

A Dutch treat: randomized controlled experimentation and the case of heroin-maintenance in the Netherlands

TRUDY DEHUE

ABSTRACT

In 1995, the Dutch Minister of Health proposed that a randomized clinical trial (RCT) with heroin-maintenance for severe abusers be conducted. It took nearly four years of lengthy debates before the Dutch Parliament consented to the plan. Apart from the idea of prescribing heroin, the minister and her scientific advisers had to defend the quite high material and non-material costs that would arise from employing the randomized controlled design. They argued that the RCT represented the truly scientific approach and was the royal way to unambiguous results. In the present article, I question this common dual justification of RCTs. First, I situate the historical origins and the basic assumptions of the ideal experiment in 20th-century economic liberalism. Secondly, using the Dutch heroin experiment as an example, I discuss human-science experimentation as an attempt to *create* reality rather than merely record it. Finally, I discuss some surprising responses by heroin users. These responses display the assumptions of RCTs discussed in the historical section, and underline the importance of the culture of heroin use. In the epilogue, I suggest that cultural aspects of heroin consumption can best be studied by thorough ethnographic research.

Key words classification, clinical trial, ethnography, heroin-maintenance, randomization

Mr Van de Camp, aren't you simply fooling yourself when you ask how the scientific protocol has been put together? As if you and I can make a judgement on this, for we are not scientists. You must just say now whether you want these experiments or not.

This provocative remark was recorded in the minutes of the Dutch Lower House, and was uttered by one Member of Parliament to another (Tweede Kamer, 1996: 47). In the second half of the 1990s, heated debates had taken place in the Dutch Parliament concerning the methodology of scientific research, including the question as to whether or not politicians could have a say in such matters. The disputes were triggered by a plan, proposed by the Minister of Health, Welfare and Sports, to conduct an experiment with heroin-maintenance for severe abusers. At the time, the number of heroin users in the Netherlands was estimated at 24,000, of whom some 12,500 participated in a methadone program on a regular basis. Of these 12,500 methadone recipients, about 8,000 did not really improve on methadone. The aim of the proposed experiment was to establish whether additional provision of heroin would enhance the medical and psycho-social condition of these 'therapy-resistant' users.

Decades of debates on heroin-maintenance preceded the proposal.¹ The minister, Els Borst, first presented the idea of an experiment in autumn 1995, and it was winter 1999 before she managed to guide the project through Parliament. Ignoring the exhortations from the minister's own party (D66)² that politicians should stick to their trade, quite a few Members of Parliament could not or would not separate scientific from political matters.

In this article, I discuss the Dutch heroin experiment as an example of human-science experimentation. I will argue that the outcomes of an experiment cannot be as neutral as many believe. Even an impeccably designed and perfectly conducted experiment necessarily involves implicit assumptions and hidden effects. The article ends with an epilogue discussing the advantages of solid ethnography, advantages which have become largely invisible due to the hegemony of experimentation. In order to clarify how the paper unfolds, it is first necessary to provide additional information about the details of the experimental design and the issues at stake in the Dutch Parliament.

METHODOLOGICAL DEBATES IN PARLIAMENT

The minister tried to convince Parliament of the benefits of obtaining outcomes that were more reliable than those achieved in earlier Swiss and British

heroin-maintenance experiments. Long-lasting debates on the effects of 'heroin on medical prescription' could be settled once and for all, she argued, if, and *only* if, a truly scientific experiment could be done this time. The minister, who once was a full professor of evaluation-research of clinical practice, explained that a truly scientific approach means conducting a randomized clinical trial (RCT). Clear-cut results, she elucidated, require a comparison between an experimental group receiving a combination of heroin and methadone, and a control group receiving only methadone. Close monitoring of both groups throughout the experiment enables the determination of any changes specifically in the heroin group. Of course, the experimental and control groups should be as equivalent as possible. For the sake of comparability, it is crucial to compile the groups on the basis of chance. If all this can be achieved, the minister pleaded, the first genuinely scientific experiment with heroin-maintenance will be conducted.

In spring 1996, the Dutch Parliament decided to postpone further decision-making. A scientific committee was asked to work out a detailed research protocol first. However, when more than a year later, this Central Committee on the Treatment of Heroin Addicts (CCBH) presented its report, the debate flared up once again. According to the CCBH, the number of participants in the experiment would have to be much larger than the minister had stated. The minister promised to convey to the CCBH that with fewer participants it would be easier to acquire political consent. She could not, however, persuade the CCBH to lower its standards. Based on statistical parameters and the requirement to demonstrate progress of at least 40% for at least 20% of the participants in the heroin group, the CCBH statisticians arrived at an absolute minimum of 656 participants. Taking into account specific dropout rates for each group, the total number to be enrolled would have to be 750 (50% of whom should be heroin inhalers, the other 50% heroin injectors).

Not only was this sample much larger than expected but *each* of the 656 to 750 participants, including the controls, would have to be promised heroin for a particular period of time. The minister had to make clear that the latter aspect of the protocol did not violate the principles of the RTC, but, on the contrary, was necessary to enable the experiment. It was not likely, she explained, that participants allocated to the control group would collaborate with thorough bimonthly physical and psychological examinations in exchange for only methadone and a standard small participant fee. This problem is well known in experimental treatment testing: if the treatment is a treat, it is hard to keep the non-treated controls involved. Before the CCBH wrote its 1997 report, methodological experts preparing the heroin experiment had already discussed the issue at length (see GG&GD Amsterdam *et al.*, 1996; Schellings *et al.*, 1996). From a methodological point of view, they found a clever win-win solution.

In order to ensure the participants' collaboration, it was argued, the controls should be promised heroin for a *following* period. This 'waiting list design' would ensure sufficient participation by the controls. Furthermore, if the original experimental group were to keep on getting heroin in this second period, it would be possible to compare a long-term heroin group with a short-term heroin group. That is, this would be possible on the condition that a *third* group would act as controls for both periods. Of course, this third group too would have to be motivated by the promise of heroin during a particular term.

It was decided that each of the terms should last six months. The CCBH report presented the idea cautiously. The need to promise heroin to the control group was a matter of 'sheer justice', and stimulating willingness to participate was only 'an additional advantage' (CCBH, 1997a: 17). Yet, originally, motivation of the controls was the primary reason for the extended RCT (GG&GD Amsterdam *et al.*, 1996; Schellings *et al.*, 1996). The minister, moreover, was quite open about it in Parliament and to the press she commented candidly: 'If you don't hold out a sausage to the controls, you can't coax any of them into participation' (quoted in the newspaper *NRC*, 4 September 1997).

