined by John McCarthy and preferred “complex information processing system” (Newell & Simon, 1956; Norberg, 2019 [1989]). However, while economic and political stakeholders engage in heated debate over the concept of artificial intelligence, they do not really argue on scientific grounds: the issues at stake are a cause of debate because they have serious consequences from a business standpoint. For example, it is not hard to see why Luc Julia (2019), head of the Samsung research centre, is pushing to replace the term “artificial intelligence” with “smart object”: his professional interests are evident as Samsung’s innovation strategy completely revolves around the Internet of Things

(IoT).<sup>(1)</sup> As paradoxical as it may seem, claiming that “artificial intelligence doesn’t exist” (Julia, 2019) therefore does not reflect, at least not in this case, a simple intention to demystify artificial intelligence: it is above all also a business move. This is why, in order to untangle the web that is AI, we wish to start from the very beginning and ask the following questions. What is AI? Is it a science? Is it a consumer item? Is it a new form of autonomous intelligence that might surpass human intelligence? We believe that trying to solve the problem of defining artificial intelligence is important, in that the virtuous quality of AI development depends on how society understands it, the meaning that we collectively give to it.

To answer these questions, we have conducted a socio-historical investigation using papers, reports, communications and videos produced by two major types of actors: those who have a long-standing interest in AI, and those who directly helped conceive and sustain this scientific discipline. This body of documentation was compiled as part of a doctoral project (Vayre, 2016), and was supplemented with research that we conducted over the last four years on the history of

<sup>(1)</sup> <https://www.strategies.fr/actualites/marques/4027180W/-l-intelligence-artificielle-n-existe-pas-luc-julia.html> (in French)

AI. Tables 1 and 2 list the resources used in our thesis which form the basis of the work presented in this paper. As mentioned previously, the work was produced with the aid of additional studies, and is therefore based on several sources not cited in Tables 1 and 2. However, these documents are systematically referenced in the body of text and in the paper's bibliography. In addition, the resources used to conduct our investigation, as a whole, were compiled using a methodology that we could deem abductive (Bruscaglioni, 2016), in that we have searched for and explored the substantive documentation to confirm or reject hypotheses made as our work progressed. In other words, and contrary to, for example, the work of Dominique Cardon and his colleagues (2018) which forms part of the development of what we may call, in reference to the French *Annales* school (Burguière, 1979), a quantitative history of AI,<sup>(2)</sup> our working approach instead follows on from the evidential paradigm proposed by Carlo Ginzburg (1980). The author of this study considers quantitative history, while having the merit of shedding light on the major structures that drive the dynamics of a given phenomenon over the long term, as tending to classify these dynamics under categories of thought that are at times far too general. Ginzburg (1980) therefore

posits that the negative impacts of this tendency may be mitigated by adopting this mindset which lies, he believes, at the root of intellectual history, and consists of reconstructing an invisible reality by interpreting traces of the past that are perceptible in the present. In the words of Denis Thouard (2007), this way of "inferring from the facts" is, at least in Ginzburg's view (1980), a paradigm for research and thinking that is particularly useful in humanities and social sciences. We have therefore tried to adopt this model across all stages of searching for, compiling, reading and analysing documents comprising our study material. In short, the investigation findings detailed in this paper are the result of selecting documents and information that reflect fragments of empirical reality that we have gradually reassembled through knowledge and intelligibility effects, characteristic of "sociological reasoning" (Passeron, 1991).

At this point, we would like to specify that although we occasionally refer below to developments in AI over the last 20 years, we are primarily interested in the period from the mid-1950s to the early 1990s, as it was during this time that AI experienced its first waves of success and failure (Cardon, Cointet & Mazières, 2018). We will set out our findings below in two large sections, enabling us to distinguish between the scientific discipline and the economic practice that is AI. However, taking into account the work of Bruno Latour (1987), we are aware that this distinction has an abstract quality: from

<sup>(2)</sup> While the authors did conduct interviews in view of tracing the history of AI, they mainly relied on the statistical analysis of a corpus of over 27,000 articles compiled in 2018 on Web of Science.

Approach type	References
Symbolic artificial intelligence	(Bickhard & Terveen, 1995), (Bonissone & Johnson, 1984), (Fodor, 1975), (Fodor, 1983), (Fodor & Pylyshyn, 1988), (Forgy, 1981), (Gaschnig, 1980), (Ince, 1992), (Laird, Newell & Rosenbloom, 1987), (Lenat, 1977), (Lenat, 1983), (Lindsay, Buchanan, Feigenbaum & Lederberg, 1993), (McCarty, 1977), (Memmi, 1990), (Minsky & Papert, 1969), (Neumann, 1958), (Newell, 1980), (Newell & Simon, 1972), (Papert, 1988), (Samuel, 1959), (Simon, 1991 [1969]), (Simon, 1992), (Tristan & Abdallah, 2009), (Turing, 1950), (Winston, 1970)
Connectionist artificial intelligence	(Ackley, Hinton & Sejnowski, 1985c), (Bechtel & Abrahamsen, 1993), (Bickhard & Terveen, 1995), (Changeux, 1983), (Fodor & Pylyshyn, 1988), (Hayek, 1952), (Hebb, 1949), (Hopfield, 1982), (Lai, 2015), (LeCun & Bengio, 1995), (LeCun et al. 1989), (McCulloch & Pitts, 1943), (Memmi, 1990), (Minsky & Papert, 1969), (Noduls, 2015); (Numenta, 2011), (Numenta, 2014), (Numenta, 2015), (Papert, 1988), (Rosenblatt, 1958), (Rumelhart, Hinton & Williams, 1986), (Smith, 1999), (Smolensky, 1988), (Widrow & Hoff, 1960)
Hybrid artificial intelligence	(Bonasso, Firby, Gat, Kortenkamp, Miller & Slack, 1997), (Cassimatis, 2005), (DePristo & Zubek, 2001), (Hawes et al. 2007), (Kubera, Mathieu & Picault, 2011), (Langley & Choi, 2006), (Müller & Pischel, 1993), (Reynaud, 2014), (Schmidt, 2005), (Silver et al. 2016), (Smolensky, 1987), (Smolensky, Legendre & Miyata, 1992)

Table 1: Documents produced by researchers or organisations with direct involvement in AI development

Resource type	References
Written documents	(Blanc, Charron & Freyssenet, 1989), (Boise, 2007), (Copeland & Proudfoot, 2015), (Dupuy, 1994), (Hodges, 2014 [1983]), (Pélissier & Tête, 1995), (Varela, 1988)
Video records	(Dammbeck, 2003), (Folgoas, 1976), (Guirardoni, 1981), (Karlín, 1971), (Lallier, 1963), (Moreuil, 1972), (Royer, 1961a), (Royer, 1961b)

Table 2: Documents produced by researchers and journalists who have a long-standing interest in AI



a purely empirical standpoint, the research work that has conceived this discipline has inextricably linked the scientific and economic logics shaping the history of AI. However, we have decided to structure this article around this distinction precisely in order to untangle these two types of logic and thereby make the forms of their entanglement and the related issues easier to comprehend. We posit in the first section of this paper that AI is first and foremost a highly competitive scientific discipline, and broadly speaking seeks to conduct highly experimental research programmes. We shall then argue that AI research was, at least initially, a resounding failure from an applied research standpoint, even if it can be deemed a success from a basic research standpoint. Following this, in the second section, we will show how AI is also an economic practice about which many promises are made. We will then explain how this practice plays a vital social role in understanding and explaining the first waves of success and failure of AI. To conclude, we will outline some considerations to give a better understanding of how AI as a scientific discipline and as an economic practice interlink, stressing that, from a socio-historical perspective, this understanding provides insight into the current success of this discipline.

## Artificial intelligence as a scientific discipline

The conference organised by John McCarthy, Marvin L. Minsky, Nathaniel Rochester and Claude E. Shannon at Dartmouth College in 1956 laid the institutional groundwork for artificial intelligence. In their proposal drawn up in preparation for this event, the four authors define this science as follows:

“[Artificial intelligence] is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer” (McCarthy et al., 2019 [1955]).

However, it must be noted that the origins of AI are more generally embedded in the history of computing and cybernetics. For example, the work of Blaise Pascal, Gottfried W. Leibniz, Charles Babbage, Augusta Ada King, George Boole, Friedrich L. G. Frege, Kurt Gödel, and, of course, of Alan M. Turing, John von Neumann, Norbert Wiener, Warren McCulloch and Walter H. Pitts played a vital role in the emergence of this science (Crevier, 1997 [1993]; Pratt, 1995 [1987]; Rose, 1986 [1984]). It is also worth noting that, since the dawn of this science, the use of the term “artificial intelligence” has not been embraced by all. McCarthy was particularly taken by this term, who eventually persuaded his colleagues to adopt it. As previously mentioned, Simon and Newell preferred to talk about a “complex information processing system” (Newell & Simon, 1956; Norberg, 2019 [1989]).

## Different styles of research

Much like other sciences, AI does not have a perfectly harmonious community: not all stakeholders collectively share the same perceptions of this science. In the words of Pierre Bourdieu (1976), AI is a scientific field the stability of which is dependent on the power struggle driving it, within which various forms of domination emerge and dissipate. This is especially true since AI is highly interdisciplinary in nature: depending on their interests, researchers in this field may stumble into such different areas as biology, psychology, anthropology, logic, philosophy, linguistics, mathematics, electronics and computing. However, the study of AI revolves around one shared goal: each and every researcher in the field has helped to test the hypothesis that a machine can exhibit behaviour that humans would generally deem intelligent. Since the discipline’s beginnings, the methods of conducting this experimentation has been the subject of intense debate.

At the Dartmouth Summer Research Project the most prominent researchers in this community were Simon and Newell, and there are many reasons why this was the case. Firstly, in 1956, Simon and Newell were the only ones to have a computer program capable of synthesising one of the aspects of intelligence that academics often consider to be the most respected: solving complex mathematical problems. The Logic Theory (LT) program (Newell & Simon, 1956) is capable of proving half of the Principia Mathematica theorems of Alfred N. Whitehead and Bertrand A. W. Russell. Secondly, the LT program was designed using expertise in the fields of humanities and social sciences because the machine incorporates some of the fundamental concepts of the bounded rationality theory (Simon, 1945).<sup>(3)</sup> However, most researchers attending the conference believed that there was no point in studying human cognitive processes to design an AI program, such as McCarthy and Marvin L. Minsky, who, in the late 1950s, shared the view that AI must focus on exploring formal logic. This idea however is just as controversial as Simon and Newell’s theory. For example, Herbert Gelernter and Nathaniel Rochester (1958), along with Oliver G. Selfridge (1959) understood AI from different perspectives. In their view, AI should not be formed by using human cognition or formal logic as a reference, but rather by using just the information processing capabilities of machines as a basis. This approach enabled them to develop their first AI programs: for Herbert Gelernter and Nathaniel Rochester, this was the Geometry Theorem Prover (GTP, Gelernter & Rochester, 1958), and for Oliver G. Selfridge, the famed pandemonium model (Selfridge, 1959).

## Three major tension points

From the outset AI has been characterised by tensions. Over time, these tensions gradually intensified and eventually gave the field of AI a lasting structure. Between the 1960s and 1990s, there were at least three major tension points that played a decisive role in shaping the dynamics of this science.

<sup>(3)</sup> It was thanks to this theory that Simon was awarded the Nobel Prize for Economics in 1978.

The first point relates to amicable disagreements that quickly arose between McCarthy and Minsky concerning the way the issue of AI can be defined: while for McCarthy the fundamental issue underlying this new discipline was primarily one of logic, Minsky did not agree with this view (Norberg, 2019 [1989]). In 1960, this first point of tension emerged between the two researchers who, from that point onwards, undertook different research trajectories. As a result, in 1962 McCarthy decided to leave the MIT AI Lab to head his own one at Stanford University, the Stanford Artificial Intelligence Lab (SAIL). It was at this point that his work on logic had a significant impact on the AI expert community. For example, thanks to the list processing (LISP) language that McCarthy developed in 1958, Douglas Lenat was able to develop his Automated Mathematician (AM; Lenat, 1977) and EURISKO (Lenat, 1983) programs. Similarly, the "IF, THEN" advice taker program proposed by McCarthy in 1959 played a key role in the development of expert systems, just like his work a few years later on circumscription,<sup>(4)</sup> streamlining the information processing performed by these systems (Crevier, 1997 [1993]). With McCarthy's departure, Minsky was heading the MIT AI Lab on his own, a rather comfortable arrangement since he was receiving sizeable investments to outdo his new rival: over several years, the Defense Advanced Research Projects Agency (DARPA) gave \$3m in funding for the Machine-Aided Cognition and Multiple Access Computer project to MIT (Flamm, 1987; Norberg, 2019 [1989]). With this funding, McCarthy's former colleague had considerable resources to establish an anti-logic approach to AI. Given the economic, technical and human resources available to the MIT AI Lab, this approach quickly became considerably popular among Minsky's peers. Many young and brilliant researchers as a result flocked to work with Minsky, including

<sup>(4)</sup> This name refers to a computing process by which the obstacles potentially impeding the logic inference engines can be isolated or minimised, to allow for navigation within a knowledge-based system.

James R. Slagle, Joel Moses, Patrick Winston and Seymour A. Papert, who respectively developed the symbolic automatic integrator (SAINT; Slagle, 1961), the symbolic integration program (SIN; Moses, 1967), the arch concept learning program (Winston, 1970) and the LOGO programming language (Papert, 1971).

