Three Major Issues Concerning Randomised Social Experimentation in France

L'expérimentation sociale aléatoire en France en trois questions

Bernard Gomel and Évelyne Serverin

Translator: Nicholas Sowels

Electronic version

URL: <http://journals.openedition.org/travailemloi/6844>

DOI: 10.4000/travailemloi.6844

ISSN: 1775-416X

Publisher

DARES - Ministère du Travail

Printed version

Date of publication: 30 December 2015

Number of pages: 85-108

ISSN: 0224-4365

Electronic reference

Bernard Gomel and Évelyne Serverin, « Three Major Issues Concerning Randomised Social Experimentation in France », *Travail et Emploi* [Online], Hors-série | 2015, Online since 30 December 2015, connection on 19 April 2019. URL : <http://journals.openedition.org/travailemloi/6844> ; DOI : 10.4000/travailemloi.6844

Three Major Issues Concerning Randomised Social Experimentation in France*

Bernard Godel**, Évelyne Serverin***

In the 1960s, tools were developed in France to test laws before they were adopted. This form of assessment eventually acquired a constitutional basis with the Act of 28 March 2003 authorising normative experimentation, both nationally and locally. This innovation was the key to a specific form of experimentation, namely the randomised experiment. It borrows its methodology from social sciences, and its perimeter from international policies on poverty. This form of experimentation raises three issues, which are discussed in three successive sections. The first one is a matter of legal science: what role does this form of experimentation play within the scope of normative experiments? The second question is a scientific one: what lessons does social experimentation bring to experimental science, applied to human behaviour? The third issue is socio-political: what is its contribution to the evaluation of public policies? The conclusion recalls the ethical and scientific requirements that are necessary in the conduct and evaluation of experiments on human behaviour.

In a democratic system citizens may demand accountability for the use of public finances via their representatives. This process involves enacting rules, evaluating their scope, and resorting to an argumentative apparatus in which statistical data occupy a central place. Figures are used *ex ante* to provide reasons to legislate, by identifying situations whose evolution calls for changes to applicable rules. Figures are also used as the basis of predictions about the effects of new rules on a particular situation. *Ex post*, statistical evaluations aim to measure the impact of rules, defined according to their use, and to assess their effectiveness, which is judged in terms of the ability of a rule to meet stated objectives. *Ex ante* assessments involve prior studies. These may concern the rationalisation of budget choices (LÉVY-LAMBERT, GUILLAUME, 1971),

* Translation: Nicholas Sowels.

Article published in French in *Travail et Emploi*, n° 135, juillet-septembre 2013.

** Centre d'études de l'emploi (Centre for Employment Studies – CEE) ; bernard.godel@cee-recherche.fr

*** Centre de théorie et analyses du droit (CTAD) ; eserveri@u-paris10.fr

impact studies which have been made mandatory under the French Constitution since 2009, in the passing of parliamentary Bills.¹ *Ex post* evaluations concern the vast field of public policy assessment (VIVERET, 1990), which is officially vested in the French Parliament under Article 24 of the Constitution.² Evaluation techniques have developed above all in the “preparation” of norms and standards, appearing as of the 1960s, with laws and rules that are very explicitly presented as “experimental”. Such evaluation was given a constitutional basis in Law No 2003-276 of 28 March 2003, which concerns the devolved organization of the French Republic, and which authorises normative experimentation both nationally and locally. This innovation opened the door to another type of experimentation, namely random experimentation, which draws on social science methodologies, and which is also applied to international policies on poverty (ALLÈGRE, 2008; GOMEL, SERVERIN, 2009).

It is now a little more than five years since random social experimentation has been used in France, within the perimeter of social policy reform. It is therefore legitimate to review the meaning and scope of this kind of experimentation. To this end, three questions may be raised:

- 1/ In terms of legal science, what is the role of this form of experimentation in the perimeter of normative experimentation?
- 2/ From a scientific point of view, what are the lessons of this experimentation for social experimentation applied to human behaviour?
- 3/ From a socio-political perspective, what has the contribution of experimentation been to the implementation of reforms?

Lastly, this article will examine the ethical and scientific demands which are inherent in all forms of experimentation relating to human behaviour.

The Place of Random Experimentation in Normative Experimentation

If we look at the chronology of experimentation in public policy in France, it can be seen that the Administration (government) preceded researchers in being concerned about anticipating the effects of reforms. Making such practices constitutional has opened the way to normative experimentation at the national and local level. Random social experimentation has emerged locally where social policies actually develop.

1. Article 39 of the French Constitution. The Organic law (or enabling Act) No 2009-403 of 15 April 2009 relating to the application of articles 34-1, 39 and 44 of the Constitution.

2. “Parliament votes the laws. It controls the actions of the Government. It evaluates public policies.”

Experimentation As a Way of Anticipating the Effects of Norms and Standards

Testing laws begins with the idea that political decisions strive for positive effects through reform, but are also concerned about negative effects. From this point of view, administrative experimentation already has a long history.

The term “experimentation” was not used in France’s laws on abortion (*Interruption volontaire de grossesse* – IVG) and Minimum Income Support (*Revenu minimum d’insertion* – RMI), but they were in fact the first examples to use testing.³ The IVG Law on abortion suspended criminal law for a period of five years, while the RMI Law provided minimum income support for three years.⁴ Once these terms ended, the respective governments introduced new Bills to perpetuate the laws.

Testing laws also arose in waivers and derogations to the rules of ordinary law affecting people in the same situation, and which therefore bypassed the principle of equality before the law.

Successive interventions by France’s Constitutional Court on such measures have recognised their experimental nature as a justification for breaching the principle of equality. In a ruling pronounced on 28 July 1993, the Court specified that:

“It is even open to Parliament to provide for the possibility of conducting experiments that involve derogations to the rules defined above, to allow it subsequently to adopt new rules, in the wake of results, which are appropriate to changes in the missions of the establishments in question. However, lawmakers have to define precisely the nature and scope of the experiments, the situations in which they may be conducted, the conditions and procedures according to which they are evaluated that may lead them to be maintained, modified, generalised or dropped.”

A second ruling on 6 November 1996 concerned the information and consultation of employees in certain companies. It reiterated these demands in situations in which social partners obtained derogations to ordinary law on collective bargaining agreements, due to the experimental nature of a measure, provided that the Government reported on experiments to Parliament within five years.⁵

These legislative practices have been supplemented by executive measures, involving the enactment of standards that are limited in time and space.

Thus, the principle of equality before the law was not deemed to have been breached in the case of the progressive application of health controls in food companies, due to a lack of veterinarians.⁶ Similarly, a decree in 1968 creating a new judge only in certain courts or courts of appeal was not considered as violating the principle

3. Law No 75–17 of 17 January 1975 introduced abortion (IVG), and Law No 88–1088 of 1 December 1988 introduced Minimum Income Support (RMI).

4. Until 1979 for the IVG Law (Art. 2), and until 30 June 1992 for the RMI Law (Art. 52).

5. *Conseil constitutionnel* (Constitutionnal Council), decision 6 November 1996, No 96-383 DC.

6. *Conseil d’État* (Council of State), Section, 13 October 1967, No 64778, Rec. CE., p. 365; *Revue de droit public*, 1968, p. 408.

of equality before the law in so far as such limitations were provisional.⁷ Lastly, an opinion given by France's Council of State on 24 June 1993 accepted different tariff measures applied to users of railway lines by the SNCF (the French railway company). Even though the users had similar characteristics, tariff differences were accepted because they were limited to an experiment lasting one year, which was to lead to the definition of criteria used in a new tariff regime.⁸

The validity of legal measures and regulations therefore depends on the existence of an assessment being undertaken by an authority which is enacting standards and the assessment must be followed up by a report.