The minister's social-liberal party D66 endorsed the plan, as did the Labour Party (PvdA), and the Green Party (Groen Links). Members of Christian parties, however, opposed it vehemently, mainly because, in their view, provision of heroin meant giving up on allegedly hopeless cases. The powerful right-wing liberals (VVD) were also against the experiment. The latter party mainly lodged objections to its high price of more than 50 million guilders (22.69 million Euro), and foresaw veritable rushes on the heroin-maintenance stations and serious disturbances in the surrounding neighbourhoods. Although having been previously rapped on his unscientific knuckles, the Christian Democrat Van de Camp insisted that 'as is well known with statistics, anything is possible'. And a colleague from the right-wing Liberal Party sided with him, wondering whether 'other scientific researchers might not have arrived at another experiment' (Tweede Kamer, 1997: 2 and 5).

The minister, however, no longer would let the opponents tamper with the research design. She insisted that a genuinely scientific experiment should be conducted, with the minimum number of participants as dictated by the established standards of valid research, or no experiment at all. She offered another compromise. If any serious trouble occurred during a planned preliminary pilot study in Amsterdam and Rotterdam, the rest of the project would be abandoned. Grudgingly, step by step, the opposing parties assented to the pilot study.

The story of the Dutch heroin experiment as told so far illustrates the importance attached in human science to randomized controlled (alias

randomized clinical) trials. Whereas in non-experimental 'observational' research 'various sources of bias occur that make the results interpretable in many ways', an RCT can guarantee 'unambiguous outcomes' (CCBH, 1997a: 12–13). In furtherance of this aim, human scientists are prepared to pay high prices for RCTs. In the present case, a large number of heroin recipients, years of laborious efforts to persuade the Dutch Parliament, a huge sum of money, and – as I will discuss below – enormous efforts to organize and carry out the experiment in various Dutch cities, were justified by the argument that the RCT represents the summit of science and provides clear-cut information.

In this article, I question this standard dual justification of RCTs. First, I discuss the historical origins of the ideal experiment. This experimental design appeared in human science textbooks from the 1930s to the 1950s, in psychology and other social sciences somewhat earlier than in medicine. Since that time, not only textbooks but also human science journals and publication manuals have conveyed the RCT as the highest standard of rigorous research and a matter of self-evident logic. In that sense it indeed became 'the' scientific approach. Historical research, however, provides a different image. As I will argue, the ideal experiment does not represent an *a-historical*, or transcendental, logic. Rather, the design originally developed as part of the 'liberal' variety of Western welfare regimes, which was distinguished by the renowned economist-sociologist Gøsta Esping-Andersen from the more caring 'social-democrat' and 'corporatist' regimes (Esping-Andersen, 1990, 1999). I argue that the RCT still carries and promotes some crucial maxims of the liberal, more reluctant, version of the Welfare State.

Whereas the next section of this article deals with the historical background and goes into assumptions about the RCT in general, the section that follows, discusses unavoidable choices in individual RCTs. Through such choices experimental human science also advocates implicit assumptions as if they were independent results. More specifically, using theoretical views on classification as a framework and the heroin experiment as an example, I argue that human science experiments, including RCTs, help to create the truth they are meant just to establish. Finally, I go into the course of events in the Dutch heroin-maintenance experiment from 1997 until 2001. Unexpected responses by the invited heroin users display the assumptions of RCTs discussed in the historical section, and the offer of heroin as an instance of the proverbial *Dutch* treat for which 'each person pays his or her own way'.³ If the usual arguments pro RCTs cannot convincingly justify their privileged status, neither can they support the rather low status of non-experimental research. The article ends with the suggestion that crucial cultural aspects of heroin consumption can be studied best with thorough ethnographic research.

HISTORICAL ORIGINS AND ASSUMPTIONS OF
THE RCT

I discuss the historical background of the RCT only as far as is necessary for the present argument.⁴ Moreover, I concentrate on the history of *social* experimentation, whereas the initiators of the heroin experiment repeatedly communicated that it was a *medical* experiment. The title of the 1997 CCBH report translates as 'Research on Heroin on Medical Prescription', and the CCBH emphasized that it preferred the latter expression to that of 'free heroin'. As the minister, characteristically straightforward, said in Parliament: 'If I just had said that we wanted to hand out free heroin to limit nuisance on the street, I am sure you would not have given permission' (quoted in the newspaper *NRC*, 4 February 1999).⁵ The CCBH chairman Van Ree explained it in more measured terms. Admitting that the experiment also tested the effects on social behaviour, he argued that this was done primarily in the interest of the heroin users themselves: 'We measure individual criminal behaviour, but not the effect on delinquency in the cities' (*NRC*, 23 August 1997). Yet, without doubting Van Ree's arguments, I focus on the history of social rather than medical treatment testing. After all, the experimental variable is not the heroin itself, but the means of acquiring it. And transforming heroin from an illegal drug into a medicine is not a matter of chemistry but of administration. Moreover, the government commissioned the experiment for the sake of policy-making. In these respects the heroin experiment indeed was a social experiment.

My history of active social experimentation begins in the last decades of the 19th century. The earliest proponent I have found was the British gentleman-economist (mathematician, philosopher) William Stanley Jevons.⁶ It is intriguing that Jevons's main example also concerned combating drug abuse through drug maintenance. In an article on 'Experimental Legislation and the Drink Traffic' (1880), he discussed the free trade of *beer* that was legalized in the 1830s in order to 'draw away people from the gin shops'. Jevons repudiated this idea as 'a salient example of bad legislation'. He vigorously argued that, rather than having been 'passed by the almost unanimous wisdom of Parliament', the Beer Act should have been tested scientifically. If that had been done, the state would have known beforehand that the 1830 Beer Act would only create 'a beastly state of drunkenness among the working classes' (Jevons, 1880: 183).

The analogy does not carry as far as to say that Jevons too considered his case a medical one, with beer as the tentative cure. Neither did he argue that a randomly composed experimental group should have been supplied with beer and compared with a control group receiving only gin. In his view, a scientific test only implied trying out interventions on a representative section of the population, and establishing any subsequent differences.

Moreover, it is important to note that, besides being an early advocate of science-based administration, Jevons was an early exponent of upper-middle-class enlightened liberalism pleading for some state protection of the working classes. In his book *The State in Relation to Labor* he stressed the fundamental maxim of individual responsibility for one's own success and failure. However, he also argued that this principle should not be driven so far as to leave factory workers completely to their fate (Jevons, 1882).

At the beginning of the 20th century, industrialized societies showed increased willingness to mitigate the excesses of unregulated capitalism. Administrative agencies were installed in order to initiate and steer social changes. Nevertheless, many contemporary commentators considered government involvement in the free market to be a hardly admissible intervention in private affairs. They did not regard the redistribution of income through central regulation as simply a matter of social justice. Many feared adverse consequences of state charity, and there was a widespread distrustful belief that administrations would squander public funds. Proponents of government interventions had to guarantee that help would not discharge people's sense of their accountability, and that expensive ameliorative policies would produce instant progressive effects. And, most importantly in the present context, the *proof* of these effects had to be of a special kind. In the words of Theodore Porter, the greater the resistance to centralized government in a society, the more the officials have to warrant their decisions in terms of standardized or 'mechanical', rather than interpretative or 'disciplinary', objectivity (Porter, 1995).

