OPINION 131

ETHICAL FRAMEWORK FOR EDUCATIONAL EXPERIMENTATION IN REAL-LIFE SITUATIONS

NATIONAL CONSULTATIVE ETHICS COMMITTEE
FOR HEALTH AND LIFE SCIENCES

C

OPINION 131

ETHICAL FRAMEWORK FOR EDUCATIONAL EXPERIMENTATION IN REAL-LIFE SITUATIONS

Opinion adopted on 27 June 2019

Rapporteurs: Lionel Naccache & Frédéric Worms

CONTENTS

SUMMARY

For a number of years, cognitive psychology and educational neuroscience have been producing results that can lead to educational recommendations. The comparison of these results with other positions, from different fields of expertise, has given rise to debate. This unprecedented situation means that all these results have to be tested in the least arbitrary way possible. There is such a rational solution: experimentation in real-life teaching conditions. The application of these methods in schools requires an ethical framework, which the CCNE sets out in its new Opinion “Ethical framework for educational experimentation in real-life situations” (Opinion 131 published on 7 November 2019). The CCNE also makes a number of recommendations.

Cognitive psychology and educational neuroscience produce results that can lead to precise recommendations on how to teach or learn. The comparison of these results with other positions, from different fields of expertise, has given rise to debate, and even to frank opposition. This situation is unprecedented and has led the CCNE to take up the issue on its own initiative.

In view of the large number of studies, their sometimes divergent interpretations and their potential applications in the world of education, the CCNE, in its new Opinion 131,[1] considers it “urgent and essential to sort out these results in the least arbitrary way possible, in order to guarantee the best educational decisions for schoolchildren”. There is such a rational solution: experimentation in a real-life situation. It originated in medicine with the clinical trials approach, and has since been developed in other disciplines, such as economics and public policy evaluation. In the field of education, a number of studies have already confirmed the relevance of such interventions. This approach makes it possible to take into account all forms of experimentation, to use a rigorous methodology (based in particular on randomisation and the concept of controlled experimentation), and even to control any bias. This in situ measurement of a teaching practice differs from laboratory experiments, which are conducted in a very different context from that of the classroom.

The CCNE’s aim is not to decide on the superiority of one teaching or learning method over another, but to propose a general ethical framework for this research, which it considers essential.

For pupils, classes and their teachers, there are a number of important questions, including:

[1] Opinion 131: “Ethical framework for educational experimentation in real-life situations” adopted by the plenary committee on 27 June 2019.

- How can the risks inherent in experimentation be minimised?

- How can we avoid interfering with current practices, delimit the strictly scientific framework of the research under way and guarantee its independence from the educational and political choices made elsewhere?

- How can we approach the issue of informed consent from minors?

- How can we ensure that the experiment and its spin-offs are socially equitable? What criteria should be used to determine the notion of educational effectiveness?

- How can we manage the risk of methodological standardisation?

The CCNE's deliberations have been shaped by a number of issues arising from the tension between the search for the most accurate pedagogical knowledge possible, ethical respect for the individual and the principle of scientific independence. A number of recommendations have been made. These include:

The ethical imperative of "beneficence" requires that the risks to which pupils are exposed be minimised. Thus the study should be preceded by research strongly suggesting the effectiveness of the planned experiment. The experiment conducted in real-life conditions should be as short as possible and involve as few pupils as possible, unless justified by the experiment. It is also suggested that the most appropriate tools be put in place to collect any undesirable effects of the studies conducted, and that the informed consent of under-age pupils be obtained, even if such consent is not legally required in France.[2]

The independence of research teams (particularly in relation to institutional practices or policies) is essential. It is up to them to establish a sound rationale for their studies, in which the effectiveness criteria being evaluated are precisely determined: while measurements of the child's "instrumental" performance (arithmetic, reading, etc.) are essential, so is the evaluation of critical thinking or creativity, which may be less easy to measure. Researchers should make the methodology and results of their work, both positive and negative, available to everyone, especially teachers.

Respecting the timeframe of this research means not confusing the timeframe of the experiment with that of any decisions to modify recommendations and current teaching practices.

Any real-life educational experiments should be supervised by an operational ethics body.

Educational experimentation should not be confused with an approach aimed at "medicalising" education or leading to methodological "standardisation". Because it is nourished by ongoing scientific work and educational experiments in the field, it seems desirable that it should take root over the long term, leading to a veritable "cultural revolution" among scientists, teachers and all those involved in education.

[2] Parental consent is mandatory.

Experimentation in real-life conditions could eventually form its own branch of educational science.

INTRODUCTION

This Opinion, which is the result of a self-referral by the CCNE, is based on five points:

(1) Over the last 20 years or so, cognitive psychology and educational neuroscience[3] have produced results that can lead to increasingly precise recommendations on the teaching methods for a wide range of knowledge and skills, particularly during the schooling period (from nursery school to secondary school).

The most resounding examples include: (i) the superiority of the syllabic method over the global or semi-global method when learning to read(Dehaene 2011); (ii) the importance of developing both an implicit sense of number (or numerosity) and explicit formal arithmetic when learning the basics of elementary arithmetic (see, for example, the experiments in real-life conditions recently conducted by the teams of Esther Duflo and Elizabeth Spelke);(Dillon, Kannan et al. 2017) or (iii) the need to adapt our teaching practices to the principles of organisation, and above all the development of memory,(Eustache and Guillery-Girard 2016) attention(Lachaux 2015) and executive functions (Houdé 2017) in children. These results are most often based on cognitive psychology experiments conducted in laboratories. These experiments involve testing several factors in order to identify their effects on learning performance. These are assessed both in terms of cognitive performance and functional and structural brain imaging (Bourassa, Menot-Martin et al. 2017).

(2) The comparison of these results, essentially from cognitive psychology and to a lesser degree from educational neuroscience, with other positions, from different fields of expertise, has given rise to debate and even to frank opposition. This situation is unprecedented.

(3) It is essential to test, confirm and compare these results in the most comprehensive and least arbitrary way possible, in order to guarantee the best recommendations for the educational decisions that will subsequently be taken. Let us not forget that the right to education is explicitly mentioned in the Universal Declaration of Human Rights (1945, article 26), and that there is an "*obligation de moyens*" ("obligation of means") here, which confronts us with an ethical problem. The time we waste resolving these decisions in the most reasonable, and therefore fairest, way places us at fault or in debt to the institutions and to the children who are

[3] By "educational neuroscience" we mean all the knowledge relating to education and learning that comes specifically from cognitive neuroscience, i.e. the combination of experimental cognitive psychology and neuroscience (including functional brain imaging and its many tools and methods). This field is undergoing significant development in France and worldwide. These empirical studies, together with their interpretations and the theories developed on the basis of them, are at the origin of debates and original proposals in the field of learning and education, whose potential relevance outside the laboratory requires further validation through experiments in real situations.

the first to be affected. The scientific nature of the results produced by educational neuroscience makes the situation even more demanding.

(4) There is such a rational solution: experimentation in a real-life situation, i.e. in real-life teaching conditions, which is similar to and forms part of the wider field of scientific experimentation carried out in real-life conditions and with real people. This solution makes it possible to take into account all forms of experimentation, to control any bias and to use a rigorous methodology (based in particular on randomisation and the concept of controlled experimentation). It differs from other forms of experimental initiatives and approaches that do not meet these methodological criteria. The challenges of educational experimentation in real-life conditions are major and include the possibility of directly measuring the consequences of an educational practice in real-life conditions (and not indirectly under the artificial conditions of laboratory experiments, which usually involve a much smaller number of individuals and a much shorter study period). A second major challenge is to open up and develop experimentation in schools and educational establishments in the broadest sense. This on-site experimentation could eventually form its own branch of the educational sciences, given the expertise and technical nature of such large-scale interventions.

(5) As soon as it is formulated, this observation of the need for educational experimentation to be carried out in real-life situations (mainly in schools, but also in extracurricular settings) introduces a second necessity: we need to think about the ethical framework raised by this very specific approach. A number of ethical questions immediately arise:

- How can the risks inherent in experimentation be minimised?

- How can we avoid interfering with current practices, delimit the strictly scientific framework of the research under way and guarantee its independence from the educational and political choices made elsewhere?

- How can we approach the issue of informed consent from minors?