The Constitutional Acceptance of Normative Experimentation

After 40 years of testing legislation of all types, it was no longer possible to continue with such methods without giving them a constitutional foundation. In its ruling of 17 January 2002, the Constitutional Council considered

“that by opening to Parliament, albeit experimentally and based on a derogation limited in time, the possibility of authorising local government in Corsica to take measures falling within the field of the law, the law in question actually intervenes in areas which belong to the jurisdiction of the Constitution”.⁹

Following this decision, Constitutional Law No 2003-246 of 28 March 2003, relating to the devolved organization of the Republic (known to some as “Act II of France's devolution”: JANIN, 2005), authorised normative experimentation both nationally and locally. On the one hand, it introduced Article 37-1 into the Constitution which states that: “laws and regulations may include measures of an experimental nature which are limited in time”. On the other hand, it added to Article 72 of the Constitution, which relates to local government, a fourth paragraph allowing local authorities the possibility to “derogate from legislative and regulatory measures which govern the exercise of their competencies, on an experimental basis for limited purposes and duration”. Applying this latter measure, the Organic Law No 2003-704 of 1 August 2003 of the General Code of Local Government added a chapter specifically dedicated to experiments, for which the Organic Law provides a general framework.¹⁰

7. *Conseil d'État*, 21 February 1968, No 68615 *et seq.*, *Ordre des avocats près la cour d'appel de Paris*, Rec. CE., p. 123; *Dalloz*, 1968, p. 222.

8. *Conseil d'État, Assemblée générale, section TP*, 24 June 1993, *avis* No 353605.

9. *Conseil constitutionnel*, 17 January 2002, No 2001-454, *statut de la Corse*, note J.-E. Schoettl : *L'actualité juridique : droit administratif*, 2002, p. 100 *et seq.*

10. “On the basis of the fourth paragraph of Article 72 of the Constitution, the law authorises local governments to derogate from the legislation governing the exercise of their powers, on an experimental basis. The law defines the purpose of such experimentation and its duration, which may not exceed five years, and lists the provisions which may be waived. The law also specifies the legal nature and characteristics of the local authorities which are allowed to participate in testing and, if necessary, the cases in which experimentation can be undertaken. It sets the time limit within which local authorities, meeting the conditions it has laid down, may apply to participate in the experiment” (Article LO 1113-1).

Observers have noted that including experimentation in the Constitution has led neither to much controversy nor to debate (STAHL, 2010). This relative silence suggests there is a kind of consensus on the fact that, following 40 years of practice, there is an interest in testing norms and standards before generalising them. Yet by setting out the virtues of “normative prudence”, constitutional texts have now opened the door to experimental methods which have never been discussed before.

The Uses of Constitutional Authorisation

Since the revision of the Constitution came into force in 2003, article 37-1 has been invoked most, mainly in administrative (*i.e.* governmental) reasoning. Social experimentation for its part is based on Article 72.

Following the tradition of texts relating to testing, **experimentation under article 37-1** has been largely geared to **administrative action**, from a technical and budgetary point of view.

The first use of constitutional authorisation occurred with the Law of 13 August 2004 concerning “local liberties and responsibilities” (JANIN, 2005). This Law expanded the field of experimentation with respect to the transfer and delegation of responsibilities and competencies, be they for the management of financing or the exercise of competencies. When it examined an appeal against this text, the Constitutional Council drew on this new foundation to reaffirm the criteria for experimentation which had been forged in previous decisions.¹¹

In another decision of 13 December 2007, the Constitutional Council states that “article 37-1 of the Constitution provides that the Social Security Financing Law includes experimental provisions [...], but that the Council would have to censor provisions that [...] are actually not about experimental application, but applications limited in time”.¹²

This position expresses an essential idea, namely that experimental measures can only function if they are accompanied by an evaluation that looks at all the consequences of the reform.

Other reforms pursuing economic or social or environmental purposes are expensive and so need to go through an experimental “airlock”. There are many examples.¹³ The latest example concerns the development of citizen assessors in criminal

11. *Conseil constitutionnel*, 12 August 2004, No 2004-503 DC: “article 37-1 of the Constitution, which follows the constitutional revision of 28 March 2003 referred to above, allows Parliament to authorise experiments which derogate from the principle of equality before the law, for limited purposes and duration, in view of their possible generalisation. [...] However, lawmakers must define the objectives and the conditions [of such experimentation] with sufficient precision and not disregard other requirements of constitutional value”.

12. *Conseil constitutionnel*, 13 December 2007, No 2007-558 DC – the Law financing public health insurance.

13. Sections 37-1 and 38 of the Constitution were covered by Ordinance No 2006-433 of 13 April 2006, “experimenting the contract on the professional transition contract” for identified employment areas (a measure repealed by Law No 2011-893 of 28 July 2011 for the development of training in sandwich courses and providing career security). Under the same heading, Law No 2008-596 of June 25, 2008 on the modernisation of the labour market,

cases, by Law No 2011-939 of 10 August 2011. This resulted in a (negative) report issued by two judges of France's Supreme court (*Cour de cassation*) (SALVAT, BOCCON-GIBOD, 2013).

In all these areas, experimentation concerns norms and standards themselves, and its inclusion in all norms affecting social, economic, and budgetary policies. The task of experimentation falls on the Administration itself, without recourse to external research. Above all, no scientific hypotheses are tested. Thus, for example, it was not planned in the experiment on citizen assessors to test the "severity" of non-magistrates, even though this had been one of the reasons for the reform. Only the consequences of the measure were to be assessed. In short, experimentation amounts to a sort of "penitence clause" to avoid undertaking future measures that are expensive, controversial, or which present serious risks of being distorted.

Random experimentation does not have its roots in article 37-1, but in **Article 72** of the Constitution. For the first time, the objective of the experiment was not to check the quality of a norm, the feasibility of reform or even its cost. Instead, experimentation was to measure the effects on the behaviour of the persons subject to the experiment.

Social experimentation with random assignment has a long international history. But it began in France with the report of the Families, Vulnerability and Poverty Committee (*Commission Familles, vulnérabilité, pauvreté*, 2005), chaired by Martin Hirsch. It suggested the government use "experimentation programmes that it [the government] would define itself but would like to test on part of the national territory" (GOMEL, SERVERIN, 2009, 2011). Lawmakers took up this advice in two stages. First with Article 142 of the Finance Law for 2007 of 21 December 2006:

"On an experimental basis, to improve the conditions of financial incentives for persons returning to work and to simplify access to assisted employment contracts, the Departments [*i.e.* French counties] mentioned by the decree provided for in Article LO 1113-2 of the General Code on Local Authorities are authorised to adopt all of the exceptions to the provisions of the Labour Code and Code for Social Action and Families in favour of the beneficiaries of minimum income, for a period of three years from the date of publication of the decree [...]"

On 4 May 2007, two Departments (Eure and Côte d'Or), which had previously declared themselves as volunteers on 1 February and 23 March 2007, were authorised by decree to conduct experiments. These related firstly to improving the conditions of financial incentives to return to employment; and secondly to simplifying access to subsidised work contracts. On 20 June 2007, the General Council of Eure adopted a resolution setting out the derogation rules for experimentation in the Department.

introduced a new employment contract related to the defined objective, and tested for a period of five years (Article 6). The Finance Law No 2009-1673 of 30 December 2009 introduced contractual income support for autonomy, on an experimental basis. Decree No 2010-1395 of 12 November 2010 on mediation and legal activities concerning family matters experimented with an obligation to see family mediators. Law No 2011-1862 of 13 December 2011 on the allocation of litigation and the simplification of legal procedures experimented compulsory family mediation in certain disputes concerning children.