Such appeals for restrained governmental interference in the free-market economy provided the context in which the randomized controlled experimentation gradually emerged. In summary, three interrelated maxims of the renewed liberalism were of pivotal importance. The *first* and most important one was that the individual is the prime focus of administrative action. The crucial notion of individual responsibility dictated that remedying policies were first of all directed at integrating people who do not keep up with the given social order, which order itself is not a central matter of concern. The *second* maxim was that of the pivotal importance of effectiveness. The condition that central government should be kept to a minimum led to emphasis on immediate results of administrative interventions. And the *third* maxim was that of the importance of mechanical, impersonal judgement. The 'unanimous wisdom of Parliament', already mocked by Jevons in the 1880s, increasingly lost its authority. To a much larger extent than their 19th-century predecessors, 20th-century administrative agencies had to base their authority on 'objective', in the sense of 'independent' or 'impersonal', knowledge.

Early 20th-century human sciences seized the opportunity to extend their social niche. These sciences too began to adhere to the three maxims of help directed at individuals, at effectiveness and at impersonality. Economists,

sociologists and psychologists increasingly replaced trust in expert judgement with reliance on quantitative data acquired via standardized methods. Here too wisdom was now associated with whim, and philosophy at best with frill (Bannister, 1987; Danziger, 1990; Haskell, 1977; Porter, 1995; Ross, 1991).

It was no coincidence that these developments took place early and strongly in the USA. Even today the USA passes for the prototype of the liberal welfare regime that offers less organized protection of the socially vulnerable than 'social-democrat' or 'corporatist' regimes (Esping-Andersen 1990, 1999; Brown, 1999). Compared with other welfare societies, the USA still spends the least money on social care but the most on scientific control of its effectiveness (Jansson, 1993).

At the beginning of the 20th century, new forms of socio-technical research flourished. Using refined statistical techniques, American social scientists devised standardized scales for measuring phenomena such as delinquency, poverty, or various unhealthy habits. Soon a new idea arose: the repeating of the measurements after the implementation of ameliorative attempts. And a further issue was raised: that measurement before and after makes it impossible to decide whether a difference was induced by a planned intervention or some other concurrent factor.

In 1917, the American sociologist F. Stuart Chapin – pithily characterized by the historian Robert Bannister (1987) as a 'patrician/technician' – published two articles on the issue in the *Scientific Monthly* (Chapin, 1917a, 1917b). Chapin envisioned experiments with minimum wage bills, eugenic marriage laws, child-labour bills, and slum-clearing projects. Unlike Jevons in the 1880s (who wanted only to investigate a representative part of the population) but similar to the Dutch Minister of Health in the 1990s, Chapin maintained that before and after measurement was insufficient.

Yet, Chapin did not deem it *feasible* to make research on the effects of social amelioration fully airtight. A disturbing incomparability of social groups, he thought, would always plague experimentation with humans. Fundamental differences 'in race, political ideals, or standard of living, constitute the uncontrolled conditions which invalidate conclusions that may be drawn from much social experimentation' (Chapin, 1917b: 244). Chapin did not yet discuss the option of ensuring comparability by composing *artificial* groups. In later writings, however, he explicitly rejected the latter procedure on moral grounds, arguing that the experimental allocation and manipulation of people for the sake of research collides with the humanitarian mores of reform (Chapin, 1938, 1947).

It was psychologists who introduced the crucial idea of composing artificial groups rather than investigating natural ones. From about the 1870s, psychologists conducting psychophysical laboratory research on cognition and perception had developed meticulous skills in subjecting volunteers to

the strictest methodological regimes (Benschop and Draaisma, 2000). During the early 20th-century social changes, they adapted their established psychophysical research protocols to the new demands (Dehue, 1997). Many of them specialized in research on the efficiency of educational techniques (Callahan, 1962; Danziger, 1990). As before, the volunteers in laboratory experiments (schoolchildren and teachers) could comparatively easily be persuaded (or forced) to keep to the rules of an experiment.

Administrators striving for efficiency in education commissioned research on ways of learning to read, the impact of large versus small classes, fresh versus ventilated air, the sex of the teachers, and many other educational variables. In this context, the standards of mechanical objectivity were further narrowed down. Educational psychologists faced the need to design methods for composing groups that were as equal as possible. For instance, in testing new ways of teaching, the experimental and control groups should not differ in mean level of intelligence. Educational psychologists sought ways further to exclude arbitrariness.

Initially the problem was handled by subjecting children to preliminary tests on factors suspected of causing bias, such as intelligence, and subsequently forming groups with equal test results. From each pair of equally bright pupils, one was assigned to the experimental group and the other to the control group. This 'matching' procedure, however, was quite laborious. Even worse, the choice of factors upon which matching pairs were created strongly depended on the researchers' imagination. Apart from their intelligence, the children's sex, or ethnic background, or any other factor might also make a difference. Matching therefore violated the twin maxims of efficiency and impersonality cherished in administrative and social science circles alike.

In 1923, William A. McCall, an educational psychologist at Columbia University in New York, published a manual, *How to Experiment in Education*, which epitomizes the dual dictate of standardization and effectiveness. McCall calculated that psychological experimentation would save trillions of dollars for generations of Americans. Subsequently, he advanced a way to make his profession even more profitable for society. In his sections on how to create comparable groups, he recommended using chance as 'an economical substitute' for matching: 'equivalence may be secured by chance, provided the number of subjects to be used is sufficiently numerous' (McCall, 1923: 42). As far as is known, McCall's was the first methodological manual to present the RCT in human science.

From the 1930s, the British statistician and eugenicist Ronald Fisher circulated his statistical techniques based on random allocation (Fisher, 1935).⁷ Fisher's analysis of variance provided further arguments to educational psychologists propagating the randomized controlled design. The idea that valid results could be gained only in this way spread among other psychologists and outwards to a wider circle of social scientists (Dehue, 1997; Lovie, 1979;

Rucci and Tweney, 1980). The more that 19th-century *laissez-faire* thinking was replaced by the new social liberalism, the more it became acceptable to experiment with humans. From the 1960s, large-scale RCTs were conducted in the USA, in which sometimes thousands of people were involved. Apart from university students and other volunteers, people with limited social power acted as experimental subjects. Schoolchildren, soldiers, slum-dwellers, chronic drunks, spouse-beaters, wild teenagers and disabled food-stamp recipients were allocated randomly to experimental or control groups in order to test the effects of sentences, rewards, special training, social housing programs, marriage-courses, safe-sex campaigns, health programs, income-maintenance and employment programs (Boruch, 1997; Bulmer, 1986; Orr, 1999).