- How can we ensure that the experiment and its spin-offs are socially equitable? What criteria should be used to determine the notion of educational effectiveness?

- How can we manage the risk of methodological standardisation? In particular, there seems to be a parallel between educational experimentation in real-life conditions and clinical trials, another form of experimentation in real-life conditions which have revolutionised medicine since the second half of the twentieth century and which have been the subject of rich and abundant reflection on the ethical framework which should govern them. But this analogy is by no means identical, and it is essential to highlight the differences between these two forms of experimentation, in order to define the boundaries of the ethical framework appropriate to the subject we are dealing with here. It should also be emphasised that the educational context is, of course, specific and that education cannot be reduced to the apprenticeships studied here.

So many questions that are already outlining the contours of the framework that the aim here is to sketch out.

In the present Opinion, we propose to raise each of the essential ethical questions raised by this observation, and to formulate recommendations of a deontological nature and appropriate principles for reflection. In itself, there is nothing technically original about this approach to defining an ethical framework, but we found no trace of it in the many available documents we consulted, or in the expert interviews we conducted. In the United States, the National Research Council has produced an essay which advocates six very general main recommendations (Education 2002), but does not contain a general framework such as the one we are seeking to propose.[4] There are several other countries in which educational experimentation in real-life situations is already being conducted, but we have not found an explicit general ethical framework.

We also feel it is important to define this Opinion in negative terms by explicitly stating what it is not:

(1) The aim of this Opinion is by no means to settle any of the questions currently being addressed by neuroscience and educational science, but to reflect on the overall framework for implementing educational experimentation in real-life conditions, so that it can be taken into account in a lucid, fair and appropriate manner.

(2) This Opinion concerns experimentation in real-life situations, but obviously does not disqualify other forms of research or existing methods in the field of pedagogy or education. It aims, on the contrary, to explicitly define a general framework that should be taken into account whenever a real-life experiment is envisaged.

(3) This Opinion in no way seeks to medicalise education by drawing certain parallels between medical experimentation in real-life conditions in medicine (clinical trials) and educational experimentation in the field, which is the subject of this Opinion. We discuss both the similarities and the profound distinctions between these two fields. The similarities seem to us to be helpful in thinking about and defining the framework we are looking for, but we are just as keen to make it clear why such an analogy

[4] Here are the six common-sense principles laid down by the US National Research Council in 2002:
- Pose significant questions that can be investigated empirically.
- Link research to relevant theory.
- Use methods that permit direct investigation of the question.
- Provide a coherent and explicit chain of reasoning.
- Replicate and generalize across studies.
- Disclose research to encourage professional scrutiny and critique.

constitutes a real risk. Problem #7, which we identify below, is entitled: "Risks of medicalising education".

Finally, we would like to shed some light on the question of the legitimacy of the CCNE taking up this issue: why would an advisory institution with an activity relating to the life and health sciences be relevant on this subject? Three points are worth mentioning:

- According to the WHO, health includes: "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity". Health is "one of the fundamental rights of every human being, regardless of race, religion, political opinion, economic or social condition". In this context, it seems clear to us that education is one of the key factors in "mental and social well-being", and as such, anything that focuses on education should not be alien to us.

- It should also be recalled that we are talking here about an ethical framework for scientific research into learning as a biological and psychological capacity, research whose methodology therefore comes under the life sciences, which are the very subject of the CCNE.

- Finally, it should be recalled that, in all its Opinions, the CCNE itself considers societal issues, including those relating to education and information; it is also certain that clear information for players on current scientific research on education is one of the ethical requirements.

<u>Finally, a few points of context:</u>

- The CCNE took up this subject and the hearings conducted by a specific working group between 6 January and 15 June 2016 in accordance with its methodology, confirmed the expectations of the players and provided input for this text. It should also be noted that this self-referral predates – and is independent of – the creation by the Ministry of Education of an interdisciplinary scientific council. Jean-Michel Blanquer and Stanislas Dehaene were also among the experts interviewed for this Opinion, even though they were neither Minister nor Chairman of the Scientific Advisory Board.

- The significant delay between this self-referral and the present Opinion stems from the fact that the CCNE organised the General Assembly to review the Bioethics Law. This intense activity has resulted in the implementation of an original methodology, and in work that has occupied all our resources: 154 hearings were held between 15 February and 24 May 2018; several meetings with experts on the issues addressed; training/discussion sessions with the citizens' committee; drafting of the Summary Report (https://www.ccne-ethique.fr/sites/default/files/eg_ethique_rapportbd.pdf) and then Opinion 129 "Contribution of the CCNE to the review of the Bioethics Law" (https://www.ccne-ethique.fr/fr/actualites/lavis-129-contribution-du-ccne-la-revision-de-la-loi-de-bioethique-est-en-ligne).

- Despite this delay, the permanence and relevance of the issues we wanted to address in the present Opinion has not only been confirmed, but seems to have been accentuated by the growing interest in educational experimentation in real-life situations (numerous national and international conferences organised since then). The Summary Report of the General Assembly mentioned above (e.g. p.73) also referred to public interest in studying the potential impact of educational neuroscience on school practices, and in taking account of the related ethical issues. The hearings also identified the desire of all the players involved (including those from the voluntary sector) to participate collectively in this discussion with the more traditional players (including teachers and other education experts, as well as researchers). This is a step in the direction of public participation in the ethics of life sciences and health in general. This citizen participation could also motivate research targeted at specific objectives, such as situations of extreme poverty or disadvantaged environments.

- We also note that the increasing presence of digital tools in the field of education (educational games, educational applications, MOOCs, etc.) is a further source of concern for children's mental health and psychological and social well-being. E-education thus seems set to play an increasingly important role in learning and education, and its tools also need to be part of an approach framed by the ethical thinking proposed here.

OUTLINE OF THE OPINION

In the first part, we will develop the rationale[5] we have just outlined very briefly, in order to highlight the importance of the issue we have taken up. The first three of the five points set out above will be detailed and illustrated: (1) the indisputable development of neuroscience, and in particular cognitive educational psychology; (2) the comparison of these results with other conceptions or results from other disciplines; and finally (3) the urgent need to confirm these different results in the most independent and least arbitrary way possible.

The second part of the Opinion will briefly present "educational experimentation in real-life conditions" through a few real-life examples with comments.

In the third part, we will set out the main issues that we feel have been raised by educational experimentation in real-life conditions.

In the fourth and final part, we propose a framework and some recommendations for conducting such experiments. Our thinking has been inspired in part by the solutions already found to certain ethical questions arising from other fields that involve experimentation in real-life conditions, such as clinical trials in medicine, and the economic and public health studies developed more recently. We have sought to highlight both the similarities and the differences between these approaches, which are already available, and the subject that concerns us here.

Rather than setting out the recommendations relating to each of the questions posed in the third part, immediately after they have been formulated, we have preferred to set out all the questions before detailing the recommendations. Indeed, there is an advantage in taking the whole picture into account, both for the questions and, subsequently, for the recommendations, which are often interlinked. Thus we find it easier to understand the question as a whole and the ethical framework for responding to it.

Finally, we will close this Opinion with a summary of this proposal for a general ethical framework, in the form of a summary sheet covering each of the problems explored and proposals for possible solutions.

[5] The term "rationale" here refers to the rationality of a proposed experimental study, i.e. the past scientific, theoretical and experimental argumentation underpinning this project.

DEVELOPMENT OF THE INTRODUCTORY RATIONALE

We have deliberately presented the rationale behind our approach in a concise manner in order to provide readers of this Opinion with the clearest possible vision of our approach.

Having done this, it is important to develop and reference the initial five-point observation that served as the starting point for our thinking, starting with each of the first three points set out. The following parts of the Opinion will develop the last two points.

1. For the last 20 years or so, educational neuroscience, and in this case cognitive psychology, has been producing results that can lead to precise recommendations on the teaching methods for a wide range of knowledge and skills.