In short, as shown in Figure 1 (see above), from the McCarthy/Minsky split was born two major working approaches to AI. According to Roger C. Schank, up until the early 1990s there were two different research styles in the AI field: the "neat" style that subscribes to the logical approach developed by McCarthy, and the "scruffy" style more associated with the anti-logic approach developed by Minsky:

"In Schank's view, the neat style is refined, focusing on superficial phenomena like logic and syntax, which can be understood and compartmentalised in pretty little boxes. The scruffy style is haphazard, and revelled in wrestling with tortuous issues such as semantics" (Crevier, 1997 [1993], p. 201).

This first point of tension could be supplemented with another that helps to define the research stream developed by Schank and the stream of his colleagues Simon and Newell. In the 1960s, Simon and Newell were teaching and researching at Carnegie Mellon University, and were highly esteemed among academics and industrialists alike (Norberg, 2019 [1989]). They both therefore quickly gained recognition for their work on problem solving. On the back of the LT program's wow factor, Simon and Newell continued to study and draw on human cognition to develop new computing programs. Their work led them to developing their famed General Problem Solver (GPS; Newell, Shaw & Simon, 1959) which planted the seed for the design of the most well-known expert systems. For example, Edward A. Feigenbaum and Bruce G. Buchanan directly based the development of the DENDRAL (Buchanan & Feigenbaum, 1978) and MYCIN (Buchanan & Shortliffe, 1984) expert systems on the GPS, as did their students Randall Davis and John P. McDermott, who respectively designed the TEIRESIAS (Davis, 1978) and "eXpert

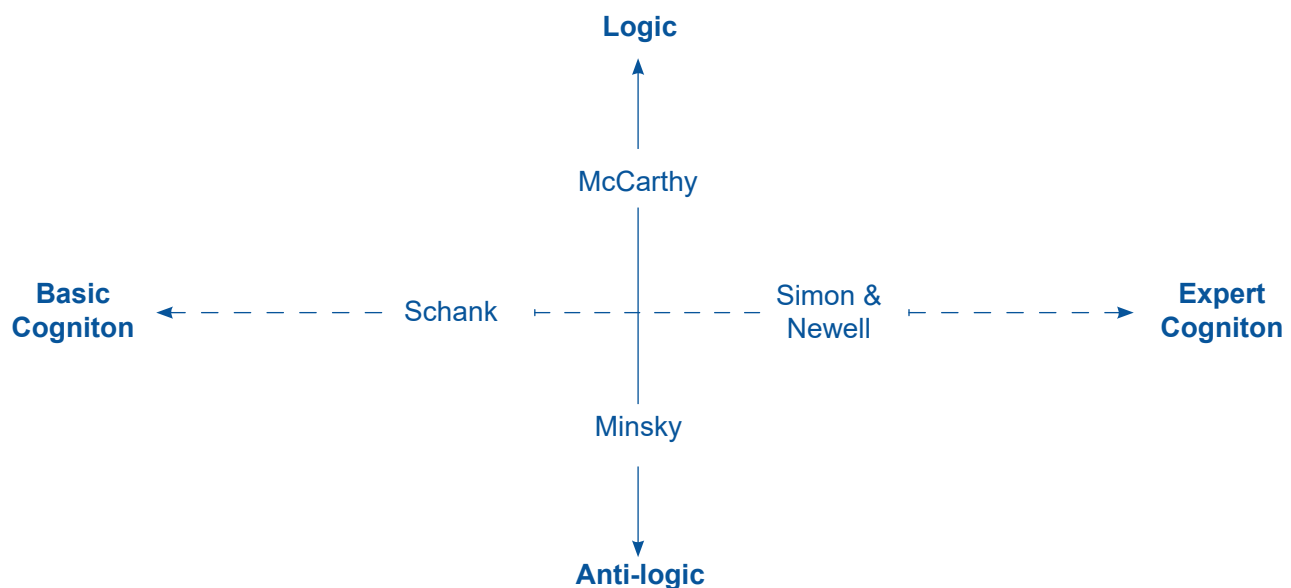


Figure 1. The four main research streams in symbolic AI

CONfigurer” (XCON; McDermott, 1982) programs. However, Schank’s (1978) work at Yale University did not relate to solving expert problems. While Schank and Simon and Newell share the view that AI should be based on the study of human cognition, Schank was less interested in so-called high-level intelligence and more so in intelligence applied in daily life. In other words, Schank did not study expert cognition, but was rather focused on basic cognition. His aim was not to develop computer programs to enable machines to assist or supplant highly qualified individuals, but rather to understand how these machines can adapt to humans from day to day to help them live together in a better society.

It should be noted that, as the work of Cardon and his colleagues (2018) demonstrate, there is a third tension point that played a fundamental role in the history of AI. This third point is the opposition between the symbolic and connectionist stances on this discipline. Indeed, the two points of tension discussed just above subscribe to the symbolic stance on AI: the aforementioned researchers all generally hold the view that intelligence is a computational system of symbolic representations of a rather deliberative nature (Fodor & Pylyshyn, 1988). This is exactly why, at this point and following Simon and Newell, many researchers had a particular interest in the expert aspect. However, while this approach dominated the field of AI from the 1960s to the 1990s, a shift occurred starting from the 2000s (Cardon, Cointet & Mazières, 2018; Vayre, 2016). The connectionist stance on AI that Frank Rosenblatt (1958), Wilfrid K. Taylor (1956), Bernard Widrow and Marcian E. Hoff (1960) struggled to defend against the criticism of Minsky and Papert (1969),<sup>(5)</sup> is now predominant, with the success of deep learning and the neurological theories explaining individual and collective behaviour (Changeux, 1983) being a reflection of this. From this viewpoint, intelligence is formed of a gargantuan network of triggers and inhibitors for basic calculation units, which is rather adaptive in nature. The collective focus of AI researchers is less on expert reasoning and more on what we may call, rather crudely, the “child’s learning experience” (Vayre, 2016). This other stance on AI, based on the biological nature of cognition and the mathematical nature of information processing, covers various research styles. So as not to spread the scope of this paper thin, we have decided, at least in this work, to not further touch upon this stance on AI.

### Successes and failures: the case of the BAIR Lab

From the outset there have been many successes in AI: the programs developed by researchers in the field are able to perform the tasks they have been assigned to do. For example, the SHRDLU program developed by Terry Winograd (1972), is able to question its motivations and therefore has something which somewhat resembles “self-awareness”; these two properties allow

<sup>(5)</sup> However, note that in spite of their criticism, the two authors have a certain interest in connectionism. It is worth mentioning that Minsky (1954) wrote his thesis on neural networks.

the program to engage in a truly constructive conversation with the user, and this was one of the first instances of this behaviour in the field of AI. For illustrative purposes, here is the sort of discussion that a user can have with SHRDLU:<sup>(6)</sup>

“23. Had you touched any pyramid before you put the green one on the little cube? YES, THE GREEN ONE [...]. 24. When did you pick it up? WHILE I WAS STACKING UP THE RED CUBE, A LARGE RED BLOCK AND A LARGE GREEN CUBE [...]. 25. Why? [...] TO GET RID OF IT [...]. 26. Why did you do that? TO CLEAN OFF THE RED CUBE [...]. 29. Why did you do that? BECAUSE YOU ASKED ME TO. (Winograd, 1972, p. 13).

As impressive as that may seem at a basic level, the problem is that this conversation can only truly be considered constructive within the bounds of the micro-world in which SHRDLU develops i.e., a virtual environment formed of several cubes, cones and spheres, a “toy world” you might say (Cardon, Cointet & Mazières, 2018). As explained previously, AI is an experimental discipline. In this respect, any successes, no matter their significance, are quickly overshadowed by failures: progress is systematically a reminder to those who make it of how many of their goals are far from being achieved. To understand this last point, we would now like to focus on the case of the Berkeley Artificial Intelligence Research (BAIR) Lab. The BAIR Lab is an equally insightful and interesting case of understanding the limitations and benefits of AI’s experimental nature.

Since the early 1980s, the BAIR Lab had been headed by Robert Wilensky, a former student of Schank. In keeping with Schank’s legacy, who was often considered the *enfant terrible* of AI, Wilensky did not care for formalism, whether it be logical or mathematical in nature. Unlike a number of his colleagues who saw in Noam Chomsky’s work (1965) the potential to formalise human intelligence, he did not believe language could be reduced down to formal syntax. While he agreed with Chomsky’s paradigm – according to which language is at the root of thought – he also believed that language poses a semantic problem and not a syntactical one. In other words, to reuse the Schanksian expression, Wilensky was “scruffy”. He had an inclination for tricky problems, and developed a take on AI in his own image, being both original and bold. Indeed, Wilensky was an unusual individual, often considered a non-conformist by a fair number of his colleagues (Rose, 1986 [1984]). He liked originality, and it was probably because of this that he was drawn to the prevailing intellectual climate at the University of Berkeley: while the institution did not have a true computing culture when Wilensky arrived, it fostered an intellectual diversity that he appreciated. At Berkeley you could find anyone: idiosyncratic anthropologists, non-conformist linguists, cognitive psychologists, and, most importantly, Hubert L. Dreyfus and John R. Searle who played an active role in stimulating Wilensky’s research. With their unrelenting criticism of AI, the two philosophers in fact fuelled the BAIR Lab in its work, and raised the profile of its director. For Wilensky, who

<sup>(6)</sup> The passages in upper case are spoken by SHRDLU.

was interested in commonsense reasoning and the issue of basic actions, Berkeley was the perfect testing ground for new ways of understanding AI. In order to fully grasp the experimental nature of the research projects conducted at the BAIR Lab, a presentation of some of the main programs developed by Wilensky and his colleagues is provided below.

Much like the Script Applier Mechanism (SAM) developed by Schank and Abelson (1977), Wilensky's Plan Applier Mechanism (PAM; 1977) had a certain capability of understanding narratives and situations that it was told. In 1980, PAM was able to have the following discussion:

"[Based on the following description:] John needed money, he got a gun and walked into a liquor store. John told the owner he wanted his money. The owner gave John the money and John left." [... and based on the following question:] "Why did the owner give the money to John?" [... PAM answered with, for example:] "The owner was scared that John would kill him" (Rose, 1986 [1984], p. 71).

Like SAM, PAM was able to exhibit a certain degree of understanding in that it demonstrated knowledge that was not explicitly contained in the statement it was told. However, PAM has a certain edge over SAM: to understand a situation, PAM did not need its creator to provide it with the underlying scenario. Naturally, much like Winograd's SHRDLU program (1972), PAM is only capable of such a feat provided that the statements it was told relate to its micro-world. Furthermore, as demonstrated by Frank Rose (1986 [1984]) with the case of "Plan ANalyzer with Dynamic Organization, Revision and Application" (PANDORA), expanding this micro-world requires a myriad of computing tricks that are just as much ways of questioning the workings of human intelligence. For example, Joe Faletti (1982), a student of Wilensky who developed PANDORA, struggled to make his program understand that the act of going to fetch a newspaper from the letterbox may require different behavioural patterns depending on the weather. Cognitively speaking, such a capacity to understand and adapt relates to significant planning issues, particularly in terms of organising the goals and sub-goals of a particular act and its constituent tasks, but also in terms of memorising relevant information – organising and applying knowledge to perform every task (Faletti, 1982). For instance, in order for PANDORA to put on a coat, it had to know that rain is wet and that being dry is a desirable state, but also that a coat protects from the rain. As odd as it may seem, for Faletti, this sort of problem was equally as important as it was difficult to resolve from a computing standpoint. The threefold benefit of the work of Wilensky and his colleagues is evident in this respect as well. Together, they underscored that:

- as basic as it may seem, an action entails different forms of problem solving which, despite being automatic in nature, are cognitively complex;
- these forms of problem solving are inseparable from the social conventions that existed before the given action;

- the coordination of cognitive and social aspects in completing any human action (even the most trivial ones) entails a form of intelligence that is extremely difficult to identify, describe, understand and formalise.

In opposition to the simplistic discourse on AI that often emerges, the case of the BAIR Lab is proof that AI is not merely a community of researchers wanting to impose their logician and mathematician viewpoints by applying them in the field of humanities and social sciences. For Wilensky and his colleagues, computers are an implement for scientific experimentation, with the heuristic benefit of helping them to question and understand what intelligence is. However, this point of view is not specific to the BAIR Lab: for example, as already noted, Simon, Newell and their Carnegie Mellon students share this viewpoint. The case of the BAIR Lab is also of interest to us for another reason, one that is embedded in this critical and original vision of AI that Wilensky and his colleagues adopted. Following on from the work of Minsky on frames (1974), PANDORA<sup>(7)</sup> was a method of representing a notion of intelligence with a computer, which used Searle's Background theory (2002). According to this theory, language is a code whose meaning cannot exist without the social conventions that enable its expression. In this respect, history has shown that Wilensky and his team failed in their project to design a computer program capable of simulating basic cognition. Can this lead us to conclude that their research program was a failure? It depends on who we ask. Investors like DARPA or IBM would say yes: it was a computer program that merely worked within the bounds of a micro-world created by a researcher, with no political or economic application. In contrast, a sociologist interested in the history of science and technology would clearly say no. Naturally, with its experiments, the BAIR Lab was unable to confirm the hypothesis that basic cognition can be represented by computers. However, we see this failure to be a huge success, since this unsuccessful venture was a stepping stone for Wilensky to more effectively test out the complexity of the interaction between cognition and culture, the difficulty in representing this complexity with computers, and in particular the concept that the effectuation of this complexity is required to correctly perform, analyse and understand the smallest basic action.

As fragile as it may be, the cognitive value behind this conclusion is particularly high since, following on from the work of Simon and Newell, it raises the question of how we conceive intelligence. The experiments conducted at the BAIR Lab lead to a hypothesis being formed: while Simon and Newell quickly managed to produce satisfactory simulations of expert cognition, this was because, contrary to the belief of potentially most academics, this cognition was probably less intelligent than it seemed. While it may seem outrageous, Wilensky and his colleagues were not so sure that solving half of the Principia Mathematica required more brainpower than going to pick up mail from a letterbox.