This brief history demonstrates that the RCT was accomplished as part of the rise of the liberal welfare regime, characterized by emphasis on individual accountability and distrust of central government. This experimental design embodied the three interrelated maxims: that 'objectivity' should be defined in terms of standardized impersonal procedures; that instantaneous effectiveness of ameliorative interventions should be demonstrated unambiguously; and that such interventions should be directed at individuals in need of integration into the given social order. Social order itself was neither a major matter of concern, nor, accordingly, a major subject of scientific research. As Kurt Danziger has cogently argued, only if the importance – or even reality – of inter-individual relations is denied, does it make sense to investigate people in artificial groups disconnected from their familiar school-classes, colleagues, friends, enemies and (in artificial settings) their familiar things. Only on the assumption that historically and culturally established relationships do not matter can the RCT be a rational research strategy (Danziger, 2000).

CREATING TRUTH

In his volume on the history of the RCT in American medicine, Harry Marks (1997) demonstrated that in medical research the design began to find acceptance somewhat later (see also Yoshioka, 1998). In this context too comparative experimentation was propagated primarily because of regulatory concerns and as a means to counter charges of subjectivity. In clinical practice, however, the procedure was not readily accepted. Physicians appealed to their expert discretion, and they maintained that medical ethics prohibited assigning diseased people to control groups or trying out treatments on healthy controls. Doctors, of course, had much more social status than their counterparts in early social experimentation. It was harder to break down their resistance than that of schoolchildren and their humble – mostly female

(Lagemann, 2000) – teachers who were allocated to experimental and control groups from the 1910s onwards. Marks describes physicians who continued to rely on clinical observation, refused to roll dice to determine their patients' lot, or used statistical techniques in incorrect ways.

Marks also recounted the many pitfalls of large-scale randomized clinical trials for testing drugs or diets. Owing to the impracticability of standardization and mutual alignment of thousands of people, the projects were often cancelled halfway or completely failed to settle the issues at hand. As to the latter problems, comparable stories can be told about the many large-scale social experiments that were conducted in the United States from the 1960s onwards. In spite of the use of the ideal experimental design, laid out, moreover, in ever more sophisticated research protocols combined with increasingly refined statistical techniques, critics frequently challenged the impartiality of the outcomes. Methodologists – whose professional speciality it is to trace remainders of bias – spotted flaws, as did those who politically opposed the tested actions from the very start. Within half a century after William McCall enthusiastically introduced random allocation to groups as a means of saving money while reducing decision-making according to whim, some social science experts began to wonder aloud whether the scarce gains in objectivity could ever outweigh the high expense of social experimentation (Lyons, 1969; Williams, 1971; Rossi and Lyall, 1976; Bulmer, 1986; Haveman, 1987).

Nevertheless, with the Americanization of the international human sciences, the RCT gradually became an international standard dissociated from its context of origination. Disappointing results in the past inspired many human scientists to further develop the methods of experimentation. Generations of human science students learned that the RCT represents the peak of methodological rigour, and quite a number of them have specialized in the increasingly complicated skills of implementing the RCT.

The Dutch heroin experiment also has been designed according to the highest standards. Detailed and ingenious protocols for every aspect of the project ensured that as little as possible was left to the discretion, preferences and interpretations of those conducting the experiment or those participating in it. The heroin had been produced in a way that allowed precise determination of quantities. The maintenance stations had been designed and equipped according to strict instructions. Base-line measurements of nuisance had been carried out in their neighborhoods. Participants had been selected following punctilious procedures, nurses had learnt exactly how to hand out the heroin and doctors were taught how to do the physical examinations. Interviewers had been trained through an internationally standardized training program, and detailed schemes were developed for reporting the results. At present the final report on the results is still in preparation. No doubt, depending on the outcomes of the experiment, either the opponents or the

proponents of heroin-maintenance will find gaps in the experiment's woven network of procedures. However, to me such infringements of the standard methodological protocol are only of secondary interest. I argue that even in perfect RCTs in which each collaborator or participant fully keeps to the experimental protocol, the results cannot represent reality as it is.

Historians, sociologists and philosophers of science have demonstrated that research methods cannot lay bare reality but actually shape the world they are supposed just to make known (Bowker and Star, 1999; Danziger, 1990, 1997; Hacking, 1986, 1999; Osborne and Rose, 1999; Rose, 1996; Stone, 1998). The basis of their argument is that every description demands classification. Describing reality means that it is ordered in categories, and the choice of categories depends on human decisions and cultural conventions. Stated differently, the distinctions applied do not simply represent natural kinds but are culturally embedded ways to make sense of human life. And, most importantly, categories or classifications do not just *reflect* conventions, but also reinforce and revise them. They can have serious normative and practical consequences.

The argument can be illustrated with a straightforward example. Whereas in many societies every heroin user is an heroin addict, some drug experts make a distinction between addicts and mere consumers arguing that it is possible to have one's use under control. Others also maintain that the usual concept of drug dependency is an expression of particular cultural views on motivation and behaviour (Peele, 1990; Gomart and Hennion, 1999). Obviously, such notions of drug use make far-reaching differences to people's rights and life. The same holds for distinctions such as the one made in Dutch law between soft drugs (for instance, hashish and marijuana) and hard drugs (for instance, heroin and cocaine). Dutch hard-drug consumers are relegated to the streets or at best to Spartan public 'user rooms', whereas consumers of soft drugs convene in nicely decorated commercial 'coffee shops'. Street dealers exploit hard-drug users, but soft drugs are inexpensive, if not home-grown.⁸ In such ways, distinctions between kinds become true. They are true, Ian Hacking argued, not because they represent natural kinds but because social processes *turn them into* independent realities. Classifications, Hacking argued, 'make up people' (Hacking, 1986, 1999).

Quantitative science is the social institution with the strongest power to reinforce or change categories. No matter how hard researchers try to avoid 'information bias' during a research project, even if it concerns a perfect RCT, the numbers resulting cannot be information-free but are significantly shaped by the culturally embedded protocols specified for acquiring them. The designing of surveys and tests demands the taking of decisions as to which categories to use, and how to further specify them in survey questions and test items. After a research project, the original decisions are removed from the construction like redundant moulds. The resulting data are hard, so to

speaking, because they have 'hardened'. Subsequently, they have the power to change reality. People react to them, and particularly so when authorities use them in social policies. The power of human science to create 'impressively resilient' new entities is 'second to none', as Theodore Porter expressed it (Porter, 1993: 93 and 95). Porter pointed in particular to administrative statistics as a vigorous means to create phenomena seemingly quantitative in themselves.