Without waiting for the end of this first experiment, the Parliament adopted a substantial measure affecting its design in the summer of 2007, by moving to a second stage of legislation. Articles 18-23 of the Law of 21 August 2007 called “TEPA” (the law in favour of work, employment and real disposable income) opened up the experimentation of an “Active Solidarity Income” (*Revenu de solidarité active* – RSA) to all volunteering Departments. Section 22 of the Law requires the government to file an evaluation report before any generalisation of the RSA. The rapporteur of the Bill generalising the RSA lauded the approach taken by this assessment:

“An evaluation committee was set up under the chairmanship of M. François Bourguignon, established from the beginning of the experimentation, and not a posteriori as is too often the case, and an evaluation protocol was drafted, based on comparisons with control areas with experimental areas”.¹⁴

The model of randomised experimentation is thus covered by the constitutional framework authorising local derogations to rules, for simple methods of evaluation. Its influence has grown with the approach of “social experimentation”, which is presented in France as a major innovation that can resolve complex social issues otherwise endlessly debated inconclusively (COMMISSION FAMILLE, VULNÉRABILITÉ, PAUVRETÉ, 2005). But it is far from having replaced traditional administrative experiments. Only these can anticipate the effects of reforms in all their dimensions, without succumbing to the pressure for change. It is here that the two forms of assessment diverge. Administrative experimentation is motivated by doubts over norms and standards: it seeks to anticipate the negative consequences of reform decisions. Random testing is guided by hopes concerning new norms and standards: in this case proposed reforms are assumed to have positive effects.

The Lessons for Experimental Social Sciences

Economics as a science is familiar with the idea of experimentation. In mainstream neo-classical economics, experimentation is one of the “empirical dimensions” of testing microeconomic models, by simulating agents’ real behaviour in a laboratory environment. This field has grown considerably since its beginnings in the immediate post-war period, leading finally to the award of the 2002 Nobel prize in economics to Vernon L. Smith “for having made laboratory experimentation an instrument of empirical economic analysis” (BIAIS, RULLIÈRE, 2003). However, while this form of experimentation does partly help with decision-making, particularly in terms of economic policy, its primary goal is knowledge: its aim is to test the validity of economic theories, or to explain the observed regularities.

14. Daubresse M.-P. (rapporteur) (2008), « Rapport fait au nom de la commission des affaires culturelles, familiales et sociales sur le projet de loi (n° 1100) généralisant le revenu de solidarité active et réformant les politiques d’insertion ». *Rapport*, No 1113, Paris: Assemblée nationale, p. 35. Online <http://www.assemblee-nationale.fr/13/pdf/rapports/r1113.pdf> (accessed 3 May 2016).

Experimentation takes on another form when it is used not merely to provide “validated” theoretical models for policy-makers, but when it aims to test the effectiveness of a given policy instrument in achieving a specific result. The process then falls within the field of experimental science, and requires the use of a robust method to create test groups. Random assignment, theorised by Ronald FISCHER in 1935, is the best means for conducting this type of research. It has the merit of rendering the various samples thus constituted identical, as “randomisers” have systematically pointed out. Precision only depends on sample size:

“Random draws, of a test group and a control group within a same population, ensure the comparability of the two groups: on average, the population characteristics of each of these groups are not statistically different. The larger the sample, the more these characteristics are likely to be similar. This property applies to the observable characteristics (gender, age, academic achievement, educational level, etc.), but also those which are not: characteristics which are specific to individuals and difficult to measure, such as motivation, self-confidence, etc.”¹⁵

Strict adherence to random assignment is a safeguard against selection biases that disrupt the value and precision of estimates made from the data of an experiment. The areas studied by FISCHER concerned experimental science: the random assignment here is a matter of procedure, of method, as is the quality of the collection of experimental data. When an experiment is conducted according to the rules of the art, the difference in values for the criterion tested between the treated and untreated groups is indeed due to the specific effects of the criterion itself.

It is tempting to consider various social policies as “treatments” applied to populations whose behaviour policy-makers want to change. Compared to other methods for proving the effectiveness of reform, “randomisers” hold out the promise of greater scientific certainty regarding the effects of a measure on specific behaviours. However, once the experiments are conducted in the real world and not in the laboratory, the evidence is more difficult to establish. As we shall see, several examples of experiments conducted in France in recent years show that their proponents have not attached due attention to the consequences of their immersion in the real world concerning the robustness of the results. In the case of the RSA, it seems that initially assessors of the experiment were only observers who had no control over the conduct of operations at any time. In the experiment on the placement of unemployed persons, the assessors were not able to control random assignment in the operation. Finally, in the testing of the “Parents’ schoolbag” (*la Mallette des parents*) project, which aims to provide parents with better information about how their children’s schooling and education operate, it was not possible to isolate statistically the specific effect of the measure.

15. Ministère des Sports, de la Jeunesse, de l’Éducation populaire et de la Vie associative, Fonds d’expérimentation pour la jeunesse (2012), *Rapport d’activité 2009-2011*, p. 35. Online http://www.experimentation.jeunes.gouv.fr/IMG/pdf/FEJ_RA_20092011_CorpsRapport.pdf (accessed 3 May 2016). The size of the sample depends on other factors such as the importance of the effect, see below.

The Difficulties of Carrying Out Experiments

The results of the experiments concerning RSA income support conducted in 33 Departments on incentives to return to work generated controversy in parliamentary discussions in the autumn 2008 (GOMEL, SERVERIN, 2009). These results were also discussed by some economists (CAHUC, ZYLBERBERG, 2009). Despite the very short duration of the experiment (five months), experimenters provided discussions with a very weighty “finding”, namely that the rate of return to employment had increased by 30%. A few months later, when results were available for a slightly longer period, the effect of the RSA turned out to be three times lower. Whatever the actual level, this percentage identified did not have the meaning which lawmakers attributed to it. But it only had a purely statistical sense, summed up by the Scientific Evaluation Committee:

“The rate of entry into employment of recipients of the RMI in the experimental areas is on average higher than in the control areas, but the gap varies quite widely between Departments and over time.”

(COMITÉ D’ÉVALUATION DES EXPÉRIMENTATIONS, 2009, p. 12.)

Nevertheless, the difference between the experimental and control areas, which was 30% on average in the first five months, was no more than 9% after 15 months, at the usual 5% threshold of significance (COMITÉ D’ÉVALUATION DES EXPÉRIMENTATIONS, 2009, p. 13):

“[...] The average monthly rate of job entry in control areas is 3.1%. The difference between experimental and control areas zones is 0.28 percentage points, or an extra 9% of job entry rates in areas experimenting the RSA. [...] The probability of being wrong in asserting that the effect of the experimental RSA on the return to work is greater than zero is 12%. This is not a high value, but nevertheless leaves limited room for uncertainty.”

This was not the only weakness of the experiment. When the policy came into force, a massive phenomenon of non-take-up of its new component –the *RSA-activité*, aimed at the working poor– cast doubt on the relevance of a model geared only to financial incentives (SERVERIN, 2012).

Apart from discussion of the results, two methodological issues raise questions.

First, the assessors were not able to determine the method of fixing the test and control areas, which is particularly disturbing given the importance the experimenters give to random assignment. In fact the law gave local authorities the freedom to choose the concrete modalities of the experimentation. In particular, the Regional Councils selected the zones for experimentation. These were often the most disadvantaged socially, while Councils showed no concern for finding a control zone for comparisons. The team responsible for selecting the control areas was not able to match each zone tested with a control area. It was not therefore possible to compare the rate of return to employment for each Department conducting the RSA experiment with the effects of the RMI, which would have provided information about variations in the RSA experiments. As the treated and control populations were comparable on a general level, the

only calculable outcomes about the return to work related to the average difference between the entire population treated and entire control population.