Our intention here is not to delve into the technicalities of the issues being debated, both for the sake of synthesis and because of a lack of expertise, but simply to highlight the many areas of education that have been the subject of recommendations from the cognitive sciences and neurosciences. <u>We feel it is important to point out here that most of these recommendations do not come from functional brain imaging studies (often carried out on fairly small numbers of subjects), but from experimental psychology studies conducted according to the rigorous principles of cognitive psychology,</u> i.e. the analysis of cognitive processes in terms of information processing and representation operations. This nuance helps to dispel a fear that is often expressed, between an approach that is more concerned with children as “subjects”, and a technoscience that sees them as “objects”. In reality, educationalists and experimental psychologists share a concern for the same “subjects”, and have the same basic interactions with them; but their methods of observation, their objectives and their respective fields of expertise differ. This proximity and difference is what justifies both the use of real-life teaching experiments, and a strict distinction with current teaching practices, without direct interference of any kind.

Here are four typical examples of experimental work from the cognitive sciences and cognitive neuroscience leading to possible recommendations of educational value, and providing objective guidance when faced with methodological trade-offs.

<u>**Understanding and learning a text by trying to remember it**</u>

In 2011, Karpicke and Blunt reported in the journal *Science* (Karpicke and Blunt 2011) a remarkable study showing that – in the experimental context they had defined – learning the meaning of a school text was clearly favoured by the use of a spontaneous retrieval method, compared with conceptual learning comparable to what is commonly practised in schools and universities. They enrolled 80 undergraduates and divided them into four groups of 20 subjects. All the students read the same text, then depending on the group to which they belonged they were exposed to one of the following four conditions: (1) no further exposure until the test phase (below); (2) repeated exposure to the same text for a total of four readings; (3) work carried out in the presence of the text, consisting of drawing conceptual maps between the ideas developed in the text, after rapid practice in this type of work familiar to these students; (4) instruction given to try to remember spontaneously the ideas in the text they had just read, then a second presentation of the text and repetition of the spontaneous memory recall exercise.

The first important result was that the immediate memorisation performance (group 4) and that of the concept maps drawn and written (group 3) during this learning phase were identical.

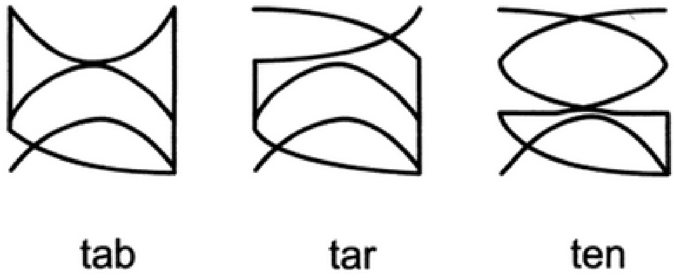
A week later, the students were questioned on the content of the text, both from a formal point of view (recall of elements and ideas) and from a more conceptual point of view (e.g. inferences deduced from the reading). Remarkably, students in group 4 were significantly better than those in the other three groups (including those in group 3, reflecting the dominant teaching methods). Interestingly, this study also included a metacognition test, carried out at the end of the learning phase, in which students were asked to identify which of four possible methods they thought was the most effective for learning and understanding this text. Unequivocally, the students (even those in group 4) were convinced – wrongly – that method 3 (and not 4) was the best.

The researchers replicated and generalised their study in a second experiment involving 120 students, reported in the same article.

This study was subsequently adapted and successfully replicated by the same group of researchers with primary school children (average age = 10) (Karpicke, Blunt and Smith 2016). Irrespective of unequal levels of reading proficiency, the spontaneous recall method led to better learning results and better comprehension of the verbal material studied. **In other words, this work suggests that the spontaneous recall method should be more widely used and developed than explicit formal learning that does not involve spontaneous memory retrieval exercises.**

<u>Superiority of an analytical reading method over a global reading method</u>

In 2006, the McCandliss team in the United States published an original and enlightening experimental study (Yoncheva, Blau et al. 2006). They developed a new alphabet made up of lines and curves organised from bottom to top (see Stanislas Dehaene's book *Reading in the Brain* for details of this work (Dehaene 2007)).



An important trick was not to separate the different letters used, so as not to necessarily make people aware of the existence of infra-lexical units corresponding to the letters making up each word. The figure above shows the equivalent of the "t" in position 1 (lower part of the word) in the first three words (tab, tar, ten). Researchers then set out to learn to read this new alphabet by comparing an analytical method with a global method. They recruited two groups of students and asked the subjects in the first group (global method) to memorise the global form of the new words, while those in the second group (analytical method) were explicitly informed of the existence of a sequence of letters written from bottom to top, which they could use to identify and learn the grapho-phonemic transformation code.

Each day, the subjects had to learn to read a list of 30 new words written in both the new alphabet and the Latin alphabet.

Three main results were observed:

Firstly, on the first day, global reading was accompanied by better word identification performance than analytical reading.

Secondly, the initial ease of learning diminished with each passing day, as new words were learnt at the expense of consolidating the words learnt in previous days.

Conversely, the analytical readers made slow but steady progress, and their learning performance was generalised both to old words (which they read better than at the start but did not revisit regularly) and, above all, to words they had never encountered before.

Finally, analysis of functional brain imaging data (in this case functional MRI) revealed that while global reading was correlated with activation of regions of the right hemisphere known for global apprehension of visual forms, only the analytical reading method was associated with activation of a region of the left hemisphere known to underlie reading expertise (the "visual word form area" (Cohen, Dehaene et al. 2000)).

Taken together, these results seem to confirm the superiority of an analytical approach based on the explicit acquisition of the grapheme-phoneme code over a

global reading method, while regaining the initial subjective and objective ease of a global approach. Initial ease quickly replaced by flagrant inefficiency.

The rich digital capacities of the very young, long unsuspected

One of the most spectacular, and long-standing, illustrations of the unsuspected results of the experimental approach used by cognitive psychology concerns the digital skills of young children (Mehler and Bever 1967, Dehaene 1997). As we know, Piaget's influential constructivist school postulated that the concept of number was not acquired by children until they were between four and five years old, as they interacted with their environment and adults. One of the classic Piagetian experiments, which seemed to confirm this conception, consisted in presenting the child with two horizontal rows of absolutely identical marbles (see figure below), and asking him/her: "Is it the same, or does one of the rows have more marbles?"

o o o o

o o o o

After the child had responded, the critical experimental condition was presented, in which the concept of length (assumed to have been acquired by the child) was contrasted with that of number (assumed not to have been acquired by the child), by presenting the child with two rows of marbles, one of which was longer but contained fewer marbles than the second (see figure below):

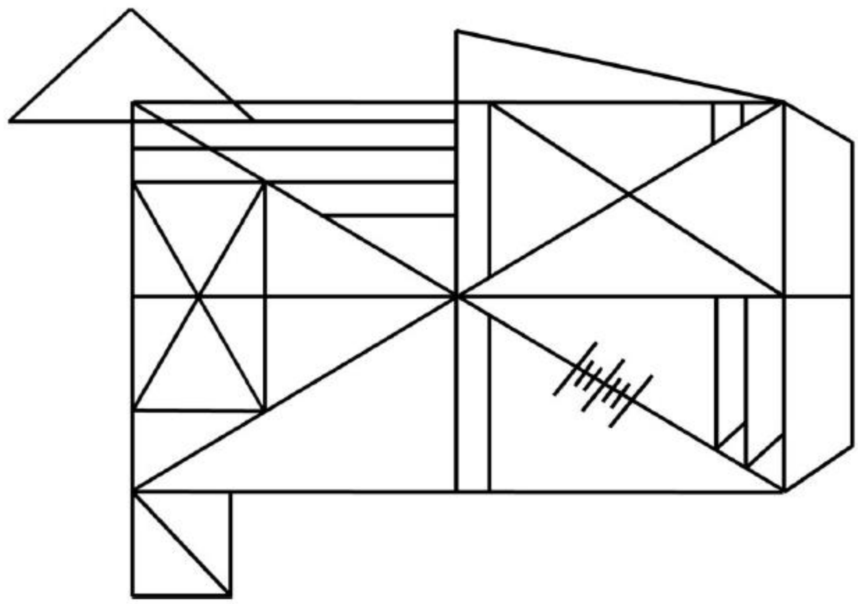
o o o o

o o o o o o

The child was then asked the same question about the new arrangement of the marbles:
"Is it the same, or does one of the rows have more marbles?"
Piaget and his colleagues observed that up until the age of around four or five, children made mistakes and tended to identify the longer row as the one with the larger number of marbles.
In 1967, Mehler and Bever (Jacques Mehler was one of the researchers who introduced the cognitive sciences to France) modified Piaget's task slightly, taking the view that the question could be misunderstood by children because of its complexity and, above all, its incongruity in the context of ordinary conversation: the first question could seem so obvious to the child that he/she might be led to try to understand something other than what was stated, and so make a mistake in the second answer, but for different reasons from those inferred by Piaget. To get away from this potential source of error and to get closer to the ecological and motivational value of the concept of number, Mehler and Bever replaced the marbles with M&M's chocolate sweets. The children were simply encouraged to choose the row containing the most sweets. The children, even the youngest aged two, overwhelmingly chose the row containing the most sweets, and therefore no longer failed this modified Piagetian test.
This historic experiment, followed by almost half a century of work on infants and young children, illustrates the very rich numerical, linguistic, social and, more broadly, cognitive capacities of the young human being, who is not, therefore, a tabula rasa.