<sup>(7)</sup> As well as the PAMELA program designed by Peter Norvig to supplement it (Faletti, 1982).

## AI as an economic practice

We have seen that AI builds on cybernetics and the history of ideas that form it, meaning that AI emerged with the development of the first computers: its origins coincide with those of computing which, we stress, embraces scientific, technological and industrial areas of activity. In this respect, it is important to keep in mind that AI is inextricably linked to the major socio-technological innovations that paved the way for the computerisation of society (Mounier-Kuhn, 2010). In the 1950s, these innovations were prohibitively expensive, which meant pioneers in AI had to quickly partner with political and economic stakeholders to fund their research. This was nothing terribly new: the forefathers of AI had to do the same thing. For example, Turing worked with the British government to crack the enigma code and help the Allied powers defeat the Nazis (Hodges, 2015 [1983]), John von Neumann collaborated with the American government to enhance the explosive power of the atomic bomb and provide the Americans with a tool to intimidate the Soviet Union (Hoddeson et al., 1992) and, more generally, in the aftermath of the two world wars, mathematicians and cyberneticians attending the Macy Conferences wanted to help establish a new world order to guarantee peace among mankind and the “mental well-being” of the people i.e. their autonomy and intellectual freedom (Heims, 1991).

In other words, building on the arguments put forward previously, it must be stressed that to have a career in AI, you cannot just be a renowned researcher who is respected by your peers; you have to also be able to draw in investors to receive funding for as long as possible (Latour, 1987). To do this, AI researchers had to navigate the political and economic spheres, particularly because DARPA was the biggest source of funding for this field for quite some time. It was precisely through this specific form of “economisation” (Akrich, 1989) that AI was able to enter into the public forum. Many controversies have as a result surfaced, leading to spillovers into other areas that have obfuscated the collective understanding of what AI is. We would like to examine the history of AI in relation to the market in order to better understand why this obfuscation was able to take root, and also what its impacts are on the advancement of this science.

### Promises to draw in investors

AI is a scientific field torn between the three points of tension detailed above (see section Disappointment and fears). It is important to understand that, in order to compete within such a tumultuous field, AI researchers had to find partners who could provide them with the suitable technological and financial resources to conduct their research programmes. They were therefore collectively compelled to lay down bridges between the scientific, political and economic worlds so that the above-mentioned divergent viewpoints could exist. We have seen that, from the outset, AI breeds a turbulent working environment: while AI researchers know how to play nice, especially when searching for partners who will help them conduct their work more effectively,

they are also well aware that in order to achieve their career goals they will have to jostle for position. The leading academic institutions and universities with an AI laboratory foster this competitive environment for at least two reasons. The first one is that, as we have seen with the MIT AI Lab, these laboratories may on occasion pledge several millions of dollars per year to the institutions and universities hosting them. The second is that, given the military-industrial complex’s interest in AI, these very institutions and universities strive to draw in the most esteemed researchers. As a result, in AI, science and the market end up sustaining each other to form a particular “opinion economy” (Orléan, 2000) in which the scientific value of the research programme conducted by a given laboratory is not the only factor that matters any more: there is also – and above all – the element of the researchers’ ability to flex their muscles before their peers, raise positive public interest, and draw the attention of political and economic stakeholders and build trust with them.

Within this highly competitive environment, arrogance can at times give an edge. This is what at least seems to be the case with Simon and Newell, who, as previously mentioned, always received considerable recognition in the field of AI. The two associates from Carnegie Mellon in fact had a reputation for self-importance. For example, in an interview, Minsky told Daniel Crevier (1997 [1993]) that Simon and Newell came across as aloof during the Dartmouth Summer Research Project. The other attendees believed that the two researchers seemed just as pleased as they were flattered that they were the only ones to present an AI program. Simon himself would go on to confirm this observation (Crevier, 1997 [1993], p. 67). In 1997, 40 years after the famed conference, Simon and Newell’s confidence in their work had not at all faltered, and in fact had only bloated. In his 1991 book, Simon said that, with their invention of a computer program capable of processing symbolic data, he and Newell had demonstrated how a system composed of matter can exhibit the attributes of thought. In Simon’s view (1991), their work held the key to unlocking the mystery of the dualism of the mind and body. This claim is naturally subject to debate, as demonstrated by the work of Daniel Dennett (1991).<sup>(8)</sup> Nevertheless, as questionable as it is from a scientific standpoint, the claim is a good reflection of the degree of confidence one has to deal with when working with AI researchers. In the vein of Simon and Newell, AI pioneers are researchers with key expertise in mathematics and computing, but also in humanities and social sciences. To make it in this extremely competitive environment, researchers have to learn to showcase their expertise and unique qualities to political and economic stakeholders. This is why, to impress investors while securing their full trust,<sup>(9)</sup> AI researchers have to provide a high level of assurance in relation to their work. In this respect, a considerable

<sup>(8)</sup> It is important to stress that this is not an attempt to disregard the key role played by Simon and Newell in the development of the philosophy of the mind.

<sup>(9)</sup> Note that the sums of money involved are huge, and it has been known, even in the 1960s, for funding exceeding \$1 million per year to be provided.

number of eminent AI researchers since the late 1950s have had to collectively make predictions that often err on the rosy side. Once again, Simon and Newell are among the researchers who did not hesitate to abuse their scientific authority to give credence to overly ambitious promises:

“1. [That] within ten years a digital computer will be the world’s chess champion, unless the rules bar it from competition. 2. [That] within ten years a digital computer will discover and prove an important new mathematical theorem. 3. [That] within ten years a digital computer will write music that will be accepted by critics as possessing considerable aesthetic value. 4. [That] within ten years most theories in psychology will take the form of computer programs, or of qualitative statements about the characteristics of computer programs.” (Simon & Newell, 1958, pp. 7-8).

Simon and Newell were of course discerning to some degree, as some of the predictions above were correct. However, in the strictest sense of the word, all the predictions were false and should have been lowered: for example, we would have to wait until 1997 – and not 1968 – for the Deep Blue supercomputer to win against Gary Kasparov in a chess match. For Simon and Newell however, whether their predictions would be proven true or false was not that important. The two colleagues quickly understood that applications of AI could transform into a market brimming with management technology serving all stakeholders in the production and distribution chains of goods and services, including consumers (Cochoy, Smolinski & Vayre, 2016). What mattered to Simon and Newell was that their predictions were equally as reasonable as they were rosy for the military-industrial complex with which they were very familiar. The two researchers knew how business worked, and more specifically how military and industrial business was run: Simon and Newell were also consultants for the RAND Corporation. In other words, even though they knew as researchers that their predictions were not true in the scientific sense, they knew as consultants that the predictions were promises likely to draw in investors.

### Disappointment and fears

AI has always been an unsettling field, since it attempts to understand human behaviour from an objective and detached perspective. In this respect, it is important to understand that when Turing argued, in 1950, that a machine has the potential to produce thoughts, his main intention was to shake up the intelligentsia of the time. This forefather of artificial intelligence knew that he was a homosexual at this point, and rebelled against the commonly held beliefs of his time: he was not convinced by the often religious, authoritarian and dubious lines of thinking that, for example, considered women, and to a greater extent animals, incapable of demonstrating intelligence (Turing, 1950). Why was intelligence considered sacrosanct by some? Was there anything that could rule out the theory that a machine can exhibit intelligent behaviour? Simon and Newell quickly realised that the provocative nature of AI research in itself could provide a socioeconomic edge. They knew that the market liked innovation and that it could be an invaluable partner in combating scientific

orthodoxy that could hinder the advancement of AI. At least at the beginning, and to establish this discipline as a scientific one, AI researchers could only partially rely on academic institutions: they had to find other means of securing the lasting future of AI. However, Simon and Newell were not alone in realising this. For example, Minsky in his own way helped to publicise AI, painting a more or less realistic picture of what it could produce. He did this most notably through science fiction, advising Stanley Kubrick during the filming of *2001: A Space Odyssey* (1968; Ganascia, 2019 [2016]). There are also other researchers who used other means to give publicity to AI. It is important to bear in mind that the overall goal for these researchers – expressed with varying degrees of clarity – was to promote the market expansion of this science by bringing it onto the communications market.

This basis caused a number of difficulties to emerge from a socioeconomic standpoint. The pioneers of AI most likely had developed communication channels between their discipline and the market too quickly. This hastiness was particularly due to the fact that they felt a sense of urgency since they needed powerful and costly machines to get ahead of their competitors. In view of this, while scientific competition is more strongly influenced by the political and economic dynamics of the capitalist system through its involvement on the market, it is clear that AI was going from strength to strength and was firmly established within academic spheres. In the early 1960s, a flurry of promises were being made, and investors were being hooked in. AI became a media sensation, and it entered into its golden age. However, this era was not set to last. After the highs of great expectations came the equally as notable lows of disappointment. For example, following the rather negative assessment made by the Automated Language Processing Advisory Committee (ALPCA; Pierce et al., 2019 [1966]) concerning the progress made in the field of machine translation, the US government decided in 1966 to halt investments which were initially intended to fund the translation of Soviet Union press releases (Hutchins, 1996). As previously stated, this was only the first in a long line of failures as economic and political stakeholders saw it. The Shakey robot (Nilsson, 2019 [1984]) for example had no military or industrial use, given that the tasks it could perform were slow and essentially a series of jerky movements. What is more, Shakey was very sensitive to changes to its surroundings: With just a slip of the wheels, its perception of its surrounding environment would no longer correspond with the actual situation (Hart & Nilsson, 1972). Speech Understanding Research (SUR) by Donald E. Walker (2019 [1973]) was also another attempt in vain to find an application for AI. This system was not viable since its users had to severely restrict their grammar usage so that SUR could process their request in real time. Ultimately, this technology was more difficult to use than the traditional menu selection systems (Crevier, 1997 [1993]). This was also the case for expert systems which, at least in the 1980s, were a great success however. One such system was XCON (Bachant & McDermott, 1984): after a number of years, updating its knowledge base became a true ordeal. In the words

of Cardon and his colleagues (2018), XCON turned into a “cathedral of rules”: this expert system was so complex that the Digital Equipment Corporation (DEC) was required to invest over \$2 million per year for its maintenance (Simon, 1987).

Nevertheless, AI has not only just been a disappointment to investors: from the outset, it has also struck fear in consumers. To understand this phenomenon, it must be first noted that the communications market revolving around AI is lucrative, particularly as it draws up fantasies, promises, but also substantiated risks of varying degree. The problem is therefore that, in the eyes of the public, this market generates a mix of information that turns AI into a catch-all and troubling concept. As the following excerpts from interviews show, AI ended up scaring consumers, and this fear can be understood in different ways.

For Michel Melkanoff for example, this fear is irrational, since the risk of AI comes not from the machines themselves but the people designing and using them.

“There are those who are afraid of machines [...] that [...] will turn into superhuman robots who will take over the world [...]. I have something to say about that. [...] [Nobody] can truly have serious concerns [...] over a bunch of wires and metal, it is an irrational fear. Interviewer: “The atomic bomb is a bunch of wires and metal too!” Sure, but it’s not the bomb that people are afraid of, it’s the people dropping them. In this respect, there is perhaps a threat posed by those able to use computers” (Michel Melkanoff, quoted in Moreuil, 1972).

Abraham Moles has a different view on the matter. For the computer and communications science expert, this fear is rather the result of what he calls a “sociological concern”, a fear of varying rationality among people relating to the forms of alienation caused by using AI:

“The public is afraid of machines [...] as they reveal their nature [...] and [...] pervade our day-to-day life [...]. As René de Possel noted, when 47 million French citizens will be classified under 1,000 or 2,000 criteria, each one stored on a punch card, no more police files, no more proceedings, everything stored in a central registry, then you will be able to identify every individual. They will no longer be anonymous. They will be [...] personalised, not able to rely on interstitial freedom or the workings of institutions. They will be prisoners! I believe that this is why [...] people are afraid” (Abraham Moles, quoted in Lallier, 1963).

In reference to the work of Madeleine Akrich, Michel Callon and Bruno Latour (2006), it is therefore through a more or less mastered economic practice that AI researchers transformed their work into promises and applications to interest and bring on board investors. This transformation was then amplified by the communications market revolving around AI. The problem is that, from the outset, this amplification was a distortion obscuring the true nature of AI. The public ended up forgetting for quite some time that AI is first and foremost a science. It is very well possible that this is still the case today: how many people have an understanding of AI research programmes, or the epistemological, social and human issues related to the field? From the 1960s to 1980s, the public saw AI at best as a kind of mechanism or energy, in vague and incom-

prehensible terms, existing within machines to regulate their operations and developing in a more or less dangerous manner... And at worse, AI was a massive scam.

### Criticisms to reassure and shake up the market

For a large section of the population, the concept of AI lost its meaning. It became a source of discomfort, but not really one for the researchers: even those who were not completely happy with this notion got used to it rather painlessly (see the section “Different styles of research”). It was industry stakeholders working in the development of computing who were uncomfortable with the concept. This was the case for IBM for example:

“The AI projects carried out within the firm [IBM] were eventually a victim of their own success. [...] During a shareholders’ meeting, Thomas J. Watson was asked to explain why the company funnelled research investments into such worthless fields. The IBM marketing department had also observed an alarming change in consumer psychology: they considered computers a threat and abandoned them out of fear. [F]or Watson, this was the last straw [...]. The firm’s future marketing campaigns [...] threw away the image adopted from science fiction of a computer acting as a giant brain and replaced it with one that was reassuring, of a machine simply processing figures. Computers, IBM unflaggingly claimed, [...] would only do what they were told. They would never oust an executive, as their sole talent was in quickly processing massive data flows” (Crevier, 1997 [1993], p. 78).