Second, qualitative research monitoring the experimentation has shown that information about the persons concerned in the areas experimenting the RSA was not better controlled for either. Yet it was precisely here the incentive effects are to be found, which the proponents of experimentation were looking for. A monographic study conducted during the experiment in five Departments (LONCLE *et al.*, 2008) showed that communication with target audiences was “prudent”, not to say opaque, whereas persons concerned should have had access to complete information. The objectives themselves were not limited to the resumption of activity, but could have been broader (changing the action of administrators accompanying the policy, ensuring predictability of income, etc.). They would then have shown greater continuity with the reorganizations already occurring as part of the devolution processes of the RMI.¹⁶

In reality, the conditions of the experiment were only controlled at the margin. This should have prompted caution on behalf of the assessors, and at the very least, it should have led them to refuse to provide any results whatsoever concerning the “incentive” nature of the measure.

The Difficulties of Controlling Objectives Through Random Assignment

A large randomised experiment was conducted in 2007 by France’s national unemployment insurance fund and the public network of jobcentres (respectively Unédic and ANPE) on enhanced individual support and accompaniment for jobseekers. This experiment testified to the changes concerning the role of random chance in experiments, the reasons for such shifts and their consequences for results.

The evaluation of the effectiveness of two procedures for the enhanced accompaniment of jobseekers was a first in France, in terms of the exceptional means mobilised (the evaluation covered altogether more than 200,000 jobseekers) and its methodological quality. On the basis of a random draw, three groups were created:

“those who were offered POP (Private Operator Placement) accompaniment; persons offered ‘company destination’ or CVE accompaniment (*Cap vers l’entreprise*), and persons provided with the traditional enhanced accompaniment provided by French jobcentres”.

(ANPE, DARES, Unédic, 2008, p. 4)

The authors stressed that the random draw ensured the comparability of the three groups (p. 4):

16. Another survey conducted by the Crédoc (Research Centre on the Study and Monitoring of Living Conditions) of beneficiaries in the same five Departments shows that claimants are very unlikely to position themselves in the simple dichotomy between work or income support given in the models. The concern here is one of regularity and certainty in receiving regular benefits as a complement to wages. The need for security and regularity is shared by all households regardless of income levels.

“The same proportion of young people, women, etc. (this can be checked) but also the same proportion of motivated, dynamic jobseekers, etc. (even if we cannot check this, as it is not really possible to measure motivation or energy, it is sure to be so if the groups are large enough and are drawn randomly). Therefore, if we observe differences in jobseekers return to work between these three groups, we know that these differences are due to the fact that these persons were not offered the same type of support, and that only.”

Despite this presentation, everyone remembers that the experiment pointed to the superiority of the public operator (the CVE programme) over the private operators, while the experiment occurred at a time when the ANPE and the Unédic were competing with private operators to obtain the market for jobseeker placement.¹⁷ Several factors explain this misunderstanding, though a direct comparison between the CVE and the POP was not on the agenda of the experiment.¹⁸ It seems there were even hesitations concerning the objectives of the study within the final report itself (BEHAGHEL *et al.*, 2009).

On the one hand, the report recalls the initial objective (p. IV):

“This protocol, of ‘controlled experimentation’, guarantees that the different situations in the labour market that can be observed between the groups after several months can only result from the benefits of the programmes. Thus the value-added of CVE and POP programmes is measured with precision and certainty”.

Rates for exiting unemployment at three, six and twelve months of compensated flows without enhanced accompaniment are given: 12% at three months, 23% at six months, and 37% at twelve months. “After 12 months, CVE support raises the exit rate from unemployment from 37% to 44%” and “the effects of POP [...] are generally weaker and later” (p. V).

On the other hand, the report emphasizes the importance of direct comparison between the two operators. For this, observation was restricted to common geographic areas to ensure that the “economic conditions and target populations were then identical” (p. V).

“In these areas, the effects of both programmes are stronger, but also more mixed: CVE increased the exit rates by 11 points from the third month, while over this horizon the effect of POP was still not significant. At 12 months, CVE had an impact of 8.5 points and POPs had an impact of 6.4 points” (p. V).

From one part of the report to another, the objective of the assessment changed, from examining the behaviour of jobseekers to looking at the operators. This shift was

17. Before the merger of the ANPE-Assedic (the payment centres of the Unédic unemployment insurance fund) disrupted the distribution of functions that was being put into place.

18. It is only really possible to compare POP support with standard jobseeker support on the one hand, and CVE support with standard jobseeker support on the other hand. By contrast, it is not possible to compare directly the results of POP support and CVE support, if the strengths and securities stemming from random selection were to be retained. This comparison could only be carried out indirectly: so in fact, two parallel comparisons were conducted to compare the respective performances of CVE and POP support relative to standard jobseeker support.

due to the conditions in which the experiment was conducted, and more specifically due to the direct intervention of project operators in the final assignment of jobseekers to the groups tested. Indeed, according to the study, only 47% of jobseekers randomly assigned to a POP were actually supported by the POP, with the support rate of CVE projects being 43%. At the same time, only a small proportion of non-entrants into projects –less than 20%– reflected the refusal of jobseekers themselves from participating in a project. In most cases, jobseekers were rejected from participating after selection was directly organized by operators, on a case-by-case basis. The latter tried to keep only those jobseekers most likely, in their view, to benefit from the enhanced accompaniment. The greater selectivity of the CVE scheme could be explained by a better knowledge of jobseekers and so explain the greater success of enhanced accompaniment. This issue is discussed at the very end of the summary of the final evaluation report. It is pointed out as a new item to be addressed in future analyses. But it actually revives questions about the ability of random assignment, which was the main innovation of the method, to help identify the direct effects of the programmes in explaining differences in the exit rates from unemployment into getting a job. As the authors of the report themselves finally state,

“it is plausible that differences in outcomes [of the POPs] compared to the CVE programmes may be interpreted by the incentives given to the different actors involved” (p. VIII).

Despite the very significant adjustments to the random assignment in the implementation of the experiment, the final report continued to focus, in its presentation of the experimental protocol, on the importance of (pp. 2-3):

“The value-added of a programme is defined as the difference in the situation of individuals benefiting from the programme compared to what would have been their situation if they had not benefited from it [...]. To re-create this hypothetical situation (*i.e.* the counter-factual), assessment generally refers to a control group. [...] Constructing an appropriate control group is difficult. [...] It is nevertheless possible to ensure, in the construction, that we can obtain a group of beneficiaries and a control group which are thoroughly comparable. This can be done using random selection from the same population. The differences in paths observed following the programme’s implementation can then be attributed transparently and robustly to the programme, and to it alone. This was the principle of the experimental protocol applied to the evaluation of the POP and CVE programmes...”

The final presentation, once the experiment was conducted, used exactly the same words to publicise the assessment (ANPE, DARES, Unédic, 2008), despite the “uncertainties” faced during the implementation.

The assessors neglected the impact the selection of operators had on the scope of the enhanced accompaniment, and were only concerned about the voluntary participation (compliance) of jobseekers:

“The assessment only provides information on the average impact of enhanced support when participation is voluntary, and only if 50% of persons to whom the programme was offered actually participated: the average effect on all jobseekers would potentially be different if participation were compulsory and if 100% of jobseekers thus participated” (p. 14).

What is ultimately calculated is the average impact of support on persons accompanied (PARIENTÉ, 2008). The conditions of the effective entry into the support programmes of the experiment were taken as given.¹⁹ Yet these conditions modified the very essence of the “support effect”.