Although the results of the Dutch heroin experiment have not yet been presented, this experiment can illustrate the issue. This project also carries a large burden of consequential conventions. To begin with, the 'intake' diagnosis of 'therapy-resistant' addiction offers a forceful example. Heroin users were eligible if they had been dependent for at least five years, participated in a methadone-maintenance program in which they tried 60 milligrams of methadone daily for at least one month without interruption, contacted the methadone program at least 50 times in the previous six months, but still used heroin (almost) daily, and did poorly on particular tests (CCBH, 1997a: 22-3). Professionals who knew what they were talking about have, of course, established this definition of a 'therapy-resistance' heroin user. Nevertheless, no one will doubt that at least the chosen periods of time and the chosen quantities of methadone might have been different. Moreover, not everyone categorizes methadone as a therapy. Many consider it a substitute drug. Nevertheless, the chance of heroin-maintenance was withheld from users classifying themselves as irremediable addicts without having tried methadone. Conversely, the label of therapy-resistant addict implying the opportunity of heroin-maintenance was *granted* to heroin-methadone users who might have been classified as curable according to other criteria. A straightforward illustration that categories 'make up' people was the story of some heroin users who quit a detoxification program when the rumour spread of the advent of free heroin (Beems, 1995: 15). The experiment's seemingly neutral inclusion criteria instantly turned their self-image of a remediable user into that of an irreparable 'junkie'.

The instruments for measuring the participants' progress also offer resources to illustrate the issue.⁹ They encompassed categories such as body weight, blood pressure, the condition of the skin, personal hygiene, social contacts inside or outside the drug scene, social responsibilities, and illegal activities such as stealing, begging, or urinating in the streets. I do not want to suggest that these criteria, and the further specifications for measuring them, are arbitrary or mistaken. The point is that medical and social categories are based on conventions. Nature does not unambiguously prescribe the criteria of health or indicate precisely which body weight, blood pressure, or skin condition is the right one. As far as such phenomena have been made quantitative and internationally standardized, much labour and fine-tuning has gone into the process.¹⁰ Likewise, nature does not tell us

exactly which psychological make-up or social behaviour is desirable. Much deliberation has been spent in selecting standards of mental health and, subsequently, on choosing the test items that further define such standards.

Most importantly, these classifications may have forceful consequences. When participants in the experiment took the tests they knew which responses would best serve their interests. The early architects of the heroin experiment had already pointed out that the prospect of later heroin-maintenance would affect the responses of the controls (GG&GD Amsterdam *et al.*, 1996). This argument gains strength from the fact that the controls knew they would only get heroin after six or twelve months if their condition was still bad. Moreover, the participants in the heroin group knew that their participation would be ended if their condition declined (CCBH, 1997b). Of course, this problem occurred because the heroin-maintenance experiment could not be conducted double blind. However, even without such inevitable departures from the really ideal RCT, the classifications used in the experiment would co-determine the outcomes. If, later on, the conclusion of the experiment is that heroin-maintenance helps (or does not help), some miserable lawbreakers will (or will not, respectively) be turned into nurtured patients. Either of the possible outcomes will be as much the result of the classifications used in the experiment as of the heroin that was provided.

UNEXPECTED RESPONSES

In November 1998, the CCBH could report that during the pilot study in Amsterdam and Rotterdam only negligible problems occurred. A 'National Committee Aspects of Command' monitored the consequences for public order, and a 'National Safety Committee' controlled for unwanted medical side-effects. Both committees noted that everything passed off quietly. That, however, does not mean that the pilot study also proceeded fully as expected. As journalists already had reported, the Amsterdam and Rotterdam users not merely refrained from beleaguering the maintenance stations but, quite the reverse, proved surprisingly reluctant to be enrolled. In addition, soon after the first participants had received their heroin the rumour spread that there was something wrong with this dope. Although the government heroin was chemically purer than street heroin, some participants commented that it did not 'taste good' or even was 'nasty rotgut' (newspaper *De Volkskrant*, 7 February 1998, 25 November 1998). Among the minor incidents reported by the Committee Aspects of Command was that of one participant trying to smuggle heroin out of the maintenance station in order to have it tested in a laboratory of his own choice (Committee Aspects of Command, 1998). Some participants even quit the project because of the inferiority of the heroin. The pilot study was conducted with 141 participants instead of the planned 150,

of whom 45 instead of 50 were assigned to the experimental group (CCBH, 1998).

Parliament once more discussed the heroin experiment in February 1999. The former worry that too many heroin enthusiasts would rush on board was now replaced by the fear that not enough participants could be attracted. This time weighing her answers most carefully, the minister argued that the number of 750 participants was not as sacred as it might have seemed. The principal concern of the project, she explained, was the effect of the twelve-month program. Therefore, the six-month heroin group was actually not all that important. If the need arose, the six-month experiment could be conducted with only the 50% heroin inhalers as injecting heroin had lately become less popular. The required number of participants would then amount to 625. In order to deal with the issue of the taste, the proportion of caffeine mixed with the heroin would be reduced (Tweede Kamer, 1999: 2 and 3).

No MP recalled that the minister once insisted on 656 participants as an absolute minimum dictated by statistics. No one suggested that if the lack of injectors was the only problem inhalers might replace them. The danger of assaults on the maintenance stations having been allayed, the liberals' minds were set at rest, and the minister finally gained a majority in the House (Tweede Kamer, 1999). When the full project took off in April 2000 in six Dutch cities, the CCBH indeed limited the minimum number of participants to 625 (Van den Brink *et al.*, 2000). That, however, did not solve the problem. Soon after the start of the experiment the city councils collectively alerted the minister. A considerable sum of extra money was needed because of the slow enrolment of participants and the consequent longer duration of the experiment (Groningen City Council, 2000).

Why had the heroin experiment to be conducted with a sample that was too small according to the original criteria? And why did the Rotterdam Junkie Union insist that 'the flash is different' although the heroin-caffeine proportion had been changed (Nora Storm, telephone interview, 23 August 2000)? Human science experiments more often struggle with recruitment problems. In this case, however, there seemed to be every reason for confidence. To recall, the number of users who took heroin without much improvement was estimated at 8,000. Only some 8.2 to 9.4% of them were needed for the experiment. And isn't it quite a treat indeed (as the liberals argued) to receive something of the very best quality for free? Moreover, in this particular experiment, even the participants with the bad luck of being allocated to the control group knew they would get the rewarding treatment. And last but not least, this was an experiment with a drug to which the envisaged participants were addicted.

Again, it is not my intention to argue that the Dutch heroin experiment did not proceed according to the established standards, but rather to question

these standards themselves. As has been discussed in the historical section of this article, it is a maxim of liberalism embodied in RCTs that social problems ought to be fought at the individual behavioural level rather than that of social change. The primary means to fight poverty, crime and deviation are rewards, punishments, education and medicalization. It has also been argued that it is a maxim of liberalism embodied in RCTs that inter-individual factors are unimportant. Only upon that assumption does it make sense to investigate people in artificial groups and settings. Recognizing such basic principles of the RCT might help to explain why it proved hard to attract participants for the heroin experiment.