<u>**The weight of gender stereotypes and how to overcome them**</u>

Experimental psychology can also be used to identify and quantify any subjective biases that interfere with learning and cognitive performance. For example, the stereotype that boys are better than girls at maths has been the subject of a great deal of empirical research. Huguet and Régner exposed male and female secondary school pupils to the same visual memory test under two different experimental conditions (Huguet & Régner, 2007). Each volunteer was tested in his/her classroom and, after giving their consent, was asked to observe and memorise as accurately as possible the following complex figure, adapted from Rey's figure and used in neuropsychology since 1941. After 50 seconds of observation, the figure was removed and the subject was asked to redraw it from memory. The quality of reproduction was quantified independently by two psychologists, using a scale that had been validated for several decades.



The trick in this experiment lay in the way the test was presented to the pupil. Randomly, each pupil was assigned to one of two conditions: in one group, the experiment was presented as a geometry test, while in the other group it was presented as a memory game. Although the initial stimulus and the task performed were identical, boys performed significantly better than girls when the experiment was presented as a geometry test, whereas the opposite performance profile was observed when it was presented as a memory game.
In several similar variants, we find the early impact of numerous stereotypes (not only of sexual gender). These experiments have also shown that minimal intervention is often enough to reduce or even eliminate them. For example, simply mentioning a little phrase to the pupil, such as "there is no difference in performance between men and women in this test", eliminates the effects of these stereotypes on pupils' cognitive performance.
This type of experiment illustrates the wide range of cognitive processes that can be explored by cognitive psychology: from the most instrumental dimensions (e.g. reading, arithmetic) to the critical sense, creativity and stereotypes that can affect a pupil's learning and mental well-being.

We could, of course, cite the numerous experimental studies carried out in the laboratory demonstrating the impact of attention, motivation, sleep quality and the importance of inhibiting automatic responses in the acquisition and performance of numerous cognitive tasks.

A fundamental point needs to be emphasised if we are to understand the demand that the present Opinion aims to regulate. These examples highlight both the strong scientific value of this work, but also its value, which is often very limited to the controlled experimental context in which it was obtained. The example of Mehler and Bever's seminal work also illustrates the distance that exists between an empirical result and its theoretical interpretation. The same reproducible experimental result can therefore be understood differently depending on the theoretical framework from which it is understood and interpreted. From the laboratory to the classroom, then (and for many reasons) there is a distance that needs to be taken into account, so that we do not simply extrapolate.

2. Comparing these findings from neuroscience with others from other fields of knowledge creates an unprecedented situation.

This proliferation of theoretical and experimental data from cognitive psychology and educational neuroscience on children's learning provides new knowledge on educational practices and concepts that needs to be compared with other knowledge. This confrontation is a challenge.

It therefore seems necessary to communicate this information to teachers and other professionals in the field of teaching and education. This access to information could be offered in both degree courses and continuing education courses. The twofold objective here would be not only to create the possibility of first-hand (and ongoing) access to this knowledge (and not just to pre-prepared material limited to the use of metaphors or analogies more or less faithful to the targeted knowledge), but above all to offer teachers the possibility of exercising a constructive critical eye on this knowledge, fuelled in particular by their own expertise. This approach is beginning to emerge in France, in philosophy for example, where some contemporary thinkers and researchers are managing to reflect on neuroscientific knowledge by avoiding these two pitfalls (see for example (Gillot 2007, Forest, 2014), which are, on the one hand, the stance of outright and systematic condemnation and, on the other, the "stupor and trembling" genuflection in the face of science. Similarly, the development of a rigorous, honest and fruitful form of "metabolisation" of neuroscience by education professionals seems to correspond to one of the desirable outcomes of the situation described above.

This situation sheds light on and reinforces the principle of the teacher's "pedagogical freedom", within a precise framework whose status was not legally defined until 2005 in the Fillon Law.[6] The first thing to remember is this: "the teacher's freedom to teach

6 Law No 2005-380 of 23 April 2005 on guidance and programmes for the future of schools

is exercised in compliance with the curricula and instructions [...] and within the framework of the school or establishment project, with the advice and under the supervision of the members of the inspection bodies" (article L912-1-1 of the Education Code), and the aim of the experiments, which is purely scientific, in no way directly determines the choices and laws established elsewhere.

However, as the SNUipp website[7] states:

"The Conseil d'Etat has reiterated on several occasions that the State defines the content of teaching and the duties of teachers. Thus the freedom to choose teaching methods, didactic approaches and types of mediation also remains. This principle has always been upheld as a guarantee of the school's independence from partisan pressures and respect for the rules of secularism. But it is also a way for teachers to ensure that 'the specificity and autonomy of their professional practice are respected' in relation to the administrative hierarchy and families. This freedom does not relieve teachers of their obligations, such as informing their pupils" (text available on the website: https://www.snuipp.fr/La-liberte-pedagogique#nb2).

It is therefore essential to share any scientific advances in fields related to their application.

3. It is becoming urgent, and essential, to decide on these results in the least arbitrary way possible, in order to guarantee the best educational decisions for the children at school.

Faced with debated methodological choices, the least bad way of solving these challenges is based on experimentation in real-life conditions. It is important to realise that the real-life experimentation described here is not based on measurements of brain activity that would be acquired in the classroom with pupils, but almost exclusively on behavioural measurements that are very comparable to those traditionally used in schools (reading performance, problem solving, development of critical thinking, etc.).

Indeed, to dispense with experimentation in such situations of methodological debate is to make our choices dependent on arbitrary factors, arguments of authority, or various power relationships. We should also note the crucial importance of using control conditions because of the frequent improvements caused by the experimental context regardless of the specificity of the experimental condition used: from the placebo effect to the Hawthorne effect (training effect), these non-specific effects,

[7] Syndicat National Unitaire des Instituteurs, Professeurs des Écoles et PEGC, affiliated to the FSU, Fédération Syndicale Unitaire.

which have led to the development of rigorous clinical trials, are also present in the field of educational experimentation.

In addition, rigorous, controlled experimentation carried out outside real-life teaching conditions runs the risk of extrapolating – wrongly – results which may be locally valid (experimental context different from real-life teaching conditions), but globally ineffective or inefficient.

EDUCATIONAL EXPERIMENTATION IN REAL-LIFE CONDITIONS

1. A REMINDER OF EXPERIMENTATION IN REAL-LIFE MEDICAL CONDITIONS.

Experimentation in real-life conditions was developed in the twentieth century, initially in agriculture (with the founding theoretical work of the statistician Ronald Fisher) before undergoing major development in medicine through what are known as “clinical trials” (Favereau 2014).

It should not, of course, be confused with medical experimentation on “human subjects” without their consent, the prohibition of which – after the Nazi crimes in particular – is at the very origin of bioethics, in 1945 at Nuremberg. The clinical trials in question have a very strong ethical framework in all countries, and this is one of the reasons, all other things being equal, why this opinion recommends a controlled transposition to educational trials.