Isolating the Specific Effect of the Variable Being Studied

The “Parents’ schoolbag” project implemented in 2008-2009, within the Youth Experimentation Fund (*Fonds d’expérimentation pour la jeunesse* – FEJ), held out the same prospects of improving problematic social situations (AVVISATI *et al.*, 2011). It was evaluated by a team from the Paris School of Economics. Funding for this experiment was justified by the fact that the involvement of parents in middle school and its effects on pupils’ behaviour is a crucial question which has long been debated. And “yet nothing was ever really tried to shed light on the question rigorously” (p. 5). For the first time in France, an experimental trial in this area would aid in “judging the effectiveness of a policy” and “the appropriateness of generalising” the policy (p. 5). This would be achieved thanks to an experimental protocol aiming at “a rigorous and transparent assessment” (p. 5) of the programme. According to the investigators, “as a random draw ensures that there is no systematic difference [...], the differences that can be observed [...] can be definitely attributed to a single cause: the benefit of the programme” (p. 5).

The trial took place in 40 middle schools belonging to the regional education Academy of Créteil (to the south-east of central Paris) (pp. 3-4).

“Earlier this year, about a hundred classes were randomly selected. Their parents were then invited by the principal to attend a series of briefings on the functioning of the middle school and to discuss how best to help children and to interact with teachers. At the end of the school year, when comparing these parents to other classes in the first year of middle school (initially similar but not randomly selected), they were characterised by being significantly more involved in their children’s schooling. [...] In particular, the data show that this increased involvement resulted in a significant improvement of children’s behaviour.”

19. The acceptance of jobseekers randomly assigned to reinforced accompaniment under CVE or POP is only a problem of the generalisation of the experiment’s results. By contrast, the further selection by the operators raises another problem: the observed effects will mix two causes: i) the quality of their enhanced support, and ii) the ability of operators to choose candidates who are best able to take advantage of the strengthening of services to help with access to employment among applicants and randomly assigned volunteers. To achieve this second selection, the CVE programmes certainly had better expertise in using the ANPE data made available to all operators.

This improvement was measured using a series of indicators of student behaviour relating to absenteeism, sanctions, “distinctions” and to school life scores.

“Whether it is absenteeism or a composite ‘quality of behaviour’ score summarising three other variables, the test groups’ advantage over the control groups is about 10% of a standard deviation. Given their magnitude, these differences are clearly greater than chance effects in the investigation” (p. 15).

Moreover, the experiments have shown that “the virtuous effect of these openness policies towards parents extends beyond the usual circle of those families who are most attentive to the tuition of their children, and affects families and pupils who are most detached from the school system (by ripple effects in classrooms)” (p. 3).

But how is this improvement to be explained? For the experimenters, there was no doubt that this was indeed a result of the parental involvement programme. However, it is far from certain that the participating population was actually more involved in middle-school life. The results of the experiment show an equivalent rate of non-responses by the parents of test groups and control groups in the final assessment, whereas one would have expected greater participation of the former. This is a weakness in the methodology which was recognised by the experimenters, but which they suggested is limited in scope:

“If the response rates were very different [...], then comparing the responses obtained in test classes and control classes could no longer be interpreted as reflecting the sole effect of eligibility for the experimental programme” (p. 10).

But even if non-significant, the fact that the rate of parent responses in test classes is still slightly lower than for the control classes is troublesome, for an experiment that specifically addresses parental involvement and its effect on children’s behaviour.

As we can see, it is not so easy in a “natural” experimental framework to ensure the validity of results, from an internal point of view (in terms of the variables used). This is even truer from an external point of view, because of the impossibility of being certain that other causal variables do not intervene. Yet it is the external validity that has to do with generalisability: “are these results valid also for the broader population for which the policy or treatment is being considered?” (RODRIK, 2008, p. 16).

The scrapping of an experiment is not easy even if Article LO 1113-6 of Local Government Code provides for this. The assessors of another experimental programme of the Academy of Créteil to tackle truancy did indeed manifest such reticence. This programme made the headlines and received a negative review in April 2010, at the end of its pilot scheme, immediately followed by justifications (BEHAGHEL, GURGAND, 2010, p. 16). This review “does not mean the measures had no effects –only an impact assessment could demonstrate this– but that the conditions of assessment were not met”. Accordingly, the conditions of the scheme’s social acceptance mentioned in the review in fact constituted two additional limits relating to the legitimacy of an experimentation in social matters: first, issues to be evaluated should be consensual (though randomised experimentation is presented in canonical way as determining issues that

divide public opinion when there is a lack of evidence); and second, projects must be defined without prior knowledge that they will be funded as part of an evaluation. This seems to contradict the whole principle for responding to the call for projects by the Youth Experimentation Fund. The assessors added another type of condition (p. 2):

“from a scientific point of view, it is clear that the conditions for an objective assessment –that actors do not change their behaviour in view of assessment– were not met during the pilot project. There is no reason to think that these conditions were met in 2010-2011”.

Knowing when and how to give up a project is indeed a challenge facing social experimenters who are officially involved in reforms.

The Contributions to Reform

One of the remarkable features of randomised social experiments is their vocation to be generalised. A guide on the methodology of the evaluation of social experiments (*Le Guide méthodologique pour l'évaluation des expérimentations sociales*)²⁰, published for project leaders in 2009, in fact set this as the primary purpose of projects (p. 2):

“This is a social policy innovation which is being initially launched on a small scale, given uncertainties about its effects and implementation under conditions that allow evaluation of the results, with a view to future generalisation, if the results are convincing”.

In its first evaluation report, the Scientific Council of the experimentation fund argued that the assessment of its action “would be judged in particular by its ability to inspire public decisions” (MINISTÈRE DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DE LA VIE ASSOCIATIVE, 2011, p. 6).

If we take the promises of social experimenters seriously, then we should soon see reforms emerge that are entirely constructed on the basis of experimental results. It is precisely this criterion of operationality that has dominated financial investment in such experiments for the last five years. It seems that today there is a growing awareness concerning the realism of these goals, and that there is a return to a more traditional approach to evaluation.

Random Experiments in Support of Decision-Making

Soon after having been promoted by the Hirsch Commission, random experiments found their place in the budgets of central government, in the form of two

20. *Guide méthodologique pour l'évaluation des expérimentations sociales à l'intention des porteurs de projets*. Online <http://www.experimentation.jeunes.gouv.fr/IMG/pdf/guide-pour-l-evaluation-des-experimentations.pdf> (accessed 3 May 2016).

dedicated funds: the Innovation and Social Experimentation Fund (*Fonds d'innovation et d'expérimentation sociale* – FIES), established in 2006; and the fund supporting experimentation for young people, better known under the name of the Youth Experimentation Fund (*Fonds d'expérimentation pour la jeunesse* – FEJ), created in 2008.

Since the 2006 Finance Law, the FIES has provided support for the development of experiments in anticipation and in support of public policies favouring solidarity and social cohesion. Experimentation is included in the Mission: “Solidarity, integration and equal opportunities”, in the Programme 304 of “Fighting poverty: active solidarity income and social experiments”, under Action No 2 called “Social experiments and other experiments in social and social economy matters”.²¹ This action is presented as an embodiment of

“Resolution No 13 of the *Families, vulnerability, poverty* report of 2005, which aimed to make ‘boldness, innovation and experimentation’ keywords in public actions, based on the fact that action required for vulnerable families had to be customised and innovative” (p. 31).

Scientific backing is widely put to use: “selected by a panel including qualified personalities, experiments supported by the innovation fund and social experimentation are to be encouraged through calls for structured projects on topics listed above”.²²

As for the fund supporting experimentation for young people,²³ it places experiments in the longer time horizon of research:

“The fund supporting experimentation for young people, established under Article 25 of the Law of 1 December 2008 referred to above, aims to fund experimental programmes geared to promoting pupils’ success in schools, to contributing to equal opportunities and to improving the sustainable social and professional integration of young people under 25 years old. The fund may therefore finance spin-off experiments in new territories [...]”.²⁴

To its promoters, the fund’s objective is not to finance projects only for their own sake, but “to learn, in order to capitalise mobilised knowledge for the design of future public policies” (MINISTÈRE DE L’ÉDUCATION NATIONALE, DE LA JEUNESSE ET DE LA VIE ASSOCIATIVE, 2011, p. 7). Budgets and actions have followed. Since its creation, the

21. The Interministerial Mission on Annual Projects for Performance (2012) (« Programme 304. Lutte contre la pauvreté : revenu de solidarité active et expérimentations sociales », In *Mission interministérielle. Projets annuels de performance. Annexe au projet de loi de finances pour 2012. Solidarité, insertion et égalité des chances*, pp. 19-43.) Online http://www.performance-publique.budget.gouv.fr/sites/performance_publique/files/farandole/ressources/2012/pap/pdf/PAP2012_BG_Solidarite_insertion_egalite_des_chances.pdf (accessed 3 May 2016).