Industry stakeholders’ discomfort with AI worsened with the many disappointments previously mentioned, to the extent that, as previously mentioned, economic and political stakeholders questioned their commitment to developing this field. These stakeholders therefore took a genuine interest in the criticisms of AI. Stuart E. Dreyfus, a consultant from the RAND Corporation, took this opportunity to put his brother Hubert L. Dreyfus into contact with the research organisation. Hubert L. Dreyfus was called upon to assess, from a philosophical standpoint, the viability of the AI project: the RAND Corporation wanted him to predict this field’s ability to confirm the theory that behaviour deemed intelligent by humans can be materially replicated. After his investigations, Dreyfus (1972) gave a resounding no: he believed that intelligence bore no relation to a system that computes symbolic representations and does not entail the performance of logic operations. Dreyfus specifically felt that in contrast to humans, machines crudely perform calculations: they are unable to distinguish between what is relevant and what is not (a problem of restriction). According to the philosopher, even though humans can translate the complexity of the world into simple responses, this same complexity has to be reduced, formalised and made plain for a machine to be able to respond similarly. In Dreyfus’ view, this was an impossible task, at least for symbolic AI which was prevalent at the time. Along these lines, he added that while humans have no trouble adapting to changing environments, this does not extend to machines, which only know how to follow explicit rules (a problem of framework). In summary, Dreyfus believed that AI could not qualify as a science given the irrational nature of its

inherent research hypothesis. AI was, in his view, a sort of overambitious alchemy. More generally, *Alchemy and Artificial Intelligence* (Dreyfus, 2019 [1965]) provides vociferous criticism of AI: it is a provocative and poorly documented report that goads readers into a scrappy debate. This paper rendered such criticism irrelevant and clumsy. Even Joseph Weizenbaum, one of the very few AI researchers to side with Dreyfus, thought that his colleague's report was poor, particularly because it demonstrated a meagre understanding of how computers work (Crevier, 1997 [1993]).

Nevertheless, by seeking the services of Hubert L. Dreyfus, the RAND Corporation legitimised his ideas and contributed to distorting the scientific debate into a socioeconomic controversy which quickly turned into an armchair debate. During such discussions, scientific arguments were thrown out for crude insults. For example, in response to Dreyfus' provocation that a six-year old could beat any computer program at chess – which, at a point in time, was actually the case – Papert publicly challenged the philosopher to beat such a child in a game.<sup>(10)</sup> Of course, Dreyfus was right in many respects, and his criticism was fundamentally interesting: his poor understanding of AI was offset by his strong knowledge of philosophy. It was this knowledge that allowed him to considerably beat McCarthy to the punch in identifying the two major limitations of AI: the aforementioned problems of restriction and framework. In any case, whether he was right or wrong matters little. The important thing is to understand that starting with *Alchemy and Artificial Intelligence* (Dreyfus, 2019 [1965]), the controversy surrounding AI spilled over from the field of science to become a socioeconomic issue.

Generally speaking, this problem relates to two major factors. The first concerns a collective form of ramping up commitments (Joule & Beauvois, 2002). This refers to the idea that some researchers, who were probably far too committed to AI development, continued to keep to their promises (so as not to lose face) while also seeking to regain the market's trust (and more specifically investors' trust). The second factor concerns a fictitious reconstruction of reality. It brings together the AI critics who wanted to play the game of industry stakeholders like IBM by denying the existence of AI. The problem was therefore that these critics had forgotten, more or less willingly, something very important: AI is not some form of mechanical autonomous thinking that computer-related technologies could develop. It is a science, the existence of which is hard to deny. The socioeconomic issue of AI is therefore twofold: on one hand (for the escalation of commitments), the issue is linked to the fact that this science's funding was reliant on bluffs that risked perpetuating the cycle of promise/disappointment until this disrupted the positive results of scientific programmes in this field; and, on the other hand (for the fictitious reconstruction of reality), the issue is related to the fact that the strategy consisting of distorting what is AI to have better grounds to deny its

<sup>(10)</sup> For more details on the concrete forms of this debate, please refer to Papert's report (2019 [1968]), written in response to *Alchemy and Artificial Intelligence* (Dreyfus, 2019 [1965]).

existence was a method of masking its development. This concealment gave vendors considerable power: without even truly realising, they became the only stakeholders who could ensure and control not only the dissemination of AI technology applications, but also the means of its funding as a result.

Following the prevalence of the commitment escalation problem for several years (AI winters), the second factor of the socioeconomic problem of AI (the fictitious reconstruction of reality) came to the fore from the 1980s. As shown in Figure 2, during this period, IT companies no longer wanted to talk about expert systems, and even less so AI. They preferred to act as mere IT solution providers in order to seem both far-sighted and diligent in the eyes of their customers. Consequently, AI programs became hidden applications, being discreetly integrated into more traditional computer programs. Patrick H. Winston for example was very familiar with this strategy. In the 1980s, like most of his colleagues, he owned a computer program development business, explaining that the programs were based on what he called a "raisin bread" system:

"AI is currently integrated into systems like raisins in a loaf of raisin bread: the raisins do not occupy much space, but they often provide the principal source of nutrition. You cannot remove the raisins from the bread; and there are many types of raisins" (Patrick H. Winston, quoted in Crevier, 1997 [1993], p. 252).

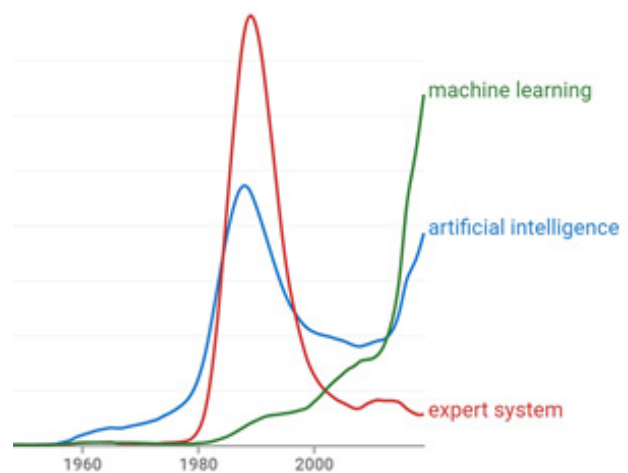


Figure 2. Graph of the number of occurrences of the terms "artificial intelligence", "expert system" and "machine learning" in Google Books Ngram Viewer<sup>(11)</sup>

A typical example of the "raisin bread" system is the commercial assistance program. Traditionally, this program would just check product availability, record the transaction, draw up the invoice, and notify the shipping service provider. This program was also able to be enhanced with a specially designed expert system

<sup>(11)</sup> The lines forming this graph are a sort of sounding board of empirical reality and therefore have a slight lag behind the figures presented in this paper, given that they are generated using statistical analysis of the corpora of texts available on Google Books. For instance, the peak of the occurrences of the term "expert system" is in 1988, while the golden age of expert systems in the US was actually from the late 1970s to early 1980s (Crevier, 1997 [1993]).



that could, for example, make suggestions for substitute products in the event of shortages. The benefit of this new way of conceiving the AI business model was that by concealing the existence of the AI technology, the media frenzy surrounding this science gradually calmed down. The confusion over what is AI dissipated at the same time as any concerns, fears and related risks. There was a drawback however: this discipline and the technological applications it created would continue to exist. As previously mentioned, the concern would still linger because the organisation of the dissemination of AI technology applications shifted between the hands of economic stakeholders whose interest should not be mistaken with those of society. It would take, as we have witnessed in the past decade, the significant and rapid increase in digital data production, storage and processing capacities – which paved the way for a new age of machine learning<sup>(12)</sup> (see Figure 2) – for society to be aware of this issue and once again question the economic, social and human stakes of AI development and of the dissemination of related applications (see the big data movement; Cardon, 2015; Vayre, 2016).

## Conclusion

AI is a scientific discipline with a research programme that was, at least at the beginning, highly experimental: it tests the hypothesis that the intelligence of humans – and, by extension, of all living beings – can be materially replicated. History has shown us that, from a purely scientific standpoint, AI had the merit of contemplating what intelligence is and therefore brought about major developments in not only the field of cognitive sciences but also humanities and social sciences. We posited that working in AI was not just engaging in scientific activity, but also in an economic practice. To carry out work in this science, there is a need to draw in investors who can fund costly equipment: in AI, science and the market are inextricably linked. This is why, in a socioeconomic context in which the major digital stakeholders (GAFAM – Google, Apple, Facebook, Amazon and Microsoft) tend to draw in leading AI researchers, we believe that it is worth stressing that historically AI was first and foremost an experimental scientific discipline which seeks to better understand what intelligence is and how we can (or cannot) synthesise it. In this respect, our work has the advantage of highlighting that, for pioneers, AI was not a consumer good that leads to the development of automated services like, as is the case currently, the platform economy for example: it was a science that lets researchers ask fundamental questions that could result in successful applications.<sup>(13)</sup> However, our studies have revealed that, from the outset, AI has also been an economic practice consisting of dressing up scientific ideas in a political and economic fashion so as to link them to socio-technical uses. We have sought to show how this window dressing has historically driven

push and pull dynamics with economic stakeholders (i.e. investors and consumers). It is a factor that should be better understood, particularly if we want to have a greater insight and grasp of the actual stakes of AI development.

## References

- AKRICH M. (1989), « La construction d'un système socio-technique. Esquisse pour une anthropologie des techniques », *Anthropologie et Sociétés*, 13(2), pp. 31-54.
- AKRICH M., CALLON M. & LATOUR B. (2006), *Sociologie de la traduction : textes fondateurs*, Presses des Mines, Paris.
- BACHANT J. & McDERMOTT J. (1984), "R1 revisited: four years in the Trenches", *AI Magazine*, 5(3), pp. 21-32.
- BOURDIEU P. (1976), « Le champ scientifique », *Actes de la recherche en sciences sociales*, 2(2-3), pp. 88-104.
- BRUSCAGLIONI L. (2016), "Theorizing in grounded theory and creative abduction", *International Journal of Methodology*, 50(5), pp. 2009–2024.
- BUCHANAN B. G. & FEIGENBAUM E. A. (1978), "Dendral and meta-dendral: their applications dimension", *Artificial Intelligence*, 11(1-2), pp. 5-24.
- BUCHANAN B. G. & SHORTLIFFE H. E. (1984), *Rule-based expert systems: the mycin experiments of the stanford heuristic programming project*, Addison Wesley, Boston.
- BURGIÈRE A. (1979), « Les Annales, 1929-1979 », *Annales. Économies, sociétés, civilisations*, n°6, pp. 1344-1346.
- CARDON D. (2015), *À quoi rêvent les algorithmes : nos vies à l'heure des big data*, Paris, Seuil.
- CARDON D., COINTET J.-P. & MAZIÈRES A. (2018), « La revanche des neurones : l'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, 5(211), pp. 173-220.
- CHANGEUX J.-P. (1983), *L'Homme neuronal*, Paris, Fayard.
- CHOMSKY N. (1965), *Aspects of the theory of syntax*, MIT Press, Cambridge.
- COCHOY F., SMOLINSKI J. & VAYRE J.-S. (2016), "From marketing to 'market-things' and 'market-ITing': accounting for technicized and digitalized consumption", in CZARNIAWSKA B. (dir.), *A research agenda for management and organization studies*, Edward Elgar, Glasgow, pp. 26-37.
- CREVIER D. (1997 [1993]), *À la recherche de l'intelligence artificielle*, Paris, Flammarion.
- Davis R. (1978), "Knowledge acquisition in rule-based systems – knowledge about representations as a basis for system construction and maintenance", *Pattern-Directed Inference Systems*, pp. 99-134.
- DENNETT D. (1991), *Consciousness explained*, Little, Brown and Co, Boston.
- DREYFUS H. L. (1972), *What computers can't do: a critique of artificial reason*, Harper & Row, New York.
- DREYFUS H. L. (2019 [1965]), *Alchemy and artificial intelligence*, available online: <https://www.rand.org/content/dam/rand/pubs/papers/2006/P3244.pdf>
- FALETTI J. (1982), "PANDORA - A program for doing common-sense planning in complex situations", *AAAI-82 Proceedings*, p. 185-188.
- FLAMM K. (1987), *Targeting the computer: government support and international competition*, Brookings Institution, Washington.
- FODOR J. & PYLYSHYN Z. W. (1988), "Connectionism and cognitive architecture: a critical analysis", *Cognition*, 28(1-2), pp. 3-71.
- GANASCIA J.-G. (2019 [2016]), « Marvin Minsky : un des cerveaux de l'intelligence artificielle », available online: <https://interstices.info/marvin-minsky-un-des-cerveaux-de-lintelligence-artificielle/>

<sup>(12)</sup> Major successes in this field include the AlphaGo, Watson and DeepL Translator programs respectively developed by DeepMind, IBM and DeepL.

<sup>(13)</sup> Even though these applications were actually rather unsuccessful at the beginning.