The most original principle of this “clinical trial” approach is based on the concept of a controlled experiment involving randomisation of the numbers involved. Typically, in a clinical trial designed to study the efficacy of treatment X on disease Y, two experimental groups of patients are randomly assigned, one of whom will receive drug X while the other will receive a placebo (a controlled, double-blind, placebo-controlled trial). Random assignment of patients to one of the two groups eliminates many of the statistical biases that could otherwise lead to incorrect interpretation of the result. In order to show that X is more effective than the placebo, we need to ensure that all the other variables likely to distinguish these two groups (age, previous state of health, progress of the disease in question, sociocultural level, place of residence, exposure to possible environmental toxins, etc.) are adjusted. The absence of randomisation thus exposes patients to selection bias, while the absence of a control or comparative group exposes patients to confounding bias and non-specific management effects. Since their massive development in the second half of the twentieth century, clinical trials have revolutionised medicine, and today they constitute a dynamic, rich and complex speciality. Clinical trials are now essential for demonstrating the causal role of a drug, a device or any other intervention in the course of a disease.

2. REMINDER OF EXTENSIONS TO OTHER FIELDS AND THEIR SPECIFICS.

Of course, this general framework of "randomised controlled experimental field trials" goes beyond the strict field of health, and can be used in other areas such as economics, public policy evaluation or education. Recently, the North American J-PAL laboratory at MIT (Abdul Latif Jameel Poverty Action Lab) made a major contribution to developing this approach to studying the effectiveness of anti-poverty measures (https://www.povertyactionlab.org/fr), and now has 131 researchers in 40 universities around the world (Duflo 2010).

Here is a classic example: the Pratham experiment (Dillon, et al. 2017).

The project involved around 1,500 children aged four to five living in India, in 214 small nursery schools run by the NGO Pratham. It consisted in proposing a curriculum to improve the learning of mathematical foundations (concepts of number, etc.). This curriculum was developed in the Harvard laboratory by the teams of Liz Spelke, a leading specialist in cognitive development. Young women have been recruited to implement this programme in schools. This mission has therefore not been entrusted to the people who usually look after the children in these schools. The children were tested on individual computers on four occasions: (1) before the experimental phase of the programme, which runs for a few hours a week for eight weeks; (2) then just afterwards, (3) six months and finally (4) a year later, when they have entered primary school.

Participating schools were randomly allocated according to three criteria:

- Introduction of a curriculum based entirely on mathematical games;
- Introduction of a control condition in the form of games stimulating social skills, built according to the same structures and the same types of rules as mathematical games, but based on the recognition of emotions. This experimental control was necessary because the simple fact that the children were stimulated by playing, according to rules, one at a time, could in itself constitute an intervention different from what they experienced in normal circumstances. This intervention could therefore produce a general beneficial effect on learning which would be reflected in the assessment tests (intervention effect), independently of any specific impact on mathematical acquisitions;
- No intervention and assessment of usual school functioning.

This trial showed that exposure to mathematical games was responsible for a significant improvement in children's performance both immediately after the end of the programme and six months and one year later, compared with the other two experimental groups. The lasting improvement observed mainly concerned non-symbolic mathematical skills (e.g. comparison of ratios, implicit addition, questions of topology), with no solid translation onto symbolic acquisitions (e.g. explicit manipulation of numerical quantities). On the basis of this experimental trial, a new trial was designed, seeking to improve the transition from non-symbolic to symbolic mathematical skills.

Many other similar trials have been carried out in various countries, particularly in France, as part of educational research undertaken or involving both public and private bodies.

The Pratham experiment we have just described meets the criteria set out above. It was randomised, controlled and prospective, and the main criteria for measuring the expected effect were defined before the start of the study.

3. THE NEED FOR A SPECIFIC ETHICAL FRAMEWORK

This trial is a perfect illustration of one of the essential characteristics of this type of experimentation: far from being a one-off event, it is part of a sustained movement over time aimed at advancing teaching methods through knowledge and experimentation. This means that familiarising ourselves with and adopting an experimental culture requires a real change of mindset on the part of the community of teachers and educational science researchers. These considerations reinforce the need to define a clear and sound ethical framework, and suggest that this ethical effort will also have to be a long-term one.

Furthermore, it is clear that we need to distinguish the concept of educational experimentation, which seeks to study the causality of a factor by manipulating it and observing the consequences of this controlled manipulation in learning, from that of educational field research, which is limited to a description devoid of any causal dimension. Surveys are valuable because of the richness of their descriptions, and they are often sources of decisive clues for tackling the question of causality, but they are different from controlled experimentation, which is the only way of demonstrating the superiority of one teaching method over another.

Finally, the ethical framework we are concerned with here covers a very wide range of types of experimentation. Although the primary target group are those being taught (usually minors), some experiments could also involve teachers or other staff involved in school education. In terms of scale, these experiments can target individuals, groups of pupils, classes, even schools or academies. In terms of the type of experimentation, it is worth noting the possibility of using digital tools, whether connected or not.

LIST OF ETHICAL ISSUES TO BE CONSIDERED

The idea of educational experimentation in real-life conditions immediately exposes a tension between three principles.

On the one hand, we can define a **principle of seeking the most accurate pedagogical knowledge possible.** Taken to extremes and considered in isolation, this principle invites unconstrained experimentation in order to identify, design and select the best educational solutions.

On the other hand, a **principle of non-harm must obviously be formulated immediately, driven by the fear of testing certain experimental teaching conditions that could prove harmful to the children concerned.** This precautionary principle obviously takes on its full meaning in a perspective that places the individual as the primary object of its concerns and can act as a brake on the application of the principle of knowledge.

Finally, there is the principle of the **independence of scientific research**, whose sole aim is to establish and advance knowledge without any direct interference with current practices, their evaluation or, lastly, the political choices made elsewhere, which may take them into account.

The tension between these principles gives rise to problems to which we do not claim to provide definitive answers, but which must be explicitly addressed in any educational experimentation. These problems form the basis for the definition of an ethical framework that concerns us in this Opinion. Some of them lead to clear recommendations that are relatively easy to implement, while others require constant and evolving attention.

Among the problems arising from the tension between the search for truth, ethical respect for the individual, and scientific independence, we have identified eight which we will now develop, before formulating specific recommendations on them in the final part of this Opinion.

Problem #1: Minimising the risks inherent in educational experimentation

The need to implement educational experimentation in real-life conditions means that the risks associated with it must be minimised, and in this case targeting schoolchildren and teenagers receiving instruction.

This major problem of minimising risks must be considered at all three stages of the experiment:

1) First of all, upstream, in order to anticipate these risks and limit them as much as possible, by designing experiments that are as useful and as low-risk as possible.
2) During the experiment itself, by respecting the principle of the minimum duration necessary to observe the expected educational effect, and by developing monitoring systems that allow an experiment to be interrupted if it proves too damaging, or if its results are obvious even before the end of the period initially planned.
3) After the experiment, in order to identify possible late educational damage, and to compensate for it through personalised follow-up measures, or even active compensatory measures (e.g. making the method identified as beneficial at the end of the trial actively available to pupils in a control group).

This problem of minimising risk is all the more important in that it constitutes the most remarkable distinction between educational experimentation and clinical trials: whereas the problem of risk in experiments imposes an obligation in clinical trials to avoid harm – a local variation on the "primum non nocere" which guides therapeutics in general – in educational experimentation there is an imperative of beneficence. The children subject to this experimentation must not only be free from undesirable effects, but they must also acquire the fundamental skills whose teaching methods are precisely the subject of these experiments.

Problem #2: Informed consent obtained from minors

The very framework of the experiment requires information and, above all, individual consent. One of the peculiarities of this problem as applied to educational experimentation concerns the fact that the vast majority of the individuals taking part are minors. We will be able to draw on the work carried out here in paediatric clinical trials (recourse to the parents; obtaining the child's consent, which is empowering for the child, rather than legal, etc.). In a recent experiment, Olivier Houdé dealt with the issue of minors' consent as follows:

"We are obviously seeking informed consent from the parents, since the study involves minors, but we are also seeking token consent from the children via small booklets. We take a lot of time to explain the protocol to the children and, once they have been briefed, we ask them to circle the answer corresponding to their choice (I have decided, yes/no, to take part in the programme) and, for the youngest children who cannot yet read, to circle the little man who smiles for 'yes' or the one who does not smile for 'no'. We also make it clear to children that they can always change their mind and stop taking part in the research at any time if they wish, which is in line with the law in force in this area. To tell you the truth, none of the children in any of the groups we observed ever said they wanted to stop taking part."

The corollaries of this problem concern the methods used to implement these experiments and the basic level at which randomisation takes place (e.g. individual, class, school, extracurricular setting, internet experiment).