22. A document for cross-cutting policies. The 2011 Finance Bill (2011) (*Document de politique transversale. Projet de loi de finances pour 2011. Inclusion sociale*, p. 17.) Online http://www.performance-publique.budget.gouv.fr/sites/performance_publique/files/farandole/ressources/2011/pap/pdf/dpt/DPT2011_inclusion_sociale.pdf (accessed 3 May 2016).

23. Created by Article 25 of the Law of 1 December 2008, modified by the Law No 2010-1658 of the 29 December 2010, Article 21.

24. Decree No 2011-1603 of 21 November 2011.

Youth Experimentation Fund has launched 11 calls for projects, received over 1,500 experimental project proposals, and has brought together over 30 expert panels. More than 380 experiments (lasting an average of three years) have been selected throughout the country (mainland France and overseas territories) addressing extremely varied themes.²⁵

While the activities of the FIES have not been accompanied by dedicated indicators, the actions of the EFJ have had two. These were set out in the 2012 Finance Bill, in Programme 163 for “Youth and associative life” of the annual performance project of the “Sport, youth and community life” Mission. The former measures the share of projects which actually began six months after their selection. The latter indicator identifies “controlled experiments” within all experiments. This indicator is accompanied by a commentary which restates the principles and precepts of the methodological guide (p. 107):

“[The indicator is] intended to support the development, via the EFJ, of a so-called ‘controlled’ experimental approach, based on the observation of a test group benefitting from a particular policy/measure and a control group that does not benefit from it. These groups are constructed by selecting people using random draws. This approach is based on quantitative assessment, and has been practiced especially in Anglo-Saxon countries for decades. It still remains to be developed widely in France, especially in the field of public policy. Based on representative samples (from hundreds of individuals to thousands, or more), the approach provides strong demonstrative scope to the policies/measures assessed.”

The emphasis given to methodology reached its high point here, virtually holding out the possibilities of an official science in the service of political action. This is clearly going a bit too far. A significant change of tone can be seen in the PAPs (*Projets annuels de performance*) of the 2013 Finance Bill, which leans towards an approach that is not so tied to scientific procedure.

Experiments in Support of Analysis

The change in presentation is particularly sensitive for the experiments included in the Programme 304 of “Fighting poverty: active solidarity income and social experiments”. From 2013 onwards, Action 12 for a “social and solidarity-based economy” of the programme has been dedicated exclusively to actions relating to the support and development of a social and solidarity-based economy. The actions relating to other experiments in social matters are included in Action 13 for “Other experimentations”, which has changed significantly in volume. The actions of Programme 304 are thus losing experimental funding: of the €5,981,487 attributed to the former action for “Social experimentation and other experiences in social areas and social economy” in 2012, €5 million has now been transferred to Action 12 for a “Social and

25. Projects available online <http://www.experimentation.jeunes.gouv.fr/>.

solidarity-based economy”. The remaining €981,947 is now only linked to experiments in a limited way:

“The 2013 budget of €981,487, for which committed authorisations = payment appropriations,²⁶ will support the development of experiments in anticipation and in support of public policies in favour of solidarity and social cohesion.

The funds of Action 13 will allow support to be obtained for the operation of the New Agency for Active Solidarity (*Agence nouvelle des solidarités actives* – ANSA). This is part of a multi-year convention currently operating, with the objective of contributing to the development of experiments, the pooling of good practices across geographical areas (particularly in terms of access to minimum benefits by claimants) and testing and evaluating of innovative projects in the field to tackle poverty, projects which are focused on preventing the breakdown of social bonds.

They must also enable the development of social engineering approaches, as part of experimental programmes to test the relevance, effectiveness, coherence and efficiency of public policies supporting social innovation on a limited scale (in four of France’s regions). The programme will aim to strengthen the tracking and support capacities of decentralised networks of social cohesion with respect to initiatives aimed at strengthening social bonds in geographical areas, and to create momentum in the field of social innovation.”²⁷

The aim of supporting local projects henceforth could not be stated more clearly, and this has more to do with administrative experimentation than science-based experimentation.

In the PAP of the Sport, Youth and Associations Mission for 2013, the tone also changed concerning the presentation of the objectives of experimentation.²⁸ To be sure, the goal was “to support and evaluate, in a specified manner, innovative measures that contribute to youth empowerment, in the context of implementing new public policies for young people” (p. 89). But the methodology is only mentioned in passing, recalling that “the external and scientific evaluation of these projects, if possible controlled, is an integral part of the selection conditions of funded projects” (p. 89). The focus is on the dissemination of the work, as a way of pushing experimenters to take on the function of disseminating information. A new indicator (indicator 4.1) is produced, whose purpose is to measure the dissemination of the results of supported experiments:

26. Commitment authorisations (*Autorisations d’engagement* – AE) are the upper limit of expenditure which may be incurred. Payment appropriations (*Crédits de paiement* – CP) are the upper limit of expenditure which can be scheduled or paid during the year to cover commitments entered into within the framework of commitment authorisations.

27. The Interministerial Mission on Annual Projects for Performance (2013) (*Mission interministérielle. Projets annuels de performance. Annexe au projet de loi de finances pour 2013. Solidarité, insertion et égalité des chances*, p. 35.) Online http://www.performance-publique.budget.gouv.fr/sites/performance_publique/files/farandole/ressources/2013/pap/pdf/PAP2013_BG_Solidarite_insertion_egalite_des_chances.pdf (accessed 3 May 2016).

28. The Interministerial Mission on Annual Projects for Performance (2012) (« Programme 163. Jeunesse et vie associative », In *Mission interministérielle. Projets annuels de performance. Annexe au projet de loi de finances pour 2012. Sport, jeunesse et vie associative*, pp. 81-115.) Online http://www.performance-publique.budget.gouv.fr/sites/performance_publique/files/farandole/ressources/2012/pap/pdf/PAP2012_BG_Sport_jeunesse_vie_associative.pdf (accessed 3 May 2016).

“to provide useful food for thought for policy-makers as part of the development of youth policies, the results of experiments supported under the FEJ must be available and accessible. Their posting on the government website (www.jeunes.gouv.fr) is therefore important” (p. 89).

The indicator includes two measures:

“The share of the experiments which were the subject of a processed and published evaluation report / The total number of funded experiments; The share of the final evaluation reports that are processed and published in the year / total number of reports expected in the year” (p. 89).

In the new terminology, the aim is now to provide “food for thought”, and not to generate strong demonstrative reach. In this context, simple random experiments become a mere reservoir of ideas, and are no longer oriented towards specific reform projects.

Finally, new experiments are emerging which are more administrative in their approach to testing standards. The Solidarity, Inclusion and Equal Opportunities Mission includes Programme 137 for “Equality between women and men”. In 2013, Action No 14 was introduced to “support actions for experimenting measures in favour of equality between women and men” (p. 126).²⁹ With a budget of €6.3 million, this action involved the creation of a budgetary fund, on 1 January 2013, for “experimentation in favour of women’s rights and equality between women and men”. Its objective is to “implement support programmes and experiments and lay the foundations of new practices promoting professional equality and the effective protection of women against violence” (p. 111 and 126).