In order to acquire their daily dose, participants had to attend a maintenance station three times a day, seven days a week. They had to report at a steel safety door where a camera recorded them. Next, they passed a metal detector and a porter behind a glass screen. Behind another door, they collected their heroin through a box-office. To make sure that no heroin was smuggled out for selling purposes, a nurse behind a screen watched the participants unremittingly while they consumed their established portion. Consumption did not take place with friends but in a randomly composed group. Once a week, participants were randomly selected for a urine test in order to control for additional use of drugs. Women were tested for pregnancy every month. Furthermore, all participants had a thorough physical and psychological examination every other month. Doctors controlled their health and hygiene. They had to open their mouth for a check of their teeth. Interviewers asked, over and over again, about their contacts outside the drug scene, their sources of income, and even the spots where they urinated.

Little wonder that the flash of the government's heroin was different. Which wine buff would enjoy the finest glass if it were handed out through a window in a maintenance station? It would probably taste like nasty rotgut. Few people would take pleasure in a free gourmet dinner followed by intrusive examinations and interviews urging them to break with the gastronomic community. The expectation that the sheer offer of free heroin would suffice to attract numbers of compliant participants (and particularly the fear that heroin users would besiege the maintenance stations for it) expressed a vision of addiction as an entirely physical, acultural affair. For this reason, a number of drug experts have contested the medical or pharmacological model of substance use. Acknowledging that drugs cause physical changes, they have emphasized the view that neophytes learn to interpret the *meaning* of these changes in historically and culturally established social settings. Stated differently, such experts argued that using drugs is also a way of living. It is part of an alternative culture, which is at least as important as the substance itself (Zinberg, 1984; Peele, 1985; Reinarman and Levine, 1997).¹¹

The drug-maintenance expert, John Kaplan, contended that the obligation of consumption in an institutional setting, as well as the role of a patient, will

keep many users from participation in a maintenance project (Kaplan, 1983: 168–75). Richard Hartnoll, who was involved in earlier heroin-maintenance experiments, noted that only a minority of the target group of problematic users wanted to be ‘registered, recognized’ heroin users (Hartnoll, 1993: 41). And Nora Storm, spokeswoman of the Rotterdam Junkie Union, used the illuminating *contradictio in terminis* of ‘Yuppie-Junkies’ in order to indicate that only those users who were capable and willing to adapt to its rules participated in the experiment (telephone interview, 23 August 2000).

Briefly, participants in the experiment had to quit their normal life and to adhere to prescriptions of regular society. Since the 1950s, theorists of society have analysed the use of the medical paradigm as an alternative means of (re)integrating people who cannot be turned into accountable individuals to a satisfactory degree. Anyone categorized as a ‘patient’ is discharged of full responsibility but, in exchange, has to submit to rules prescribed by others, and to refrain from remaining an outsider (Parsons, 1952; Goffman, 1963; Freidson, 1972; Foucault, 1975). Taking this into account, the expression ‘heroin on prescription’ was indeed much more adequate than ‘free heroin’. The heroin provided by the government was a ‘Dutch treat’ for which the recipients had to pay in terms of adhering to the mores of other people’s culture. Not all of the invited heroin users were prepared to do so.

EPILOGUE: SOME ADVANTAGES OF SOLID ETHNOGRAPHY

In the USA where the drug problem is much more vast and grim than anywhere in the Netherlands, fears of drug-related diseases such as HIV and AIDS have instigated an increasing interest in ‘hidden and marginalized populations’ that cannot be studied with conventional methods. For instance, the American National Institute of Drug Abuse (NIDA), which traditionally focused on quantitative methods, sponsored the publication of a volume *Qualitative Methods in Drug Abuse and HIV Research*. The editors of this volume argued that currently there were not enough social scientists ‘with the requisite methodological expertise to study complex human behaviours related to drug abuse and HIV prevention’. New ethnographical methods needed to be developed to gain understanding of the meanings that drug users ascribe to their way of life, and of the cultural context of abuse (Lambert *et al.*, 1995: 4).

The work of Philippe Bourgois, an American researcher, is of particular interest. He profitably combined old- with new-style ethnography. Bourgois went so far as to move into Harlem (New York) for three years and, some years later, to immerse himself in the San Francisco heroin community. In this way he also met users who do not attend clinics or maintenance stations.

Moreover, he studied natural groups rather than artificial ones. In order to avoid steering responses as much as possible, Bourgois refrained from using pre-structured surveys. Nor did he pay for co-operation with money or drugs because this easily leads to strategic reactions. He preferred not to interrogate, measure, or test, but to observe and interpret. Hanging out with users and dealers on the streets, in their homes, in crackhouses, and shooting camps, he learned the language and gained the trust of the drug community. Subsequently, by virtue of his acquired social bilingualism, he could act as an interpreter to the regular world. Without acquitting or glorifying drug offenders, Bourgois raised thought-provoking issues for traditional technological research and treatments. Most importantly, he combined his views on the drug culture with macro-analyses of social power relations. Generally speaking, his argument is that treatments based on diagnoses of physical or psychological causes lead to barely effective symptom therapies for structural rather than individual problems (Bourgois, 1995, 1999, 2000; Bourgois *et al.*, 1997).

To be sure, the results of ethnographic research are not free of the researchers' views. However, the researchers do not claim that the results are thus free; they are meant rather to instigate debate. And ethnographical research cannot proceed without classifications either, but the difference between it and the experimental paradigm is that in solid ethnography the effects of classification are openly part of the argument. Ethnographers also stress that their findings are of culturally limited value.¹² This implies that conclusions based on studies in the USA are not straightforwardly applicable to the Netherlands. For instance, in Esping-Andersen's terms, the latter country is not a liberal welfare state but represents a mixture of corporatist and social democrat regimes (see also Goodin *et al.*, 1999). This has important consequences for the nature of its drug problems (Cohen, 1997).¹³ Therefore, studies of foreign cultures are important only as examples of the possibilities of sound ethnography.

These possibilities also include questions concerning the effects of heroin-maintenance. Whereas in an RCT the differences between kinds of user need to be cancelled out, ethnographic studies can take them as a focal point of research. The Dutch researchers Otto Janssen and Koert Swierstra conducted an ethnographic study of different 'life-styles' of groups of heroin users (Janssen and Swierstra, 1982), and next presented some preliminary conclusions on the differential effects of heroin-maintenance on these varied subcultures (Janssen and Swierstra, 1983: 24–6). This work did not have much influence, nor was it continued on a large scale.¹⁴ The prevailing experimental definition of science prescribes that most of the brainpower and resources for research are spent on plugging the gaps and cracks through which interpretations ('distortions') might slip in. I have argued that this will always be in vain, and, moreover, may debar indispensable understanding.

NOTES

I would like to express my gratitude to Ad Prins, Ted Porter and Roger Smith for their helpful comments, as well as the reviewers for *History of the Human Sciences*. I am particularly indebted to Gijs Meijer, who wrote a Masters' thesis on the Dutch heroin experiment in 1998 (unpublished), for his useful views and for sharing his archival material. An earlier and less extensive version of this article appeared in the *Maandblad voor Geestelijke Volksgezondheid* [Monthly for Mental Health] under the title 'Over de assumpties van RCTs, geïllustreerd aan het Nederlandse experiment met verstrekking van heroïne' [On the Assumptions of RCTs, Illustrated by the Dutch Experiment with Heroin-maintenance].