It should be noted that this question of the type of consent must be considered in the light of current developments in school education. Thus the development of training courses delivered over the internet (e.g. MOOCs for schools) now raises the question of whether individual consent can be obtained for an experiment carried out over the internet.

Problem #3: Independence from institutional practices, evaluation and implementation

Experimental projects must be developed within the framework of existing and independent research, and must demonstrate that there is no interference of any kind (including possible conflicts of interest) with existing or future institutional practices and policies. Let us take a brief look at the ideal way in which teaching practices could evolve as a result of advances in scientific knowledge, through a four-stage process:

1) Research in experimental psychology formulates precise hypotheses based on theorising empirical data, most often obtained in laboratories;

2) This preliminary stage then leads to the implementation of an experiment conducted in real-life conditions, which is the subject of this Opinion. At this stage, it is obvious and fundamental that the research should not interfere with current recommendations or teaching practices, outside the strict framework of the experiment in question;

3) Analysing and interpreting the results of this experiment can then lead researchers to formulate pedagogical recommendations and communicate them to teachers and educational institutions;

4) It is only at the end of this process that changes to the recommendations and current teaching practices can be made. This stage must be carried out independently of the research carried out.

Finally, even if this research is carried out, it should be remembered once again that education is not just about learning, and that other parameters, such as economic, social and cultural factors, come into play, and other sciences (sociology, economics, history, for example, among others) are needed.

Problem #4: Social equity of the experiment and its spin-offs

Optimistically, the knowledge gained from this type of experimentation should lead to improved teaching methods and environments for all schools. This raises the question of the absence of an identity relationship between the subjects tested and those benefiting from these tests. This link between the conditions under which this knowledge is acquired and its spin-offs raises a question of social equity. Let us imagine, at the extreme, a school A undergoing systematic and uninterrupted experimentation, and a school B exempt from any experimentation and building its educational programmes on the precious results of the experimentation carried out in school A. The unfairness of such a situation, in which some schools would be places of intensive experimentation, while others would reap the benefits without contributing directly to them, is obvious.

Problem #5: Determining educational effectiveness criteria

Another aspect of educational experimentation is the criteria defined and used to assess the quality of any teaching method. Should we limit ourselves to performance criteria that are operational and fairly simple to use, or should we extend this metrology to values that refer to the child's autonomy to think and learn beyond the strict content of what is to be studied, to his/her critical spirit, his/her motivation and his/her subjective well-being? There is a possible bias here, which would tend to favour experimentation with certain dimensions of learning to the detriment of others, not because the former are intrinsically more interesting than the latter, but because they are easier to measure or quantify. In short, it is a bias similar to that of the individual who looks for his/her lost keys at night under a lit street lamp, not because there is a greater chance that his/her keys will be there, but because it is the best-lit place, the one where it is easiest to look. This is a major point, because the existence of such an effect on the scale of all the experiments conducted could ultimately tend to give a biased value to certain educational objectives. Generally speaking, if real-life experimentation corrects the biases of the laboratory, it can also bring about others (for example, greater motivation on the part of the participants) which it is up to the experimenters, as is always the case, to take into account (see in particular the discussion of the placebo and Hawthorne effects above, p.14).

The duration of observation, and therefore of measurement of these criteria, also appears to be a source of questioning: should the acquisition of a skill be verified at a distance from the end of the experiment? It is interesting to note that these questions are not new, but that their importance is increased tenfold by the idea of a pedagogical experiment aimed at selecting the most beneficial teaching methods in a non-arbitrary manner.

Problem #6: Objectives set at group or individual level? The risk of methodological standardisation

Extending the question about the nature of the criteria leads almost naturally to the question of normative risk and the relationship between the individual and the group. Should the experiment be designed in terms of objectives set for a group of children (e.g. percentage of children having acquired a particular skill), or should it be tailored to the individual status of each of the children involved in the experimental protocol? It seems quite legitimate to consider that the ideal answer to this question will vary according to the type of learning: standardising skills in terms of mastery of grapheme-phoneme transcoding seems *a priori* laudable, whereas standardising the way of using one's imagination to solve a novel problem seems much less so.

Problem #7: Making a lasting contribution to the concept of educational experimentation

To the ethical issues described above, we must add another that brings them all together in a lasting way. In fact, educational experimentation as we have explored it is not so much about answering – once and for all – a pre-established list of closed questions. Rather, it appears to be the method we need to face up to the educational challenges of the present, but also of the future, which are largely unknown to us. In other words, we need to think of this experiment as a structural solution, not a short-term one. We need to see it as a sustainable solution that we can incorporate into our thinking and practices. This "problem" of sustainably integrating the method raises a number of cascading issues: how can this culture of educational experimentation be integrated into all the stakeholders involved (in particular teachers, pupils and parents, but also other players (headteachers, academic inspectors, rectors, etc.) and, more broadly, civil society)? How can we ensure that these experiments are carried out in a controlled, supervised, monitored and transparent way? How can priorities for experimentation be established?

Problem #8: Risks of medicalising education

Finally, the last problem we have identified is the risk of medicalising education. Importing, for the legitimate reasons set out above, an experiment largely inspired by the clinical trial exposes us to a transformation in our relationship to teaching and learning: the way we look at a difficulty or failure encountered by a child could slide towards a representation of a pathological nature, independently of any proven medical pathology. The problem raised here is therefore to be distinguished from that of learning disorders, which seem to be related to neurological disorders, often developmental, such as certain syndromes including dyspraxia, dyscalculia, or disorders of executive functions or theory of mind, and for which specific support measures need to be proposed as soon as possible.

MAIN RECOMMENDATIONS: TOWARDS AN ETHICAL FRAMEWORK FOR EDUCATIONAL EXPERIMENTATION IN REAL-LIFE CONDITIONS

Identification and definition of the eight risks, outlined above, or problems raised by educational experimentation in real-life conditions seems to us to constitute the first layer of the ethical framework we are seeking.

It seems to us that each experiment should explicitly address each of these risks (for example in the form of a documented file) and provide the most appropriate responses. For most of the problems raised, these responses will obviously vary from one experiment to another, depending on the objectives and the precise context of their implementation. This necessary flexibility further reinforces the need to formulate these questions and the answers given explicitly, in order to ensure the ethical validity of the experimental project in question. This observation immediately leads to another: not only does this ethical framework need to be configured individually for each project, but it should also be analysed by an independent and competent body. This recommendation would take the form of ethics committees for educational experimentation in real-life conditions, independent of the educational structures where the experiments are to be conducted, and whose membership should include teachers, specialists in educational science, statisticians and specialists in educational neuroscience.

With this in mind, it now remains for us to address each of the problems raised in the previous section in a serial manner.

Problem #1: Minimising the risks inherent in educational experimentation

In order to minimise the risks detailed above, we make the following recommendations:

<u>A solid and promising rationale before experimenting</u>

It is obvious, but fundamental, to submit to the framework of experimentation in real-life conditions projects that are highly likely to improve the quality of teaching in light of the criteria that will have been defined (see problem #4 and the proposals made on this subject below). This level of evidence, which should be available before the experiment begins, will ideally be based on a combination of sound theoretical considerations, descriptive observations and empirical data acquired in the laboratory outside real-life conditions. According to the framework proposed here, this supported rationale should be examined by the ethics committee whose creation we recommend.

It should be stressed here that this rationale should be informed by data from the educational sciences and cognitive neuroscience, but that it should also draw on the expertise and initiative of teachers in the field. This last point, which underlines the importance of both a top-down and a bottom-up approach, leads us to recommend easy access to teachers who would like it, with contacts from the research community, so that they can test these ideas from the field before experimentation in real-life conditions, and thus involve them more directly in the birth of original experimental projects that are likely to succeed.

The shortest possible experiment

The second principle of risk minimisation is based on the choice of the shortest possible experimentation period. This obvious point reinforces the need to carry out a statistical power study beforehand, in order to estimate this minimum duration as accurately as possible. This recommendation is in no way intended to prohibit large-scale studies or studies involving large numbers of children, if the experiment in question can justify the importance of this parameter. We simply recommend that this essential factor be determined taking into account this principle of risk minimisation.