These experiments are defined by their objectives, without specifying any method.³⁰ There is no longer any question of using randomisation, only support for testing measures in the true administrative tradition.

This change could augur a redistribution of responsibilities in their respective fields: decision-makers decide, social scientists conduct scientific experiments, which they must validate according to the rules of the art. However, researchers involved in

29. The Interministerial Mission on Annual Projects for Performance (2013) (« Programme 137. Égalité entre les femmes et les hommes », In *Mission interministérielle. Projets annuels de performance. Annexe au projet de loi de finances pour 2013. Solidarité, insertion et égalité des chances*, pp. 107-130.) Online http://www.performance-publique.budget.gouv.fr/sites/performance-publique/files/farandole/ressources/2013/pap/pdf/PAP2013_BG_Solidarite_insertion_egalite_des_chances.pdf (accessed 3 May 2016).

30. They will relate to “the development of agreements in companies and the improvement of their quality; to orientation and gender diversity to expand the share of girls following a scientific and technical education and their share in the corresponding professions, but also to promote female-dominated occupations among boys; on the training of the beneficiaries to complement their free choice to work during parental leave, in order to reduce the effects of being away from work” (p. 110).

In addition, “other experiments may be implemented in the field of workplace equality or in the fight against violence, particularly with regard to extending measures such as the programme providing women in danger with mobile telephones (*Téléphone grande danger*).

An experimental programme will be based on the principles of partnership and support/accompagnement (with social partners, associations and communities). The experimental programmes will be most often selected through calls-for-project procedures” (p. 127).

experimenting reforms are still subject to the risk of their findings being instrumentalised: for example, as with the researchers involved in an experiment on the use of anonymous Curriculum Vitae (CVs), prior to the publication of a government decree (BEHAGHEL *et al.*, 2011).³¹ By invoking the results of the study, the Commissioner for Diversity and Equal Opportunities, Yazid Sabeg, renounced publishing the decree that would have made anonymity mandatory for candidates applying to the companies concerned. However, the authors of the study had not given a conclusion “on the full effects of generalising anonymous CVs”, while they specified that the study did not allow “for the existence of discrimination in hiring to be tested” as they had stated a few months earlier, in a long letter to the press.³² Researchers therefore clearly face difficulties in controlling the interpretations of their results by their sponsors. But such differences in interpretation are hardly surprising: in fact they bear out the autonomy of political justification relative to scientific justification, and should lead researchers to look elsewhere to obtain the validation of their work, rather than in political decision-making.

Finding a Place for Random Social Experimentation

The major criticism that can be made concerning social experiments, in the French tradition, is that they have blurred the line between research and policy making (GOMEL, SERVERIN, 2009). To restore the distinction, two dimensions should be considered: i) the ethical concerns in the conduct of research, and ii) the necessary control of experiments by scientific communities, to ensure independent research.

The Necessary Consideration of Ethics

Various kinds of ethical control may be exercised in experimental social research. Apart from the ethical commitments of researchers themselves with regard to their sponsors (public or private), these involve especially authorisation and control mechanisms that govern research on human beings, in particular in the area of biomedical research. Can these rules be applied to social experiments? There is no simple answer to this. Rules are set out in legal codes relating to the ethics of experimental research focusing on human life (the human body and health), but not on human behaviour. *A priori*, the definition of biomedical research in Article L. 1121-1 of the Public Health Code (in the version derived from the Ordinance of 23 February 2010) does not seem

31. Article L. 1221-7 of the Labour Code (modified by a Law of 22 March 2012) states that in “companies with at least 50 employees, the information mentioned in Article L. 1221-6 and communicated in writing by jobseekers can only be examined by preserving the candidate’s anonymity”, and adds that “the modalities for applying the present article have been determined by decree by the Council of State”.

32. Behaghel L., Crépon B., and Le Barbanchon T. (2011), « CV anonyme : ce que dit l’évaluation », *Libération*, 27 April 2011.

suitable to this type of research: “The research conducted and practiced on humans for the development of biological and medical knowledge is authorised under the conditions provided for in this book and is designated below as being ‘biomedical research’”. However, Article L. 1121-3, paragraph 4 of the Code also clearly provides an (indirect) reference to experiment on behaviour: “in the sciences of human behaviour, a qualified person, together with the investigator, may direct research”.³³

This simple statement has found a normative extension in Law No 2012-300 of 5 March 2012 on research involving human beings, following the Bill submitted to Parliament by Olivier Jardé (Member of Parliament), in 2009. Its author wanted to give a status to non-interventional studies (such as monitoring cohorts), to allow them to be reviewed by ethics committees. The law introduced in Article L. 1121-1 of the Public Health Code distinguishes between interventional research (which involves intervention on humans) and non-interventional research, “in which all acts are performed and products used in the usual way, without additional or unusual procedures for diagnosis, treatment or monitoring”.³⁴

This text puts experimental research under the supervision of committees protecting persons. No experimentation on people’s behaviour, claiming to be “scientific” will henceforth be able to escape presentation before ethical bodies. This therefore only excludes administrative-type experiments, relating to the organization of government bodies, including assessments which do not seek to be scientific but which are technical and administrative, and which are similar to impact studies (GOMEL, SERVERIN, 2011).

The law thus also supports ethics, as has now been identified for several years by the CNRS Ethics Committee (*Comité d'éthique du CNRS – Comets*). In a first opinion published in 2007, the committee recommended particular caution concerning experiments on behaviour which may involve psychological risks when applied to vulnerable persons, such as “persons in social difficulty, immigrants, prisoners, drug addicts”.³⁵ In 2010, a second opinion focused on the ethics of research in social experimentation, and also recommended the submission of projects to ethical principles, and their publication in journals specialised in the experimental sciences.³⁶

33. This specification was introduced by Law No 94-630 of 25 July 1994, and was upheld by the Law of 7 July 2011 in paragraph 4 of the same article.

34. Article L. 1121-3 includes a 7th paragraph stipulating that “non-interventional research may be carried out under the direction and supervision by a qualified person. The Committee for the Protection of Persons will ensure the appropriate qualifications of investigators and the characteristics of research”.

35. COMETS, *Réflexions sur éthique et sciences du comportement humain*, 23 February 2007, p. 23 and: “A risk to mental integrity exists whenever an individual is subjected to conditions that may permanently alter their emotional equilibrium [...]. Risks involved include [...] exposure to emotional or aversive stimuli, using lures or distortions of reality (virtual reality), situations leading to the systematic failure of individuals, and putting individuals into competition or conflict with others”. Online http://www.cnrs.fr/comets/IMG/pdf/14-comportement_070226-2.pdf (accessed 3 May 2016).

36. COMETS, *Éthique de la recherche dans l'expérimentation sociale*, 19 January 2010. Online <http://www.cnrs.fr/comets/IMG/pdf/07-experimentation-sociale-20100119-2.pdf> (accessed 3 May 2016).

The Requirement for Monitoring Results by Peers

Independent research is not synonymous with research without control. Assessment and evaluation have always been included in the core institutional arrangements for public research, through two complementary processes: the control of expenditure, and control of relevance. The first is normally a responsibility of the public authority that allocates funds to research organizations; the second is provided by what is commonly called the “scientific community”. These principles of assessment are found in a specific chapter of France’s Research Code: the “Evaluation and monitoring of research and technological development”. Article L.114-1 in particular states that

“research activities funded in whole or in part by public monies, undertaken by public or private operators, are assessed on the basis of objective criteria tailored to each of them and based on the best international practices. Among these criteria, the contributions to the development of scientific culture [...] are taken into account”.