- GELERTER H. L. & ROCHESTER N. (1958), "Intelligent behavior in problem-solving machines", *IBM Journal of Research and Development*, 2(4), pp. 336-345.
- GINZBURG C. (1980), « Signes, Traces, Pistes – Racines d'un paradigme de l'indice », *Le Débat*, 6(6), pp. 3-44.
- HART P. & NILSSON N. (1972), *Shakey: experiments in robot planning and learning*, available online: <https://www.youtube.com/watch?v=GmU7SimFkpU>
- HEIMS S. J. (1991), *The cybernetics group*, MIT Press, Cambridge.
- HODDESON L., HENRIKSEN P. W., MEADE R. A. & WESTFALL C. (1992), *Critical assembly: A technical history of Los Alamos during the Oppenheimer years, 1943–1945*, Cambridge University Press, Cambridge.
- HODGES A. (2015 [1983]), *Alan Turing : le génie qui a décrypté les codes secrets nazis et inventé l'ordinateur*, Paris, Payot.
- HUTCHINS J. (1996), "ALPAC: the (in)famous report", *MT News International*, n°14, pp. 9-12.
- JOULE R.-V. & BEAUVOIS J.-L. (2002), *Petit traité de manipulation à l'usage des honnêtes gens*, Grenoble, Presses Universitaires de Grenoble.
- JULIA L. (2019), *L'intelligence artificielle n'existe pas*, Paris, First.
- LALLIER J. (1963), *Visa pour l'avenir : des machines à penser ?*, available online: <https://www.ina.fr/video/CPF86656441/des-machines-a-penser-video.html>
- LATOUR B. (1987), *Science in action: how to follow scientists and engineers through society*, Harvard University Press, Cambridge.
- LENAT D. B. (1977), "Automated theory formation in mathematics", *Proceedings of the 5th international joint conference on Artificial intelligence*, vol. 2, pp. 833-842.
- LENAT D. B. (1983), "Eurisko: a program that learns new heuristics and domain concepts", *Artificial Intelligence*, 21(1-2), pp. 61-98.
- MCCARTHY J., MINKSY M. L., ROCHESTER N., SHANNON C. E. (2019 [1955]), *A proposal for the Dartmouth summer research projet on artificial intelligence*, available online: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>
- McDERMOTT J. (1982), "R1: a rule-based configurer of computer systems", *Artificial Intelligence*, 19(1), pp. 39-88.
- MINSKY M. L. (1954), *Theory of neural-analog reinforcement systems and its application to the brain model problem*, Princeton University, Princeton.
- MINSKY M. L. (1974), *A framework for representing knowledge*, Massachusetts Institute of Technology, Cambridge.
- MINSKY M. & PAPER S. A. (1969), *Perceptrons*, MIT Press, Cambridge.
- MOREUIL F. (1972), *Un certain regard - Les machines et les hommes : l'intelligence artificielle*, available online: <https://www.ina.fr/video/CPF86655105/l-intelligence-artificielle-video.html>
- MOSES J. (1967), *Symbolic integration*, Massachusetts Institute of Technology, Cambridge.
- MOUNIER-KUHN P.-E. (2010), *L'informatique en France de la seconde Guerre Mondiale au Plan Calcul. L'émergence d'une science*, Paris, Presses de l'Université Paris-Sorbonne.
- NEWELL A. & SIMON H. A. (1956), "The logic theory machine: a complex information processing system", *IRE Transactions on Information Theory*, 2(3), pp. 61-79.
- NEWELL A., SHAW J. C. & SIMON H. A. (1959), "Report on a general problem-solving program", *Proceedings of the International Conference on Information Processing*, pp. 256-264.
- NILSSON N. J. (2019 [1984]), *Shakey the robot*, available online: <http://www.ai.sri.com/pubs/files/629.pdf>
- NORBERG A. L. (2019 [1989]), *An interview with Marvin L. Minsky*, available online: <https://conservancy.umn.edu/bitstream/handle/11299/107503/oh179mlm.pdf?sequence=1&isAllowed=y>.
- ORLÉAN A. (2000), « L'individu, le marché et l'opinion : réflexions sur le capitalisme financier », *Esprit*, 11(269), pp. 51-75.
- PAPER S. A. (2019 [1968]), *The artificial intelligence of Hubert L. Dreyfus: a budget of fallacies*, available online: <https://dspace.mit.edu/bitstream/handle/1721.1/6084/AIM-154.pdf?sequence=2&isAllowed=y>
- PAPER S. A. (1971), *A computer laboratory for elementary schools*, Massachusetts Institute of Technology, Cambridge.
- PASSERON J.-C. (1991), *Le raisonnement sociologique. L'espace non-poppérien du raisonnement naturel*, Paris, Nathan.
- PIERCE J. R., CARROLL J. B., HAMP E. P., HAYS D. G., HOCKETT C. F., OETTINGER A. G. & PERLIS A. (2019 [1966]), *Language and machine: computers on translation and linguistics*, available online: <http://www.mt-archive.info/ALPAC-1966.pdf>
- PRATT V. (1995 [1987]), *Machine à penser : une histoire de l'intelligence artificielle*, Paris, Presses Universitaires de France.
- ROSE F. (1986 [1984]), *L'intelligence artificielle : histoire d'une recherche scientifique*, Paris, Payot.
- ROSENBLATT F. (1958), "The Perceptron: a probabilistic model for information storage and organization in the brain", *Psychological Review*, 65(6), pp. 386-408.
- SCHANK R. C. (1978), "Computer understanding of natural language", *Behavior Research Methods & Instrumentation*, 10(2), pp. 132-138.
- SCHANK R. C. & ABELSON R. P. (1977), *Scripts, plans, goals, and understanding*, Lawrence Erlbaum, Hillsdale.
- SEARLE J. R. (2002), *Consciousness and language*, Cambridge University Press, Cambridge.
- SELFRIDGE O. G. (1959), "Pandemonium: a paradigm for learning", *Proceedings of the Symposium on Mechanisation of Thought Processes*, pp. 511-529.
- SIMON H. A. (1945), *Administrative behavior: a study of decision making processes in administrative organization*, The Free Press, New York.
- SIMON H. A. (1991), *Models of my life*, Basic Books, New York.
- SIMON H. A. & NEWELL A. (1958), "Heuristic problem solving: the next advance in operations research", *Operations Research*, 6(1), pp. 1-10.
- SIMON R. (1987), "The morning after", *Forbes*, 140(8), pp. 164-168.
- SLAGLE J. R. (1961), *A heuristic program that solves symbolic integration problems in freshman calculus: symbolic automatic integrator (SAINT)*, Massachusetts Institute of Technology, Cambridge.
- TAYLOR W. K. (1956), "Electrical simulation of some nervous system functional activities", *Information theory*, n°3, pp. 314-328.
- THOUARD D. (2007), *L'interprétation des indices. Enquête sur le paradigme indiciaire avec Carlo Ginzburg*, Villeneuve-d'Ascq, Presses Universitaire du Septentrion.
- TURING A. M. (1950), "Computing machinery and intelligence", *Mind*, vol. 59, pp. 433-460.
- VAYRE J.-S. (2016), *Des machines à produire des futurs économiques : sociologie des intelligences artificielles marchandes à l'ère du big data*, Toulouse, Université Toulouse Jean Jaurès.
- WALKER D. E. (2019 [1973]), *Speech understanding research*, available online: <https://apps.dtic.mil/dtic/tr/fulltext/u2/757049.pdf>
- WIDROW B. & HOFF M. E. (1960), "Adaptive switching circuits", *1960 IRE WESCON Convention Record*, n°4, pp. 96-104.
- WILENSKY R. (1977), "PAM - A Program That Infers Intentions", *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, vol. 1, pp. 1-15.
- WINOGRAD T. (1972), "Understanding natural language", *Cognitive Psychology*, 3(1), pp. 1-191.
- WINSTON P. H. (1970), *Learning structural descriptions from examples*, Massachusetts Institute of Technology, Cambridge.

# From Organisational Silence to Deontic Anger: When Whistleblowers Speak Up

by Philippe Jacquinot

associate professor, Paris-Saclay University

and Arnaud Pellissier-Tanon

associate professor, University of Paris I Panthéon-Sorbonne

Original article in French published in *Gérer & comprendre*, December 2021, n° 146, pp. 27-38.

A whistleblower speaks up: they claim to have witnessed a breach of a rule in force in their social setting and report it to the competent regulatory authority, even at the risk of incurring the displeasure of the alleged perpetrators. The literature explains the decision to alert as a calculation that weighs up the advantage of putting an end to the violation against the disadvantage of risking exposure by potentially breaking the organisational silence. It adds that the purpose of the disclosure is seldom achieved and whistleblowers are often subject to retaliation (harassment, isolation or even dismissal). How can the prosocial advantage of having the violation curtailed or any possible personal advantage resulting from the alert outweigh the risk of major personal losses? This analysis of a biographical account of a whistleblower puts forward the concept of “deontic anger” which was provoked by organisational behaviour in lawsuits. Decision-making by one possible type of whistleblower (who, sensitive to values, has prosocial motivation) is analysed. Their decision to get higher authorities involved stems from deontic anger driven by their sense of duty when they think they have witnessed a violation. In conclusion, it is pointed out that it would be worthwhile for research on whistleblowing to take account of emotional intelligence and practitioners should understand the reasons behind the anger, justified or not, that moves certain employees to take action.

Whether or not a company’s ethical guidelines include a formal reporting procedure, the literature written in English defines whistleblowing as “the disclosure by organization members (former or current) of illegal, immoral or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action” (Near and Miceli, 1985, p. 4). Whistleblowing is underpinned by the belief, which is common in American culture, in the effectiveness of remedial actions taken at the initiative of each and every one, in addition to, or at the same time as, actions by the authority (Charreire Petit and Surply, 2008). The 2002 Sarbanes-Oxley Act, passed in the wake of scandals such as that having caused the fall of Enron, forced listed companies in the United States to introduce a procedure governing whistleblowing in all their subsidiaries.

The practice spread to Europe under the initial impetus of the presence of American subsidiaries of groups listed in the United States and European companies listed in the US. Many other firms subsequently rolled out in-house whistleblowing systems to increase legitimacy vis-à-vis their stakeholders (Pittroff, 2014)

or, at least, to avoid losing their “license to operate” (Cramer, 2002, p. 103) in the event of the public revelation of questionable practices (Heineman, 2007). A number of European countries, beginning with the United Kingdom, have adopted provisions to spur the disclosure of wrongdoing in companies and government departments (Boyer, 2013).

However, in 2010, the Parliamentary Assembly of the European Union criticised the fact that most Member States had no comprehensive laws for the protection of whistleblowers<sup>(1)</sup> and it was only in 2019 that the Council approved a directive in this respect<sup>(2)</sup> and asked Member States to ensure that enterprises having 50 or more workers and municipalities with at least 10,000 inhabitants implemented effective reporting channels.

<sup>(1)</sup> Resolution 1729: <http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-EN.asp?fileid=17851&lang=en>

<sup>(2)</sup> Directive (EU) 2019/1937 published in the Official Journal of the European Union of 26 November 2019: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2019:305:FULL&from=EN>

The literature looks into the reasons that prompt whistleblowing. Since the work of Latané and Darley (1968, 1970), it has been sharpening the analysis of the process by which the whistleblower decides to sound the alarm.

For Miceli *et al.* (2008, 2012), the process starts when a witness notes that a violation has been committed and they feel that this is detrimental for the organisation or for the wider society. The literature considers the witness's personality. Near and Miceli (1996) posit that, broadly speaking, whistleblowers are not considered as being exceptional people before they make their disclosure. Rothschild and Miethe (1999) maintain that a witness to a violation who instigates an alert has almost no sociodemographic characteristics that distinguish them from the silent observer. That said, they do claim that whistleblowing is dictated by the personal values of the whistleblowers in 79% of the cases they examined. They highlight a majority profile of whistleblower with the other profiles being driven by the promise of a reward, by fear of being sanctioned for failing to disclose wrongdoing or by personal differences with their management.

According to Miceli *et al.* (2008, 2012), the process continues when the witness notes that the people tasked with immediate regulation (line manager, ethics correspondent, local HR, auditor, etc.) do not take action to stop the disruption caused by the violation and they consider that it is their responsibility to refer the matter to a higher regulatory body. They show loyalty to all the company's external and internal principals, throughout the delegation chain from the firm as a whole to local managers, via shareholders, managers and the remainder of the hierarchy. This means that whistleblowers demonstrate prosocial behaviour (Miceli *et al.*, 1991), not vis-à-vis those involved in the violation they report but vis-à-vis the social setting which enables them to act within a given framework.

For Miceli *et al.* (2008, 2012), the decision to sound the alarm ends when the witness weighs up the benefits and risks and decides whether or not to proceed. Miceli *et al.* (2008, 2012) specify that the benefit is the discontinuation of the violation and the risk is dismissal or other personal consequences. They reason in terms of the likelihood of getting the wrongdoing stopped and the risk of retaliation. However, they fail to state how the discontinuation of the violation can be a source of satisfaction for the potential whistleblower. In this respect, the percentages identified by Rothschild and Miethe (1999) provide insight: 79% of whistleblowers are motivated by their personal values, 11% from fear of being criticised for remaining silent, 3% by their resentment of management and 2% in the hope of a promotion or raise. As regards the drawbacks, the literature suggests that whistleblowers run the risk of being excluded from their organisation, with a certain amount of emotional distress (Peters *et al.*, 2011; Park and Lewis, 2018) related to the vehemence of the group's reaction (Rothschild and Miethe, 1999) or to the loss of the benefits from belonging to that group (Charreire Petit and Cusin, 2013). As they break the law of silence (Cailleba, 2017), whistleblowers are considered to have betrayed the group's unwritten

rules (Schehr, 2008). Near and Miceli (1995) and Miceli and Near (2002) show that the probability of reaping the benefit and avoiding the risk is contingent on the whistleblower's authority vis-à-vis the perpetrators of the violation.

The literature therefore emphasises the reasons for whistleblowing and identifies the features of one type of whistleblower, namely a strong sense of values and pronounced prosocial motivation. In an exploratory study based on ten life narratives, Hennequin (2020) pinpointed four profiles on the basis of the extent of their compliance with ethics or simply the law in the company as well as the societal or organisational nature of their motivation; a strong sense of values and prosocial behaviour are flagged up. However, the issue of the actual calculation of benefits and risks is not settled. Scheetz and Wall (2019) noted that a substantial number of witnesses do not report wrongdoing in spite of the fact that, every year, the Securities and Exchange Commission pays hundreds of millions of dollars in awards. Conversely, Hennequin (2019) observed that witnesses continue proactively with their whistleblowing despite evidence of reprisals and suffer with growing anger what they feel to be an injustice (p. 8).