- 1 For an overview of the discussion in the Netherlands, see Derks (1990). For information on English language publications on heroin-maintenance, see <http://www.lindesmith.org/> and <http://www.druglibrary.org/schaffer/>
- 2 D66 stands for Democrats '66, a party founded in 1966 as a compromise between economic liberalism and labour. Henceforth, I will refer to D66 as 'the social-liberal party'. The other parties mentioned in this article are the right-wing liberal party (VVD), the Labour Party (PvdA), the left-wing Green Party (Groen Links), and the Christian Democrats (CDA). Since 1994, the Cabinet has consisted of a coalition of right-wing liberals, social liberals and labour.
- 3 See Merriam-Webster's Dictionary: <http://www.m-w.com/cgi-bin/dictionary/>
- 4 For an historical analysis of the emergence of the RCT in American medicine see Marks (1997); in psychology and social science see Dehue (1997, 2000, 2001) and Danziger (2000).
- 5 In defense of the minister: she said this in response to MPs' sudden demand that the effects on criminality be of primary concern.
- 6 The expression 'social experimentation' already appears in the writings of 19th-century scholars such as Auguste Comte, John Stuart Mill and George Cornewall Lewis. On the difference between its early and later meaning see Dehue (2001).
- 7 Other statisticians, such as William Gosset, objected that it is more informative to imagine which factors might cause bias rather than just cancel them out. Fisher, however, maintained that it is safer to trust an automatic elimination of unwanted differences between groups than to rely on personal judgement. The statistician Egon Pearson commented on the issue: 'Few practicing statisticians have the ability and experience of a Gosset, Fisher, or Yates, and therefore to safeguard against blunders it is no doubt important to teach would-be applied statisticians the value of randomization where it is possible in design of experiments. But that does not prove that Fisher was "right" and Gosset "wrong"' (Pearson, 1990: 109).
- 8 In the Netherlands, growing or selling soft drugs (marijuana and hashish) is illegal. However, possession of less than 5 grams is not punished and licensed coffee shops are not penalized for selling up to 5 grams to adults (Ministry of Justice: <http://www.minjust.nl>).
- 9 The experiment used the internationally standardized Addiction Severity Index (ASI), European version, as well as the Composite International Diagnostic Interview (CIDI) issued by the World Health Organization, and the Symptom Checklist (SCL-90) designed in the Netherlands.

- 10 For instance, Beevers, Lip and O'Brien (2001) discussed various difficulties in standardizing the means of blood-pressure measurement, and the rise of blood pressure in some people by the very act of measurement ('white coat hypertension').
- 11 The Centre for Drug Research (CEDRO) in Amsterdam also criticizes the medical model in drug policy (see <http://www.cedro-uva.org/lib/>). Van de Port (1998) also combined the medical-technical assumptions of the Dutch heroin experiment with the recruitment problems during the pilot study.
- 12 For qualitative drug research in European countries, see the website of the Qualitative European Drug Research Network (<http://qed.emcdda.org/>).
- 13 Contrasting the Netherlands to the USA, the Amsterdam drug researcher Peter Cohen (1997) argued that in the Netherlands the drug problem is comparatively small. He ascribed this to the fact that it has always been considered as an indication of shortcomings in the system of social care rather than a problem of particular individuals. However, in a Dutch-language article, he fulminated against the government coalition of labour, right-wing liberals and social liberals (nicknamed the 'Purple Cabinet'). Cohen's scorn concerned this Cabinet's 1995 memorandum on drug policy, which also first presented the plan of the heroin-maintenance experiment. He cynically asserted that 'Purple' was 'rediscovering the whip as a magic panacea'. The 1995 policy intentions on drugs, he maintained, focus only on the users themselves (Cohen, 1995).
- 14 Other Dutch ethnographic studies by these authors are Swierstra *et al.* (1986), and Swierstra (1990).

BIBLIOGRAPHY

- Bannister, R. C. (1987) *Sociology and Scientism: The American Quest for Objectivity, 1880-1940*. Chapel Hill, NC: University of North Carolina Press.
- Beems, D. (1995) 'Gratis Heroïne, Wilt u Dat Nooit Meer Zeggen?' [Free Heroin, Would You Never Say That Again?], *Binnenlands Bestuur* 23-4: 14-15.
- Beevers, G., Lip, Y. H., and O'Brien, E. (2001) 'Blood Pressure Measurement. Part 1: Sphygmomanometry: Factors Common to All Techniques', *British Medical Journal* 322: 981-5.
- Benschop, R. J. and Draaisma, D. (2000) 'In Pursuit of Precision. The Calibration of Minds and Machines in Late Nineteenth-Century Psychology', *Annals of Science* 57: 1-25.
- Boruch, R. (1997) *Randomized Experiments for Planning and Evaluation*. London: Sage.
- Bourgeois, P. (1995) *In Search of Respect: Selling Crack in El Barrio*. New York: Cambridge University Press.
- Bourgeois, P. (1999) 'Theory, Method, and Power in Drug and HIV-Prevention Research. A Participant-Observer's Critique', *Substance Use and Misuse* 34: 2155-72.
- Bourgeois, Philippe (2000) 'Disciplining Addictions: the Biopolitics of Methadone and Heroin in the United States', *Culture, Medicine and Psychiatry* 24: 165-95.