An experiment carried out on the lowest possible number of pupils and/or classes

Similarly to the question of the duration of the experiment that we have just discussed, this power study should also make it possible to estimate the minimum number of subjects or classes necessary to demonstrate the expected effect. This recommendation is in no way intended to prohibit the possibility of conducting large-scale trials when rationally justified, but to reiterate the importance of minimising risks.

Identifying any undesirable effects in real time

All experiments must monitor any undesirable effects that occur during the experiment, and analyse the degree to which these unforeseen events are attributable to the experiment in progress. This monitoring will have to be carried out, if not in real time, at least at short intervals, so as to be able to decide to interrupt the trial if necessary.

<u>Searching for any undesirable effects after the end of the experimentation</u>

This monitoring of undesirable effects possibly linked to the experiment should continue after the end of the study, for a variable length of time, without being burdensome for those involved. One solution might be for teachers to report such events anonymously and confidentially to the sponsor of the experimental trial concerned.

Problem #2: Informed consent obtained from minors

Consent will have to be obtained from subjects who have reached the age of majority, but informed information will have to be given to minors taking part in the experiment. These minors will also have to give their consent, even if this does not constitute valid consent in the absence of the consent of the adults responsible (most often the parents), and therefore does not dispense with the need to obtain the latter's consent.

Problem #3: Independence from institutional practices, evaluation and implementation

Declaring any conflicts of interest between the researchers responsible for the experiment and the educational institutions could help protect this independence.

A strict distinction between the scientific context of experimentation, in the context of ongoing research, and the prospects for application would also be verified by the relevant ethics committees.

On the other hand, we believe it is vital to make the methodology and results of this research available, whatever the outcome: publication of all the results obtained, both positive and negative, reinforces the quality of this work and minimises the risks of inappropriate use of this data to guide future teaching practices.

Problem #4: Social equity of the experiment and its spin-offs

At an individual level, this principle could be guaranteed by a very simple protective measure: a pupil could only be included in a limited number of educational experiments during his/her school career. If it were to be applied, this measure would require a national centralisation of completed and ongoing experiments.

At the level of the educational institution, the question of social equity is more complex to address. In order to avoid discrimination between schools, on the one hand, and to avoid introducing a culture of coercion, on the other, we could think about rewarding experimentation by allocating specific material and human resources, in proportion to the number and quality of the teaching experiments carried out. This approach, which needs to be sustained over time, could lead to the most involved centres also being the most innovative and beneficial for pupils. This dynamic of experimentation that benefits children seems possible.

Problem #5: Determining educational effectiveness criteria

The criteria used will obviously depend on the educational objective in question, which is the real issue to be defined here. More often than not, a major objective will lead to several elementary or intermediate objectives, which are often organised in a hierarchical manner. Thus learning to read a text does not have a single objective, but involves a whole range of skills, from learning how to transcode between words and sounds, to acquiring syntax, the ability to summarise and reconstruct the content of a text, and critical analysis. Knowing how to read refers to all these skills, the acquisition and mastery of which obviously cannot be assessed using the same criteria. While it is quite easy to assess the performance (quality/speed/effort) of the grapho-phonological transcoding, it is obviously more difficult to assess the capacity for analysis and critical perspective. The educational objective chosen and the proposed criteria must therefore be clearly formulated. Beyond a specific experiment, it is also important to determine whether all the teaching experiments undertaken around a single general skill (e.g. learning to read) are concentrated around a limited number of basic objectives (e.g. acquisition of grapheme/phoneme transcoding). Such biases could arise as a result of pressure being exerted on a few objectives to the detriment of others, and also because some of these objectives are much simpler to assess (and perhaps to achieve) than others. This reflection on the supervision of experiments that have already been carried out or are under way leads directly back to our recommendation in problem #6 (see below), which sets out the need for a form of cultural “revolution” regarding the sustainable integration of concepts and practices associated with educational experimentation.

Problem #6: Objectives set at group or individual level? The risk of methodological standardisation

In our view, there is no single answer to this important problem. One way of preventing this is to explicitly ask this question before any educational experimentation, and also to seek the informed opinion of a panel of specialists (see next point) before implementing it.

Problem #7: Making a lasting contribution to the concept of educational experimentation

Awareness of the ethical issues raised by educational experimentation in real-life conditions has led us to the following conclusion, which has already been mentioned in several places: educational experimentation is not a one-off or limited stage over a short period of time, but rather consists of a lasting change in our relationship to educational practices. This observation implies a "cultural revolution" which involves the long-term integration of experimentation, and therefore also the long-term integration of the ethical issues raised by its application. This spirit of dynamic change in our relationship with education and learning practices is also reflected in a number of initiatives, such as *Les Savanturiers*, a programme of education through research created in 2013 by François Taddei and Ange Ansour;[8] the more targeted programme to renew science education, *La main à la pâte*, launched in 1995 by physicist Georges Charpak;[9] and the recent report entitled "Towards a learning society: Report on research and development of lifelong learning", written by Catherine Becchetti-Bizot, Guillaume Houzel and François Taddei.(Becchetti-Bizot, Houzel et al. 2017)

We therefore make the following five recommendations:

(1) It would be desirable for a personal protection committee (inspired by the PPC model) to supervise these experiments. Based on the tried and tested model of research conducted on human subjects, the aim would be to submit planned projects to this type of PPC, which would also be involved in monitoring the quality of the research carried out. The composition of these committees should obviously include teachers, specialists in educational sciences and educational neurosciences, but also specialists in epidemiology and clinical trial statistics, specialists in the human and social sciences, as well as people from associations, particularly parents'

---

[8] The programme can be consulted on the website hosted by the Centre de Recherches Interdisciplinaires: https://les-savanturiers.cri-paris.org/a-propos/presentation/

[9] You can consult the Fondation La main à la pâte website: https://www.fondation-lamap.org/

associations. The number and organisation of these committees could be determined gradually, depending on the momentum of the educational experiment.

(2) With regard to the problem of the independence of the research conducted (problem #3), the projects submitted to these PPC-type committees should also provide proof of their scientific independence, and in particular set out any conflicts of interest relating to this experimentation.

(3) In order to supervise all the trials carried out, and also to satisfy the recommendation made in response to the problem of social equity, anonymised data should be centralised within a single structure. This would make it possible to dynamically analyse the quantity and quality of searches carried out.

(4) Critical analysis, communication and discussion of this research should be encouraged through publications in specialist peer-reviewed scientific journals, as well as publications aimed at civil society. This essential stage, which is also open to similar research carried out abroad, contributes directly to the richness and relevance of this experimental approach.

(5) Lastly, training in the concept and practice of experimentation should be provided for teachers, either during their training or during in-service training activities.

Problem #8: Risks of medicalising education

Let us not forget that educational experimentation is not medical experimentation, because the children (or adults) involved in an educational project are not sick people, and education is not a treatment, but a right enshrined in the 1948 Universal Declaration of Human Rights. It is important to remember and emphasise this difference in order to keep in mind the non-medical dimension of education. In fact, this trivial observation risks being undermined by the methodological development of educational experimentation which, as we have seen, shares many features in common with the clinical trial born in the context of medical experimentation. The protocols tested should not be thought of as treatments, and individuals who show different results in these experiments should not be considered as patients.

One way of marking this difference could be to improve the screening and treatment of genuine learning pathologies and disabilities, such as dyspraxia, dyscalculia and dyslexia, or those more widely encountered in neurological or psychiatric pathologies. The present Opinion is not concerned with this other crucial issue relating to education, but improving these two aspects (experimentation on the subject on the one hand, and care for learning disabilities/pathologies on the other) could make it easier to distinguish between them. This final recommendation would also pave the way for possible crossovers between these two issues, by considering experimentation with teaching practices specifically designed for pupils affected by these disabilities.

SUMMARY OF THE ETHICAL FRAMEWORK AND PROPOSED RECOMMENDATIONS

Problem #1: Minimising the risks inherent in educational experimentation

1-1. Drawing up a sound and encouraging rationale before experimentation; in this rationale, strictly delimiting the scientific scope of the experimentation
1-2. Conducting the shortest possible experiment
1-3. Carrying out the experiment carried out on the lowest possible number of pupils
1-4. Identifying any undesirable effects in real time
1-5. Searching for any undesirable effects after the end of the experimentation

Problem #2: Informed consent obtained from minors

2.1. Obtaining consent from adults (parents, legal guardians)
2.2. Providing clear information to the minors and adults concerned
2.3. Also obtaining the consent of minors, despite the fact that this has no legal value

Problem #3: Independence of research

3.1. No interference with current practices
3.2. Independence from decision-making institutions
3.3. Examination of any potential conflicts of interest
3.5 Publishing all the methodologies and results obtained

Problem #4: Social equity of the experiment and its spin-offs

4.1. Including a pupil in only one teaching experiment during his/her school career.
4.2. Promoting the experiment within participating schools.