Hennequin (2020) paved the way for research to better understand decisions to disclose taken by whistleblowers who are sensitive to values and have social motivation: the anger that drives them when they consider that they have witnessed an injustice causes them to make a report whereas, to an external observer, the drawbacks seem to outweigh the benefits. This perspective encourages researchers to return to the field and, in particular to use qualitative methods. We have compiled the biographical account of a person claiming to be a whistleblower. This case provides an in-depth understanding of one aspect of a whistleblower's reasons for taking action. Through her statements, we highlight the role of her indignation and even her anger. We will also be examining the literature on organisational behaviour and will put forward the notion of deontic anger. We will be offering a complete analysis of a type of whistleblower and, in the conclusion, we will make recommendations for being vigilant to the ethical nature of the anger that motivates certain whistleblowers.

## Inès de Chambertin's biographical account

Inès de Chambertin was born at the end of the 1960s. She is the younger of two children. Her and her brother had the solid education of an affluent background in which it was important to maintain social status and to complete "appropriate" studies for girls and "outstanding" ones for boys. She says that "When we were brought up, we were told 'You must work, you have to succeed'". After her baccalaureate, she enrolled at university: "It was decided that I would go to Dauphine University, [... because] the course had a good reputation". This was at a time when the newspaper headlines heralded the successes of the golden boys. As the logical continuation of her studies,

## Research methodology

One of the authors came across the whistleblower when reading an article in a popular weekly magazine. As he thought that the story could interest his students taking an ethics and CSR course that he gives at Masters 2 level, the author contacted the whistleblower. The latter reacted very positively to the request to testify. She was clearly still affected by the events and wanted to recount her experience unlike what is usually the case for sensitive topics (Hennequin, 2012). This contact provided us with a real opportunity to learn as defined by Stake (1994).

We began by carrying out a documentary study based on the five press articles devoted to this whistleblower, her speech at a colloquium of the CGT trade union and a 32-minute television programme. We then compiled a biographical account as defined by Bertaux (1997) during an interview on neutral ground that lasted 7h34 over a single day including the lunch hour. The day-long interview was recorded with the whistleblower's agreement with the express goal of establishing research work. The whistleblower talked about her experiences starting with her studies and her first job. At our request, she also explained how she had been brought up and her values. She handed us a copy of the file that she had submitted to the dismissal appeals commission of the professional federation in her sector of activity (80 pages of exchanges of emails, reports of meetings and other sundry documents). The whistleblower subsequently gave a lecture as part of one of the two authors' ethics and CSR course. We did not note any contradictions in the facts stemming from the two sources and data triangulation was therefore possible. This enabled the points in the timeline of events to be specified and for the whistleblower's story to be confirmed.

We used this biographical account to analyse a "category of situations" (Bertaux, 1997, pp. 13 et seq.), that of a type of whistleblower, as well as the "social trajectory" (Bertaux, 1997, pp. 13 et seq.) which turns a person who witnesses what they consider to be a violation into a whistleblower. The data specifically focuses on how the whistleblower viewed the issue and her scope for action. It allows for a blanket analysis of this case of whistleblowing, by looking to identify what prompted the whistleblower to go ahead and make the disclosure.

The interview was fully transcribed and the names have been changed, as we undertook to do vis-à-vis the whistleblower, to ensure that the data from the interview does not influence the ongoing legal proceedings.

she joined a bank without having really thought about her career choices. She gained 18 years' experience and was promoted from assistant to senior analyst. She married Xavier de Chambertin and they had four children over a 14-year period. She went back to work after each of her periods of maternity leave. She stayed with the same banking group and worked on financial analysis, specifically counterparty risk analysis.

Inès de Chambertin changed positions during the mergers and acquisitions which enabled the group to expand its business activity, in particular onto the financial markets. She kept her responsibilities as financial analyst and alternated between positions with the retail bank and with the investment bank. When she returned from her third maternity leave in early 2007, she says that she sensed that the bank's mindset had changed. She felt out of step with her colleagues.

"When I joined the bank 20 years ago, [...] there was an image of elegance, [...] staff in smart suits. [...] You had to instil trust. People didn't entrust their money to just anyone! [...] Now, traders have anything but a good image. [...] If they look like thugs, then that's perfect. It's very well perceived as it's a sign that they will earn a lot of money and are prepared to go to any ends to succeed".

She brought her boss and colleagues back into line when, according to her, they were talking behind people's backs and this compounded her image as the old-school mother of a large family

"I had a boss who was quite nice, very cool, who came and began talking about one of our contacts and said many horrible things [about him] [which] as usual [were] very, very funny. [...] It's very easy to laugh heartily and to even elaborate. [...] It made my blood boil [...] and I said "I'm sorry Simon" – my boss's name was Simon –, "but it's against my ethics, you can't

talk like that in front of me". [... After that,] whenever he was talking about someone, he said [with a false tone of innocence]: "Ah no, Inès shouldn't be here, we mustn't shock her...We can't say that in front of her!" But I did win this combat against my boss. Nobody spoke ill [of others] in front of me".

Inès de Chambertin loves her profession and is proud to belong to her banking group. Nonetheless, she thinks about the differences that she believes exist with the practices in force a decade ago. According to her, previously, analysts had to justify the risks that they made the bank take. Now, and again according to her, it is the strictness of the rating granted to counterparties that they have to justify. [She describes, for instance,] "committees with many members where it's impossible to say who is in charge", or "everyone has a comment to make", or "heavy pressure is applied to make you think: "OK, why don't we give a higher rating"?", where "I have to constantly prove why I'm giving a negative opinion". She refers to a "reversal of the burden of proof" leading to "reckless risk-taking".

"There was therefore a tipping point. I'd found out that, from a human perspective, the bank was operating in reverse. I was shocked and had to react. [...] At that time, I wrote a short text about the problems I had noted. [...] I put forward a solution for each of these problems as I had told myself: "You have to demonstrate that there are solutions, that it's easy". [...] I sent it to quite a few people who [in fact] didn't care at all. But I also took it to a manager. [...] He told me: "Yes, it's very interesting". But then, he did nothing. [...] I said to myself [then]: "You have to act yourself". [...] I waged my little battle by email which was not seen from the exterior. [...] I didn't allow myself to judge whether the actions of my line manager and second line manager were good or bad. I simply asked them to assume their responsibilities".

"In March 2010, a new manager arrived. His lack of scruples and conscience were unbelievable. The guy went full steam ahead". She said that her new boss stepped up productivity requirements. Everyone was supposed to study the files more rapidly. He restructured the department by gradually eliminating the assistants' positions. The analysts had to do their research themselves and submit their ratings directly. Inès de Chambertin was surprised by what she saw as a breach of the ethical four eyes principle. She claims that her boss criticised her ratings as being too low without ever substantiating this.

"He wanted to force my hand on a bad file but I argued [and] told him: "As a risk analyst, who defends the bank's long-term interests, meaning the interests of our depositors, I consider that it's dangerous to grant this limit". Deep down, I was really angry. [...] Having someone with real awareness of risks in a risk department changes that department: risk awareness will return and the staff will be happy because they'll finally be able to start working again!"

After her fourth maternity leave, Inès de Chambertin went back to work but says that she felt that her boss was annoyed about her return. She adds that he immediately asked her to take account of the bank's commercial interest. Time went by, she gave a number of negative opinions but, despite what she perceived as insistence from her boss, she refused to amend her ratings. She has the following comments to make on her appraisal interview in early 2013:

"[My boss] said horrible things about me and I told him "I don't agree". And then I had a kind of knee-jerk reaction [in that case...] – because women can cry, can't they – [...] I was fed up with being insulted so I cut the interview short and went back to my office. [...] He wrote [in the interview report] that I should decide on counterparty ratings "on the basis of the salespersons' interests and prudential ratios, although the logic of risks is still the main logic". That's what he said. It's interesting because there are two points here. [...] He claimed to be defending the interests of the sales departments. [...] Salespersons are supposed to receive their bonuses. But we are supposed to be independent. [...] And he also claimed to be favouring prudential ratios. [...] He told me: "So as not to undermine the bank, to benefit the bank, put higher ratings and then there'll be less equity requirements". That meant that he was asking me to lower the calculation of prudential ratios. [...] He was asking me to overvalue the ratings. [...] That's the crux of this case. This is where there's attempted corruption! He was asking me to do something unethical. [...] But, he was like that every day! He always put ratings one notch higher! Colleagues received more bonuses and all that benefitted everyone!"

She states that she constantly reminded her boss of the ethical requirements for analysts to be independent which throws up an Ethical Wall between the risk analysis department and the sales departments. According to her, her boss was irritated by this. She claims that he rewrote her analyses and reversed her conclusions. He took away her bonus and would not let her work with the department's last assistant. He ordered her to draw up the ratings that were previously established by the assistant and that, previously, she only had to check. She refused to do so and invoked the four eyes

principle. He insisted and rewrote "15 times the message: "Do your job, I'm not satisfied with your work, you must do this job". She describes how she reacted.

"I separated the human aspect from the substantive disagreement aspect as regards professional issues. And that was where the strength of my strategy lay. [...] From the outset, I refused to complain. There was the issue of the basic disagreement but I wasn't about to cry on the human side. This meant that none of the intimidation worked".

According to her, the situation became more tense. He asked his own boss (Inès de Chambertin's second line manager) to summon her "on the grounds of insubordination". In the days prior to the interview, he "shouted at her saying 'You're really in for it now'", in front of everyone, until the day when she says that he "got right up close to her" and was about to hit her. She stopped him by saying in front of the two colleagues in her office, as she recalls "You're going to end up hitting me". Her manager did not attend the interview but an "HR minion" was present.

"He had me summoned by his own line manager for insubordination. [...] The latter] shouted, shouted and shouted at me. [...] In the middle of the interview, after an hour and a half or an hour, I said [to him]: "Listen, look, as this is where we're at, I want to tell you that I'm concerned about my manager who, in my opinion, is asking me to violate the ethical definition of my duties". At the time, I was very politically correct. I didn't say: "He's corrupt". I said: "He's too business oriented. I don't believe that he has a risk-based mentality. He replied: "I'm not able to reply on the underlying issue".

Inès de Chambertin says that her second line manager did not specify how she was at fault and did not respond to her concerns about her boss.

"So I said to myself: "My second line manager has no authority in these matters so I'll go and see his line manager, the head of the risks department". [...] I said [to that manager]: "The thing is that I have concerns about such and such a file, such and such a way of calculating risks". [...] We listed the points. [...] He had an answer to everything. [...] His main argument was: "Your manager is a true professional, I'm in very regular contact with him and everything runs smoothly, everyone appreciates his work". From the outset, I replied: "But sir, obviously you appreciate his work and don't question his performance as you've allowed him to remain in his position. That's precisely why I've come to talk to you. It's because I have concerns that I'm warning you".

She says that the very next day they moved her office and isolated her one floor down. She went back to see her third line manager but, according to her, he took no further action. She states that she wrote a long letter to her fourth line manager, the deputy managing director, setting out her doubts and asking for a meeting. She says that, ten days later, she was summoned by the HRD for an interview prior to dismissal. She was fired for misconduct without having to work out her notice. The grounds: "having made serious accusations against her superiors". She considered that her behaviour was exemplary, that she exercised her rights under the collective bargaining agreement and that she complied with her duty of loyalty to her employer. She unsuccessfully submitted an application to the mixed committee of the French Banking Association. She

filed an appeal with the labour tribunal. She also says that, during the conciliation procedure, she asked to be reinstated without compensation, but that her employer refused.

“They don't care at all. From the moment they are acting illegally, they couldn't give a damn. From the moment they begin negotiating, there's no longer an ounce of justice. [...] One day, because of the weight of this injustice, people will kill themselves by jumping into the Seine. So, that's why I'm fighting. Because they're exceeding the bounds of injustice”.

She then brought an action for attempted corruption before the criminal courts.

## Understanding the whistleblower's motivation

We will now conduct a sweeping analysis of the reports instigated by Inès de Chambertin. To do so we will examine her biographical account to map out her perceptions, expectations, judgments and decisions so as to better explain her behaviour. This method could suggest that we share her point of view and agree with her moral stance whereas, in fact, we are simply relating her point of view in order to explain her behaviour. At the same time, we will be looking at the literature on the concepts pinpointed by this analysis, namely organisational silence, justice and anger.

### Three cases of whistleblowing in one biographical account

In fact, Inès de Chambertin's biographical account features three cases of whistleblowing as identified by Near and Miceli (1985), namely: 1) disclosure by a person 2) of practices deemed to be illegal, immoral or illegitimate by that person 3) under the control of their employer 4) to persons or organisations that may be able to effect action. For instructional purposes, we will be using this matrix in the order 2, 3, 4, 1, without forgetting that this is only her version of what happened.

**First case:** she brings her boss, who is so funny, back into line when he talks about people behind their backs:

- An illegitimate practice: belittling, whether backbiting or defamation, according to Inès de Chambertin.
- The employer's control: it can fight interpersonal injustices such as belittling through value-based management, in particular by the exemplary behaviour of line managers.
- The person or organisation able to effect action: if the employer or its representatives allow the belittling to continue, the opinion leaders of the group in which it is occurring can turn the situation around.
- Its disclosure: by simply asserting her values, Inès de Chambertin reminded everyone that she considered her boss's statements as belittling. She spoke up as an opinion leader against her boss and managed to stop the practice, at least in her presence.