These assessments are carried out in full transparency, as provided by Article L.114-1-1:

“The procedures and results of the evaluation of a research activity funded in whole or in part by public funds provided for in Article L.114-1 are made public under conditions ensuring respect for secrets protected by law and the confidentiality clauses in contracts with third parties [...]”.

While social experiments do clearly fall within the domain of scientific research, as claimed by their promoters, and if they are necessarily publicly funded, then they should follow two paths of scientific assessment: i) assessment by peers, through publications in scientific journals with referee committees; and ii) the budgetary assessment of the funds they use, whose proportion should be measured relative to other public research funding. The first type of assessment clearly requires defining the disciplinary scope of studies (experimental psychology, social psychology, economics, statistics, etc.), in order to identify their contribution to knowledge. As for budgetary assessment, it seems, as seen above, that the purpose of the expenditure should shift from pure social engineering towards creating knowledge. This is a welcome development. The expenditure on random testing could then return to its natural place in public spending, that is to say as part of the “research and higher education” mission of government.

REFERENCES:

ALBEROLA, E., ANGOTTI, M., BRÉZAULT, M., and CREDOC (2008). « Enquête qualitative auprès des bénéficiaires du rSa, anciennement au RMI et à l’API, et de ceux qui n’ont pas pu y recourir. Synthèse des résultats intermédiaires. » In COMITÉ D’ÉVALUATION DES EXPÉRIMENTATIONS, *Rapport d’étape sur l’évaluation des expérimentations rSa*, Annexe 4. Online <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/084000607/0000.pdf> (accessed 3 May 2016).

ALLÈGRE, G. (2008). *L'Expérimentation sociale des incitations financières à l'emploi : questions méthodologiques et leçons des expériences nord-américaines*. Document de travail, No 2008-22, Paris: OFCE.

ANPE, DARES, and UNÉDIC (2008). *L'Accompagnement renforcé des demandeurs d'emploi : l'évaluation des expérimentations*, No 1. Online http://travail-emploi.gouv.fr/IMG/pdf/Evaluation_OPPCVE_n1.pdf (accessed 3 May 2016).

AVVISATI, F., GURGAND, M., GUYON, N., and MAURIN, É. (2011). *Quels Effets attendre d'une politique d'implication des parents d'élèves dans les collèges ? Évaluation de l'impact de la Mallette des parents. Rapport d'évaluation finale remis par l'École d'économie de Paris au Fonds d'expérimentation pour la Jeunesse dans le cadre de l'appel à projets lancé en 2007 par le ministère chargé de la Jeunesse*. Paris: Ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative. Online http://www.experimentation.jeunes.gouv.fr/IMG/pdf/APDIIESES_11_Rapport_final_evaluation_Mallettes.pdf (accessed 3 May 2016).

BEHAGHEL, L., CRÉPON, B., and LE BARBANCHON, T. (2011). *Évaluation de l'impact du CV anonyme. Rapport final*, mars. Online http://www.crest.fr/ckfinder/userfiles/files/Pageperso/rapport_eval_CVA_20110320.pdf (accessed 3 May 2016).

BEHAGHEL, L., and GURGAND, M. (2010). *Programme expérimental « Bourses aux projets de classe » : bilan de la phase pilote du point de vue de l'évaluateur*. Paris: École d'économie de Paris. Online <http://www.parisschoolofeconomics.eu/IMG/pdf/Pilote-BoursesProjets-PSE-juin2010.pdf> (accessed 3 May 2016).

BEHAGHEL, L., CRÉPON, B., and GURGAND, M. (2009). *Évaluation d'impact de l'accompagnement des demandeurs d'emploi par les opérateurs privés de placement et le programme Cap vers l'entreprise. Rapport final*, septembre. Online http://travail-emploi.gouv.fr/IMG/pdf/Rapport_Final_CRESENTSEE.pdf (accessed 3 May 2016).

BIAIS, B., and RULLIÈRE, J.-L. (2003). « Approches expérimentales en économie et en finance. » *Lettre du département Sciences de l'Homme et de la société*, 66, 48-50.

CAHUC, P., and ZYLBERBERG, A. (2009). *Les Réformes ratées du président Sarkozy*. Paris: Flammarion.

COMITÉ D'ÉVALUATION DES EXPÉRIMENTATIONS (2008). *Rapport d'étape sur l'évaluation des expérimentations rSa. Synthèse*, septembre. Online <http://lesrapports.ladocumentationfrancaise.fr/BRP/084000607/0000.pdf> (accessed 3 May 2016).

COMITÉ D'ÉVALUATION DES EXPÉRIMENTATIONS (2009). *Rapport final sur l'évaluation des expérimentations rSa*, mai. Online <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/094000222/0000.pdf> (accessed 3 May 2016).

COMMISSION FAMILLE, VULNÉRABILITÉ, PAUVRETÉ (2005). *Au Possible nous sommes tenus. La nouvelle équation sociale : 15 résolutions pour combattre la pauvreté des enfants*. Paris: Ministère des Solidarités, de la Santé et de la Famille. Online <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/054000264/0000.pdf> (accessed 3 May 2016).

FISHER, R. A. (1935). *The Design of Experiments*. Edinburgh, London: Oliver and Boyde.

GOMEL, B., and SERVERIN, É. (2009). *Expérimenter pour décider ? Le RSA en débat*. Document de travail, No 119, Noisy-le-Grand: Centre d'études de l'emploi.

GOMEL, B., and SERVERIN, É. (2011). *Évaluer l'expérimentation sociale*. Document de travail, No 143, Noisy-le-Grand: Centre d'études de l'emploi.

JANIN, P. (2005). « L'expérimentation juridique dans l'acte II de la décentralisation. Observations sur une réforme. » *La semaine juridique – Administrations et collectivités territoriales*, 41, 1334.

LÉVY-LAMBERT, H., and GUILLAUME, H. (1971). *La Rationalisation des choix budgétaires*. Paris: Presses universitaires de France.

LONCLE, P., MUNIGLIA, V., and RIVARD, T. (2008). « La mise en œuvre de l'expérimentation du RSA à partir des enquêtes qualitatives réalisées dans cinq départements. » In COMITÉ D'ÉVALUATION DES EXPÉRIMENTATIONS (2009), *Rapport final sur l'évaluation des expérimentations rSa*, Annexe 3. Online <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/094000222.pdf> (accessed 3 May 2016).

MINISTÈRE DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DE LA VIE ASSOCIATIVE (2011). *Rapport du conseil scientifique du Fonds d'expérimentation pour la jeunesse pour la période mai 2009-décembre 2010*. Paris: Ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative. Online <http://www.experimentation.jeunes.gouv.fr/IMG/pdf/rapport-cs-fej-2010.pdf> (accessed 3 May 2016).

PARIENTÉ, W. (2008). « Analyse d'impact : l'apport des évaluations aléatoires. » *STATECO*, 103.

RODRIK, D. (2008). *The New Development Economics: We Shall Experiment, but How Shall we Learn?* Revised paper for the Brookings Development Conference, John F. Kennedy School of Government: Harvard University.

SALVAT, X., and BOCCON-GIBOD, D. (2013). *Rapport à Madame la garde des sceaux, ministre de la justice, sur l'expérimentation des citoyens assesseurs dans les ressorts des cours d'appel de Dijon et Toulouse*. Paris: Ministère de la Justice. Online <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/134000144/0000.pdf> (accessed 3 May 2016).

SERVERIN, É. (2012). « Les causes et les effets du non-recours au RSA activité. » *Revue de Droit sanitaire et social*, 4, 637-645.

STAHL, J.-H. (2010). « L'expérimentation en droit français : une curiosité en mal d'acclimatation. » *Revue juridique de l'économie publique*, 681.

VIVERET, P. (1990). *L'Évaluation des politiques et des actions publiques. Rapport au Premier ministre*. Paris: La Documentation française.