Problem #5: Determining educational effectiveness criteria

5.1. Clearly formulating the educational objective chosen and the proposed criteria.

5.2. Determining whether the educational experiments undertaken are neglecting certain important objectives by concentrating on a limited number of them.

Problem #6: The risk of methodological standardisation

6.1. Explicitly asking the question of the risk of standardisation upstream of the experiment.
6.2. Obtaining the informed opinion of a competent independent panel before the experiment.

Problem #7: Making a lasting contribution to the concept of educational experimentation

7.1. Creation of a specialised national operational ethics committee (based on the model of the Personal Protection Committee or PPC set up to oversee biomedical research on human beings) to supervise these experiments.
7.2. Centralisation of anonymised data within a single structure.
7.3. Encouraging critical analysis, communication and discussion of this research.
7.4. Providing training for teachers in educational experimentation.

Problem #8: Risks of medicalising education

8.1. Regularly reminding all those involved (pupils, teachers, researchers, civil society) why education is a right and not a medical treatment, and why this distinction is essential to protect.
8.2. A reminder of the other factors in education and real teaching life that go beyond strictly cognitive learning (social contexts, issues of justice, teaching content and curricula and its place in culture, environments, particularly digital environments).

CHRONOLOGY OF THE OPINION AND LIST OF HEARINGS HELD IN 2016

The general idea for this Opinion was drawn up and proposed by Lionel Naccache in the spring of 2015, and the CCNE's decision to consider it on its own was taken in the autumn of 2016, with Lionel Naccache and Frédéric Worms as the two rapporteurs. The Neuroscience working group worked on this project and then scheduled the following hearings:

Hearing schedule:

- 6 January 2016: Esther Duflo
- 7 April 2016: Olivier Houdé
- 20 April 2016: Stanislas Dehaene, Anne Christophe
- 31 May 2016: Pierre Léna
- 15 June 2016: Philippe Meirieu, Jean-Michel Blanquer
-

We invited the Director General of School Education on 23 November 2016, proposing two dates, and received a negative response to this request for a hearing. A second letter, sent on 3 December 2016 and opening up a number of possible dates between January and March 2017, remained unanswered.

Esther Duflo is Professor of Economics at the Massachusetts Institute of Technology (MIT) and a founding member of the Abdul Latif Jameel Poverty Action Laboratory (J-PAL), an institution specialising in the randomised evaluation of anti-poverty programmes. She studied at the École Normale Supérieure and DELTA (Paris) as well as MIT. Her research focuses on household behaviour, educational choices, school enrolment, policy evaluation, decentralisation and microfinance. Esther Duflo is the first holder of the annual "Knowledge against poverty" chair, supported by AFD, at the Collège de France. She is a pioneer in the development of experiments in real-life situations, on a limited and precise question, with a comparison between a control group and an experimental group, as a method of analysis in economics.

Olivier Houdé is Professor of Psychology at Paris Descartes University, Sorbonne Paris Cité, founder and Director of LaPsyDÉ (CNRS), and a specialist in the development of children's intelligence and learning (experimental psychopedagogy). A teacher before moving into psychology and research, Olivier Houdé has a wealth of experience on the issues raised by children's learning.

Stanislas Dehaene is a cognitive psychologist and neuroscientist, Professor at the Collège de France, holder of the Chair of Experimental Cognitive Psychology, and a member of the Académie des Sciences; he is conducting leading-edge research into the brain bases of arithmetic and numbering, reading and consciousness. Stanislas Dehaene is the author of a number of scientific works on learning, as well as essays promoting the use of this knowledge in schools. Since December 2017 he has chaired the Conseil Scientifique de l'Education National created by Minister Jean-Michel Blanquer.

Anne Christophe
A CNRS researcher and co-director of the Laboratoire de sciences cognitives et psycholinguistique (LSCP), her work focuses on language acquisition in babies. Anne Christophe leads a research team dedicated to this topic.

Jean-Michel Blanquer
A former education authority rector and former director of the École supérieure des sciences économiques et commerciales (ESSEC), he was appointed Minister of Education in 2017.

Philippe Meirieu
A researcher specialising in educational sciences and pedagogy, Philippe Meirieu was the inspiration behind a number of educational reforms, including the introduction of modules at lycée, the creation of IUFMs in the early 1990s, and the introduction of supervised personal work (TPE) and civic, legal and social education (ECJS) as part of the 1998–1999 lycée reform.

Pierre Léna
An astrophysicist by training and a researcher at the Laboratoire d'études spatiales et d'instrumentation en astrophysique de l'Observatoire de Paris, Pierre Léna has been involved in science education for young people since the 1990s. He is Honorary Chairman of the Fondation La main à la pâte, created in 2011.

REFERENCES

Becchetti-Bizot, C., G. Houzel and F. Taddei (2017). Vers une société apprenante: Rapport sur la recherche et développement de l'éducation tout au long de la vie. Paris, Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche, 85p.

Bourassa, M., M. Menot-Martin and R. Philion (2017). Neurosciences et éducation: pour apprendre et accompagner. Bruxelles, De Boeck, 592.

Cohen, L., S. Dehaene, L. Naccache, S. Lehéricy, G. Dehaene-Lambertz, M. A. Henaff and F. Michel (2000). "The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients." Brain **123 (Pt 2)**: 291-307.

Dehaene (1997). "La bosse des maths.". Ed. Odile Jacob, Paris, 363p.

Dehaene, S. (2007). "Les neurones de la lecture."

Dehaene, S. (2011). Apprendre à lire: Des sciences cognitives à la salle de classe. Paris, Odile Jacob. OUvrage collectif, 155p.

Dillon, M. R., H. Kannan, J. T. Dean, E. S. Spelke and E. Duflo (2017). "Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics." Science **357**: 47-55.

Duflo, E. (2010). Lutter contre la pauvreté (tome 1), le développement humain.Paris, La République des idées, Seuil, 103p.

Education, (2002). National Research Council,Scientific research in education. Committee on Scientific Principles for Education Research. Washington, DC National Academy Press, 204p.

Eustache, F. and B. Guillery-Girard (2016). La neuroéducation : la mémoire au coeur des apprentissages. Paris, Odile Jacob, 172p.

Favereau, J. (2014). La lutte contre la pauvreté à l'épreuve des essais cliniques. Réflexion sur l'approche expérimentale de l'économie du développement. Paris alshs-00974686. https://halshs.archives-ouvertes.fr/halshs-00974686/document

Forest, D. (2014). Neuoscepticisme. Les éditions d'Ithaque, Paris, 208p.

Gillot, P. (2007). L'Esprit, figures classique et contemporaines. CNRS, CNRS Éditions, Paris, 320p.

Houdé, O. (2017). Apprendre à résister. Paris, Le Pommier, 144p.

Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*(3), 545-560.

Karpicke, J. D. and J. R. Blunt (2011). "Retrieval practice produces more learning than elaborative studying with concept mapping." Science **331**(6018): 772-775.

Karpicke, J. D., J. R. Blunt and M. A. Smith (2016). "Retrieval-Based Learning: Positive Effects of Retrieval Practice in Elementary School Children." Frontiers in psychology **7**: 350.

Lachaux, J. (2015). Le cerveau funambule : Comprendre et apprivoiser son attention grâce aux neurosciences. Paris, Odile Jacob, 312p.

Mehler, J. and T. G. Bever (1967). "Cognitive capacity of very young children." Science **158**: 141-142.

Yoncheva, Y., V. Blau, U. Maurer and McCandless. BD (2006). "Strategic focus during learning impacts the neural basis of expertise in reading." poster presented at the Association for Psychological Science Convention, New York, US.

OPINION 131

NATIONAL CONSULTATIVE ETHICS COMMITTEE
FOR HEALTH AND LIFE SCIENCES