**Second case:** she sends out her “short text” on the bank's failings:

- An illegitimate practice: “the reversal of the burden of proof”, namely encouraging, by new collective decision-making methods, a more generous rating of counterparty risks and the setting of broader risk limits than allowed by the “ethical design of risk analysis”, which led, according to Inès de Chambertin, to increased short-term turnover and profits but also to “reckless risk-taking”. In this respect, this is a distributive injustice.
- The employer's control: it introduces or restores practices enabling the analysts to clearly set out the risks being run and decision-makers to grant loans within limits that safeguard the bank's long-term financial balance.
- The person or organisation able to effect action: the compliance department or senior management can become involved. Externally, the banking sector's regulatory authorities, or even the criminal justice system, can intervene.
- Its disclosure: the “short text” itself, that Inès de Chambertin sent to her contacts amongst the managers. Apparently, the people who she contacted were not interested in this issue. It appears that, at least initially, she did not want to report the matter to the highest echelons.

**Third case:** she advises her hierarchy of her doubts about her line manager's ethics:

- An illegitimate practice: the pressure exerted by her line manager for her to increase her ratings, contrary to what she believes to be “the ethical definition of her duties” and the real “awareness of risks”, in the shape, according to Inès de Chambertin, of verbal harrying, the withdrawal of benefits, emotional intimidation and physical violence. Inès de Chambertin refuses to consider the interpersonal injustice of this pressure but instead invokes distributive injustice which, according to her, they tried to force upon her.
- The employer's control: it upholds the independence of the financial analyst against any pressure from their line manager by setting out their respective ethical obligations, ideally within its compliance system. This means that it sets the boundaries for legitimate hierarchical pressure.
- The person or organisation able to effect action: if the superiors allow the managers to apply unethical pressure, the compliance department or senior management can get involved. Externally, professional ethics commissions can promote best practices and lawmakers can impose them.
- Its disclosure: the reporting of pressure exerted by the line managers: second and third at interviews, fourth in a letter. Then, externally, the mixed committee of the French Banking Association – Inès de Chambertin's application failed. She referred the case to the criminal courts on the legal grounds of attempted corruption.

### Organisational silence

Inès de Chambertin's biographical account flags up an overdetermined nature. She describes the "solid education" she received. She fitted into the bank's hierarchical culture with its onus on obedience. She acknowledges that she upheld the bank's political correctness. She thinks that her colleagues saw her as "an extremely conscientious mother, a hard worker with irreproachable behaviour". She very probably appears to be emotional but certainly not dangerous. Her personality in no way suggests that she would enter into conflict or report violations. But, on three occasions, she broke with the apparent consensus and went against her managers. She brought her boss back into line when, according to her, he talked about people behind their backs and used her colleagues' assent to make him stop when she was present. She waged a "little battle by email" to make her managers consider the reckless risk-taking that was fostered, in her opinion, by the "reversal of the burden of proof". Basically, Inès de Chambertin's action did not call into question what can be referred to, according to her, as the new consensus regarding risk-taking. However, her resistance to the pressure applied by her line manager causes a reaction from the "HR minion" who attended the interview with her second line manager: "But Inès, how can you dare to say something like that?". She broke away from the social models and went against the consensus. Inès de Chambertin specifies that "It was very funny as it was representative [of the culture of obedience, against which] great courage was needed".

Management science literature shows that silence may sometimes be the rule in companies when the employees do not talk about problems with their superiors (Morrison and Milliken, 2000). It identifies the various reasons why employees choose not to speak up: passiveness when faced with orders from bosses, fear of displeasing them, wish to act in the group's interests, opportunism or simply working as well as possible according to the corporate culture (Cailleba, 2017). The literature also considers the silence of managers and, in particular, their "moral muteness" (Bird and Waters, 1989), meaning their reluctance to describe their actions in moral terms, even if these actions are spurred by moral reasons: the most well-intentioned managers may just want to preserve the organisation's harmony and avoid complicating their decision-making process. This moral muteness of managers may cause staff to believe that doing business is an immoral activity and can be conducive to organisational silence. Less well-intentioned managers can stealthily shape corporate culture by attitudes, expressed beliefs, language and behavioural patterns to obtain the tacit cooperation of staff for unethical activities (Paine, 1994).

We can conclude, as does Moberly (2006), that organisational silence occurs when the leaders play on the need for acceptance by peers to assert a view of the group's unconditional loyalty. According to Grima and Glaymann (2012), the literature follows the line of Hirschman (1970). It puts forward "loyalty" as allegiance to various social groups to which each

employee belongs, in an interlinked manner, from the work team to the company as a whole. It construes the "voice" as a conflict of allegiance (Schehr, 2008). This means that whistleblowers speak up and break the organisational silence. They step away from the group to which they directly belong to prove their loyalty to the superior group. We can fully understand, to quote Inès de Chambertin, that the witness of a violation needs a certain amount of "courage" to blow the whistle. They must be driven by enough motivation to break the organisational silence and contact the higher regulatory authority.

### The witness's deontic anger

Let's go back to Inès de Chambertin's biographical account. We need to understand from which source of energy she draws her "courage". She considers the practices she reported to enshrine injustices. She does not always use this term but her feelings remain just as strong: belittling by her "so funny" boss, "reversal of the burden of proof" which encouraged the over-rating of risks and reckless risk-taking, pressure from her line manager to give ratings contrary to what she considered to be "the ethical definition of her duties" and the real "awareness of risks". Faced with her superiors who "shout" at her or have "an answer to everything": "no longer an ounce of justice [...] they're exceeding the bounds of injustice". These feelings of injustice gave rise to anger. Let's look again at the quotes. When, according to her, her boss spoke about people behind their backs: "It made my blood boil". In light of what she saw as a "reversal of the burden of proof": "I was shocked and had to react". Faced with the pressure she claims to have suffered from her line manager: "I was really angry".

The literature on organisational behaviour posits that the subject feels anger about the injustice and that it is this anger that gives them the energy to correct it. If the subject suffers the injustice themselves, then we refer to personal anger (Batson *et al.*, 2007). This is not the case with Inès de Chambertin. If the subject reacts out of empathy with the person suffering the injustice, then the literature talks about empathic anger (Hoffman, 1989). This does not apply to her either. If the subject reacts as a mere witness, without empathy for the victim of the injustice, then we are dealing with deontic anger (Folger and Cropanzano, 1998, 2001; Folger *et al.*, 2005).<sup>(3)</sup>

This is the case with our whistleblower. According to her statements, she appears to have very strong moral and legal values which are a source of legitimacy. She demonstrated a strong normative commitment believing that these values should permeate the bank and the entire financial community. Her anger originated

<sup>(3)</sup> Folger *et al.* (2005) state that the term "deontic anger" stems from the Greek root of words referring to duties binding people and representing the basis for their mutual obligations. They specify that the expression does not refer to a particular ethical perspective, such as Kantian deontology. As they look into the anger felt by the witness to an injustice, they place greater stress on proscriptions (avoiding vice) than on prescriptions (pursuing virtue), but this does not mean that they stop addressing the concept of moral duty or moral obligation.



from the sole conviction that justice is a concept that should be respected or restored in the event of a violation.

Lindebaum and Geddes (2016) take a closer look at the issue of the anger felt when faced with the injustice. They do not base their work on either Folger and Cropanzano (1998, 2001) or on Folger *et al.* (2005), but come to the same main conclusion. They refer to moral anger, as being distinct from personal and empathic anger, to describe the feeling that pushes the witness to want to correct the act which, in their opinion, represents an injustice that is prejudicial to a third party. What interests us is that they use whistleblowing as an example of moral anger. For their part, Gundlach *et al.* (2003, 2008) emphasise the anger which motivates the whistleblower when faced with a prolonged organisational violation which is able to be corrected but they do not closely examine the injustice caused by the violation nor highlight the deontic nature of the anger that drives the whistleblower.

Inès de Chambertin's biographical account provides an understanding of one type of whistleblower: driven by deontic anger against the injustice which they believe to have witnessed, they look to have this perceived injustice corrected. This means that the whistleblower does not always act impulsively. They may be undecided for a certain period of time and feel anxious (Park and Lewis, 2018). They are torn between the impetus of the deontic anger that drives them to act and the difficulty in knowing how exactly to act. They need to decide how they will have the violation curtailed. Will they speak up and blow the whistle?

### The whistleblower's decision

Let's go back to our analysis of Inès de Chambertin's biographical account. Driven by deontic anger, she brings her "so funny" boss back into line, sends out her "short text" on the "reversal of the burden of proof" and informs her superiors about her doubts concerning the ethics of her line manager. It is clear in her mind that she must take action. She scarcely wonders how to act; she decides as she is acting.

When faced with what she considers as being belittling talk from her boss, she realises that no one is brave enough to contradict him despite the fact that, deep down, no one really approves of it. So, she spoke up: "I said: '[...] it's against my ethics, you can't talk like that in front of me'". She relied on the assent of her colleagues. She spoke to them indirectly and asked them to endorse her point of view. She spoke directly to the perpetrator of the violation deeming that he was best placed to modify his behaviour. She was successful, at least on the face of it as, according to her, he was ironical about her prudishness, in her presence, and very probably, when she was not there, continued to talk about people behind their backs. She was not taken in by his double standards. There may be some doubt surrounding the effectiveness of her speaking out. Nevertheless, under the impetus of deontic anger, Inès de Chambertin was convinced of the importance of her action.

Faced with "reckless risk-taking" that, according to Inès de Chambertin, was permitted by the "reversal of the burden of proof", she was surprised that those of her colleagues, who still followed the bank's line as it was when she arrived, accepted the situation. She says "At that time, I wrote a short text about the problems I had noted". She sent it to her contacts, showed it to a director and conducted a "little battle by email which was not seen from the exterior", to encourage them to assume their responsibilities. She appealed to her contacts' professional conscience but nothing changed. She appeared unaffected by this; she did what she was responsible for. Here again, we can wonder as to whether her speaking out was effective. As she was driven by deontic anger, she still considers that it was justified.

Faced with what she considered to be pressure from her line manager to increase her ratings, Inès de Chambertin focused on the substantive issue, namely the "ethical definition of her duties" and the "awareness of risks" that all analysts should have, rather than on what she saw as the verbal harriving, the withdrawal of benefits, emotional intimidation and even physical violence. She started with passive resistance, leaving him at liberty to revise her ratings himself. She behaved in the same way as before and simply pointed out his responsibilities, once again in vain. However, when he complained about her insubordination, she felt obliged to set out the facts as she saw them. She explained to her second line manager that "He's too business oriented. I don't believe that he has a risk-based mentality". She thought that her superiors would act in good faith and be prepared to reconsider this issue, and even find that her line manager was in the wrong. She did not imagine that her superiors would carry out their threats ("even at the dismissal interview, I still thought that it was intimidation and that they wouldn't dare to fire me"). The facts proved her wrong. It was only later, when she had tried all the internal appeal channels, that she brought the case before the courts. It was only step by step that she appealed at a higher regulatory level. And, each time, she was acting out of a sense of duty, under the impetus of deontic anger.

To sum up, Inès de Chambertin believed that there were violations which were being covered up by organisational silence. She felt deontic anger and spoke up. She initially talked to the perpetrator of what she deemed to be a violation and then to her superiors and so on and so forth. She considered that it was her duty not to let the alleged violation continue.

### Discussion, recommendations and conclusion

Let's summarise what we have learned from Inès de Chambertin's biographical account. We were looking to better understand the decision to speak out taken by whistleblowers who are sensitive to values and who have a social conscience. We chose her case as she appears to be representative of this type of whistleblower. The examination of her biographical account suggests that they perceive the act of whistleblowing as a matter of

justice. Having witnessed what they consider to be an injustice, they feel deontic anger and disclose what they deem to be a violation to the in-house persons who can put a stop to it, or then to external bodies in the event of organisational silence.

### The role of deontic anger in the disclosure decision

We still have to examine the question of weighing up the advantages and disadvantages. Nothing in Inès de Chambertin's account suggests that she did this or even that she was aware of this factor. Of all the literature, only Henik (2015) states that whistleblowers do not compare the pros and cons before making the disclosure. She conducts a quantitative analysis of 47 cases and cites Goldberg *et al.* (1999) by using the terms "strategic moral guardian" and "fed-up vigilante" to distinguish two whistleblower profiles. The first behave strategically when it comes to speaking out outside the firm. They weigh things up but this calculation does not relate to whether or not to make a disclosure, as the majority of the literature posits, but to the best way of achieving the result by mitigating reprisals. The second act out of anger at the reprisals and do not weigh things up.

We consider that Inès de Chambertin is more of a "strategic moral guardian" than a "fed-up vigilante", or at least she tries to be. She sought to distance herself from any personal anger so as to avoid it being said that her disclosure was for ends other than remedying the violation that she mentions. It is clear to her that she had to do her utmost to bring an end to the violation despite the reprisals she faced. If she did indeed reflect or deliberate, it was about the best way to blow the whistle and not about whether or not to make the disclosure.

Our examination of Inès de Chambertin's biographical account supplements the quantitative work of Henik (2015) which gauged the extent of the overall anger of whistleblowers and matched it to the fact that they usually follow extra-organisational principles when making an external disclosure. Our work enables us to describe Inès de Chambertin's anger in detail and to classify it as deontic in reference to the concept put forward by Folger and Cropanzano (1998, 2001) and Folger *et al.* (2005). The analysis shows that she felt such anger in respect of successive issues even before an external disclosure. This means that her case, together with the results of Henik (2015), point to a type of whistleblower who does not weigh up advantages and disadvantages but is strongly motivated by deontic anger.

Does this conclusion conflict with the remainder of the literature which asserts that witnesses who are sensitive to values and display prosocial behaviour weigh up the advantages of having the violation stopped against the various disadvantages connected with its disclosure? In itself, stopping the violation was an advantage for Inès de Chambertin due to her prosocial motivation. What is more, it would appear that she underestimated the risk of reprisals. Does this mean that she assessed one or the other and then compared them? If this was

the case, this was not how she explained the situation. Perhaps she weighed things up without being aware that this is what she was doing. The focal point of her case is the deontic anger that drove her when she decided to blow the whistle. We wonder if her deontic anger did not make her, more or less consciously, over-estimate the advantage and under-estimate the drawbacks to such an extent that stopping the violation became, in her opinion, self-evident. We touch upon the matter of the perception and expression of feelings and their inclusion in understanding the events and their analysis, namely emotional intelligence (Mayer and Salovey, 1997). The notion of emotional intelligence could be used in future research to examine the whistleblowing decision and the part played by deontic anger.

### Factoring in employees' deontic anger

Inès de Chambertin's biographical account highlights the importance of acknowledging the deontic nature of the anger that drives the whistleblower in order to, at least, avoid a personal injustice and, at best, collect the information provided by the whistleblowing (Lindebaum and Gabriel, 2016).

To look into this importance, let's describe the potential reaction of a manager when they receive a disclosure. They consider the alert as an explicit questioning of a part of the company and an implicit criticism of their failure to act. Their initial reaction is doubt: why trust the whistleblower rather than teams that have proved their worth? (Miceli *et al.*, 2009). Their second reaction is fear of the extent of the whistleblower's anger or even their vengeance (Geddes and Stickney, 2011). They may try to silence the whistleblower or to have them dismissed. In this case, the manager does not see the deontic nature of the anger that can drive whistleblowers whose nature is characterised by sensitivity to values and prosocial motivation. They ignore the information they provide on the seriousness of a potential violation and lose the opportunity of making their company more compliant with the expectations of its stakeholders. If they carry out reprisals, they add personal injustice to the deontic injustice felt by the witness and personal anger to their deontic anger, and provoke the situation they were worried about. This is what Inès de Chambertin felt, especially at the time of her third disclosure. According to her, her superiors ignored the importance of the facts that she brought to their attention. And, as she persisted, she was dismissed for, again according to her, "having made serious accusations against her superiors". On the basis of her biographical account, we can infer that they did not understand the rationale for her approach, that they failed to see that she was driven by deontic anger and that they had provoked her personal anger.

Although Inès de Chambertin's disclosures are clearly whistleblowing as defined by Near and Miceli (1985), they are not ethics alerts within the meaning of French regulations at the time which were highly restrictive for reasons dating back to dramatic events in France's history (de Bry, 2008). The examination of the biographical account does not reveal whether Inès de Chambertin's

superiors acted in good faith. If they did, they could have viewed her disclosures as deviance (Babeau and Chanlat, 2008, 2011) from the counterparty risk analysis practices that they considered to be set in stone: she was objecting to business practices which they felt to be normal. There was a conflict as the parties' explanations and justifications were unable to convince them mutually (Chateaufreynaud, 1991). According to Inès de Chambertin, her bosses attempted to force her hand. She says that she resisted but did not have sufficient authority to succeed as she lacked the required support and allies (Boltanski *et al.*, 1984).

The question remains as to whether Inès de Chambertin can be considered to be a moral entrepreneur or at least a rule promoter. On the one hand, she "call(s) the public's attention to these matters [which were harmful to the group in question], [attempts to] supply the push necessary to get things done, and [to] direct such energies as are aroused in the proper direction" (Becker, 2020, p. 162). However, on the other hand, she sought, in her opinion, to restore the entire rule rather than change it. If we look at the interlinking of social groups, the "voice", as we have seen, can be viewed as a conflict of allegiance. The inner group no longer recognises the former rule and may consider the whistleblower as a moral entrepreneur. The outer group may believe that whistleblowers are not rule promoters but that they supplement and pave the way for the action of "professional [rule] enforcers" (Becker, 2020, p. 163) to restore the rule within the inner group. Both deviance and loyalty are relative to the group in question.

## Conclusion

If a business is looking to consolidate its "license to operate", it must take account of disclosures by its employees of actions that its stakeholders would deem to be illegal, immoral or illegitimate. It must pay attention to the alerts, especially when they are driven by anger, even if this anger can be frightening or appear inappropriate. Anger can, in fact, point to an injustice of personal or deontic origin which the company has every interest in dealing with. It should be particularly attentive to the whistleblower's deontic anger spurred essentially by prosocial motivation.

## Bibliography

BABEAU O. & CHANLAT J.-F. (2008), "La transgression, une dimension oubliée du management", *Revue française de gestion*, 34(183), pp. 201-219.

BABEAU O. & CHANLAT J.-F. (2011), "Déviance ordinaire, innovation et gestion. L'apport de Norbert Alter", *Revue française de gestion*, 37(210), pp. 33-50.

BATSON A. P., KENNEDY C. L., NORD L. A., STOCKS E. L., FLEMING D. A., MARZETTE C. M., LISHNER D. A., HAYES R. E., KOLCHINSKY L. M. & ZERGER T. (2007), "Anger at unfairness: Is it moral outrage?", *European Journal of Social Psychology*, 37(6), pp. 1272-1285.

BECKER H. S. (2020), *Outsiders. Études de sociologie de la déviance*, Paris, Métailié.

BERTAUX D. (1997), *Les récits de vie*, Paris, Nathan.

BIRD F. B. & WATERS J. A. (1989), "The moral muteness of managers", *California Management Review*, 32(1), pp. 73-88.

BOLTANSKI L., DARRE Y. & SCHILTZ M.-A. (1984), "La dénonciation", *Actes de la recherche en sciences sociales*, 51, pp. 3-40.

BOYER T. (2013), "Les dispositifs d'alerte dans les entreprises : whistleblowing vs. droit d'alerte", *Management et Avenir*, 4(62), pp. 91-111.

BRY F. de (2008), "Salariés courageux oui, mais héros ou délateurs ? Du whistleblowing à l'alerte éthique", *Revue internationale de psychosociologie*, 34, pp. 135-150.

CAILLEBA P. (2017), "Lanceur d'alerte et silence organisationnel", *Revue internationale de psychosociologie et de gestion des comportements organisationnels*, 56(23), pp. 309-334.

CHARREIRE PETIT S. & CUSIN J. (2013), "Whistleblowing et résilience : Analyse d'une trajectoire individuelle", *M@n@gement*, 16(2), pp. 142-175.

CHARREIRE PETIT S. & SURPLY J. (2008), "Du whistleblowing à l'américaine à l'alerte éthique à la française : enjeux et perspectives pour le gouvernement d'entreprise", *M@n@gement*, 11(2), p. 113-135.

CHATEAUFREYNAUD F. (1991), *La faute professionnelle*, Paris, Métailié.

CRAMER J. (2002), "From financial to sustainable profit", *Corporate Social Responsibility and Environmental Management*, 9(2), pp. 99-106.

FOLGER R. & CROPANZANO R. (1998), *Organizational Justice and Human Resource Management*, Thousand Oaks, Sage.

FOLGER R. & CROPANZANO R. (2001), "Fairness theory: Justice as accountability", in GREENBERG J. & CROPANZANO R. (éd.), *Advances in Organizational Justice*, Palo Alto, Stanford University Press, pp. 1-55.

FOLGER R., CROPANZANO R. & GOLDMAN B. (2005), "What is the relationship between justice and morality?", in GREENBERG J. & COLQUITT J. A. (éd.), *Handbook of Organizational Justice*, Mahwah, Lawrence Erlbaum Associates Publishers, pp. 215-245.

GEDDES D. & STICKNEY L. T. (2011), "The trouble with sanctions: Organizational responses to deviant anger displays at work", *Human Relations*, 64(2), pp. 201-230.

GOLDBERG J. H., LERNER J. S. & TETLOCK P. E. (1999), "Rage and reason: The psychology of the intuitive prosecutor", *European Journal of Social Psychology*, 29(5-6), pp. 781-795.

GRIMA F. & GLAYMANN D. (2012), "Une analyse renouvelée du modèle Exit, Voice, Loyalty, Neglect : apports d'une approche longitudinale et conceptuellement élargie", *M@n@gement*, 15(1), pp. 1-41.

GUNDLACH M. J., MARTINKO M. J. & DOUGLAS S. C. (2003), "The decision to blow the whistle: A social information processing framework", *Academy of Management Review*, 28(1), pp. 107-123.

GUNDLACH M. J., MARTINKO M. J. & DOUGLAS S. C. (2008), "A new approach to examining whistle-blowing: The influence of cognitions and anger", *SAM Advanced Management Journal*, 73(4), pp. 40-50.

HEINEMAN B. W. Jr. (2007), "Avoiding integrity land mines", *Harvard Business Review*, 85(4), pp. 100-108.

HENNEQUIN E. (2012), *La recherche à l'épreuve des terrains sensibles : approches en sciences sociales*, Paris, L'Harmattan.

HENNEQUIN E. (2019), "L'alerte comme source d'une crise biographique : quel processus de bifurcation pour les salariés ayant subi des représailles menant à la perte de leur emploi", communication au 30<sup>e</sup> congrès de l'AGRH "La GRH peut-elle sauver le travail ?", Bordeaux, France, novembre.

HENNEQUIN E. (2020), "What motivates internal whistleblowing? A typology adapted to the French context", *European Management Journal*, 38(5), pp. 804-813.

HENIK E. (2015), "Understanding whistle-blowing: a set-theoretic approach", *Journal of Business Research*, 28(2), pp. 442-450.

HIRSCHMAN A. (1970), *Exit, Voice and Loyalty. Responses to Decline in Firms*, Cambridge, Harvard University Press.

- HOFFMAN M. L. (1989), "Empathic emotions and justice in society", *Social Justice Research*, 3(4), pp. 283-311.
- LATANE B. & DARLEY J. M. (1968), "Group inhibition of bystander intervention", *Journal of Personality and Social Psychology*, 10(3), pp. 215-221.
- LATANE B. & DARLEY J. M. (1970), *The unresponsive bystander: Why doesn't he help?*, New York, Appleton-Century-Crofts.
- LINDEBAUM D. & GABRIEL Y. (2016), "Anger and organization studies: From social disorder to moral order", *Organizational Studies*, 37(7), pp. 903-918.
- LINDEBAUM D. & GEDDES D. (2016), "The place and role of (moral) anger in organizational behavior studies", *Journal of Organizational Behavior*, 37(5), pp. 738-757.
- MAYER J. D. & SALOVEY P. (1997), "What is emotional intelligence?", in SALOVEY P. & SLUYTER D. J. (éd.), *Emotional development and emotional intelligence: Educational implications*, New York, Basic Books, pp. 3-34.
- MICELI M. P. & NEAR J. P. (2002), "What makes whistleblowers effective? Three field studies", *Human Relations*, 55(4), pp. 455-479.
- MICELI M. P., NEAR J. P. & DWORKIN T. M. (2008), *Whistle-blowing in Organizations*, New York, Routledge/Taylor & Francis Group.
- MICELI M. P., NEAR J. P. & DWORKIN T. M. (2009), "A word to the wise: How managers and policy-makers can encourage employees to report wrongdoing", *Journal of Business Ethics*, 86(3), 379-396.
- MICELI M. P., NEAR J. P., REHG M. T. & VAN SCOTTER J. R. (2012), "Predicting employee reactions to perceived organizational wrongdoing: Demoralization, justice, proactive personality and whistle-blowing", *Human Relations*, 65(8), pp. 923-954.
- MICELI M. P., NEAR J. P. & SCHWENK C. R. (1991), "Who blows the whistle and why?", *Industrial and Labor Relations Review*, 45(1), pp. 113-130.
- MOBERLY R. E. (2006), "Sarbane-Oxley's structural model to encourage corporate whistleblowers", *Brigham Young University Law Review*, 5, pp. 1107-1180.
- MORRISON E. W. & MILLIKEN F. T. (2000), "Organizational silence: A barrier to change and development in a pluralistic world", *Academy of Management Review*, 25(4), pp. 706-725.
- NEAR J. P. & MICELI M. P. (1985), "Organizational dissidence: The case of whistleblowing", *Journal of Business Ethics*, 4(1), pp. 1-16.
- NEAR J. P. & MICELI M. P. (1996), "Whistle-blowing: Myth and reality", *Journal of Management*, 22(3), pp. 507-526.
- PAINE L. S. (1994), "Managing for organizational integrity", *Harvard Business Review*, 72(2), pp. 106-117.
- PARK H. & LEWIS D. (2018), "The negative health effects of external whistleblowing: A study of some key factors", *The Social Science Journal*, 55(4), pp. 387-95.
- PETERS K., LUCK L., HUTCHINSON M., WILKES L., ANDREW S. & JACKSON D. (2011), "The emotional sequelae of whistleblowing: Findings from a qualitative study", *Journal of Clinical Nursing*, 20(19-20), pp. 2907-2914.
- PITTROFF E. (2014), "Whistle-blowing systems and legitimacy theory: A study of the motivation to implement whistle-blowing systems in German organizations", *Journal of Business Ethics*, 124(1), pp. 399-412.
- ROTHSCHILD J. & MIETHE T. D. (1999), "Whistle-blower disclosures and management retaliation. The battle to control information about organization corruption", *Work and Occupations*, 26(1), pp. 107-128.
- SCHEETZ A. M. & WALL J. (2019), "Making crime pay: Timing of external whistleblowing", in BAKER C. R. (éd.), *Research on Professional Responsibility and Ethics in Accounting*, 22, Bingley, Emerald Publishing Limited, pp. 1-30.
- SCHEHR S. (2008), "L'alerte comme forme de déviance : les lanceurs d'alerte entre dénonciation et trahison", *Déviance et Société*, 32(2), pp. 149-162.
- STAKE R. (1994), "Case studies", in DENZIN N. K. & LINCOLN Y. S. (éd.), *Handbook of Qualitative Research*, Thousand Oaks, Sage, pp. 236-247.