- Bourgeois, P., Lettiere, M. and Quesada, J. (1997) 'Social Misery and the Sanctions of Substance Abuse: Confronting HIV Risk among Homeless Heroin Addicts in San Francisco', *Social Problems* 44: 155–73.
- Bowker, G. C. and Star, S. L. (1999) *Sorting Things Out. Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Brown, M. K. (1999) *Race, Money, and the American Welfare State*. Ithaca, NY: Cornell University Press.
- Bulmer, M. (1986) 'Evaluation Research and Social Experimentation', in M. Bulmer, K. G. Banting, M. Carley and C. H. Weiss (eds) *Social Science and Social Policy*. London: Allen & Unwin, pp. 155–79.
- Callahan, R. E. (1962) *Education and the Cult of Efficiency*. Chicago, IL: University of Chicago Press.
- CCBH (1997a) 'Onderzoek naar Heroïne op Medisch Voorschrift. I Het Onderzoek. A. Referentie document' [Research on Heroin on Medical Prescription. I: The Research Project. A: Reference Document]. Unpublished report. Utrecht: CCBH.
- CCBH (1997b) 'Onderzoek naar Heroïne op Medisch Voorschrift. I Het Onderzoek. D. Deelnemersinformatie' [Research on Heroin on Medical Prescription. I: The Research Project. D: Information for Participants]. Unpublished report. Utrecht: CCBH.
- CCBH (1998) 'Voortgangsrapportage CCBH' [Progress Report CCBH]. Unpublished report. Utrecht: CCBH.
- Chapin, F. S. (1917a) 'The Experimental Method and Sociology I. The Theory and Practice of the Experimental Method', *Scientific Monthly* 4: 133–44.
- Chapin, F. S. (1917b) 'The Experimental Method and Sociology II. Social Legislation is Social Experimentation', *Scientific Monthly* 4: 238–47.
- Chapin, F. S. (1938) 'Design for Social Experiments', *American Sociological Review* 3: 786–800.
- Chapin, F. S. (1947) *Experimental Designs in Social Research*. New York: Harper.
- Cohen, P. (1995) 'Paars ontdekt de knoet, het nieuwe wondermiddel' ['Purple Discovers the Whip, the New Magic Panacea'], *Vrij Nederland*, 21 October 1995.
- Cohen, P. (1997) 'Crack in the Netherlands. Effective Social Policy is Effective Drug Policy', in C. Reinerman and H. G. Levine (eds) *Crack in America. Demon Drugs and Social Justice*. Berkeley: University of California Press, pp. 214–25.
- Committee Aspects of Command (1998) 'Onderzoek naar Heroïne op Medisch Voorschrift: Rapportage Landelijke Commissie Beheersaspecten' [Research on Heroin on Medical Prescription. Report of the National Committee Aspects of Command]. Unpublished report. Utrecht: CCBH.
- Danziger, K. (1990) *Constructing the Subject*. New York: Cambridge University Press.
- Danziger, K. (1997) *Naming the Mind. How Psychology Found its Language*. London: Sage.
- Danziger, K. (2000) 'Making Social Psychology Experimental: a Conceptual History, 1920–1970', *Journal of the History of the Behavioral Sciences* 26: 329–49.
- Dehue, T. (1997) 'Deception, Efficiency, and Random Groups', *Isis* 88: 653–73.
- Dehue, T. (2000) 'From Deception Trials to Control Reagents. The Introduction of the Control Group about a Century Ago', *American Psychologist* 55: 264–9.

- Dehue, T. (2001) 'Establishing the Experimenting Society', *American Journal of Psychology* 114: 283–302.
- Derks, J. (1990) *Het Amsterdamse Morfine-verstrekingsproject* [The Amsterdam Morphine-maintenance Project]. Utrecht: NcGv.
- Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Esping-Andersen, G. (1999) *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Fisher, R. (1935) *The Design of Experiments*. Edinburgh: Oliver & Boyd.
- Foucault, M. (1975) *Surveiller et Punir*. Paris: Gallimard.
- Freidson, E. (1972) *Profession of Medicine*. New York: Dodd.
- Geertz, C. (1983) *Local Knowledge: Further Essays in Interpretive Anthropology*. New York: Basic Books.
- GG&GD Amsterdam, GG&GD Den Haag, GG&GD Rotterdam, GG&GD Utrecht, GG&GD Regio Stedendriehoek (1996) 'Basisdocument voor een Onderzoek naar de Effecten van Gecontroleerde Verstrekking van Heroïne in Nederland' [Preliminary Document on a Research Project on the Effects of Controlled Heroin Maintenance in the Netherlands]. Unpublished report.
- Goffman, E. (1963) *Stigma*. Englewood Cliffs, NJ: Prentice-Hall.
- Gomart, E. and Hennion, A. (1999) 'A Sociology of Attachment: Music Amateurs, Drug Users', in J. Law and J. Hassard (eds) *Actor Network Theory and After*. London: Blackwell, pp. 220–48.
- Goodin, R. E., Bruce Headey, R. M. and Henk-Jan, D. (1999) *The Real World of Welfare Capitalism*. New York: Cambridge University Press.
- Groningen City Council (2000) 'Voortgangsrapportage Onderzoek naar Gecontroleerde Heroïneverstrekking in Groningen' [Progress Report Research Project Controlled Heroin-maintenance in Groningen]. Unpublished report. Groningen.
- Hacking, I. (1986) 'Making up People', in T. C. Heller (ed.) *Reconstructing Individualism*. Stanford, CA: Stanford University Press, pp. 222–36.
- Hacking, I. (1999) *The Social Construction of What?* Cambridge, MA: Harvard University Press.
- Hartnoll, R. L. (1993) 'Heroin Maintenance and AIDS Prevention. Going the Whole Way?', *International Journal of Drug Policy* 4: 36–41.
- Haskell, T. (1977) *The Emergence of Professional Social Science: The American Social Science Association and the Nineteenth-Century Crisis of Authority*. Urbana, IL: University of Illinois Press.
- Haveman, R. H. (1987) 'Social Experimentation and "Social Experimentation"', in W. R. Shadish and C. S. Reichardt (eds) *Evaluation Studies*, Vol. 1. London: Sage, pp. 608–27.
- Janssen, O. and Swierstra, Koert E. (1982) *Heroïnegebruikers in Nederland. Dl. I Een typologie van levensstijlen* [Heroin Users in the Netherlands. Part I: A Typology of Life-styles]. Groningen: Rijksuniversiteit Groningen.
- Janssen, O. and Swierstra, Koert E. (1983) *Uitgangspunten voor een Integraal Drugsbeleid* [Starting Points for an Integrated Drug Policy]. Groningen: Rijksuniversiteit Groningen.
- Jansson, B. (1993) *The Reluctant Welfare State. A History of American Social Welfare Policies*. Pacific Grove: Brooks/Cole.

- Jevons, W. S. (1880) 'Experimental Legislation and the Drink Traffic', *Contemporary Review* 7: 177–92.
- Jevons, W. S. (1882) *The State in Relation to Labor*. London: Macmillan.
- Kaplan, J. (1983) *The Hardest Drug. Heroin and Public Policy*. Chicago, IL: University of Chicago Press.
- Lagemann, E. C. (2000) *An Elusive Science. The Troubling History of Education Research*. Chicago, IL: University of Chicago Press.
- Lambert, E. Y., Ashery, R. S and Needle, R. H., eds (1995) *Qualitative Methods in Drug Abuse and HIV Research*. NIDA Monograph 157. Washington, DC: US Government Printing Office.
- Lovie, A. D. (1979) 'The Analysis of Variance in Experimental Psychology: 1934–1945', *British Journal of Mathematical and Statistical Psychology* 32: 151–78.
- Lyons, G. M. (1969) *The Uneasy Partnership: Social Science and the Federal Government in the Twentieth Century*. New York: Russell Sage Foundation.
- McCall, W. A. (1923) *How to Experiment in Education*. New York: Macmillan.
- Marks, H. M. (1997) *The Progress of Experiment. Science and Therapeutic Reform in the United States, 1900–1990*. New York: Cambridge University Press.
- Orr, L. L. (1999) *Social Experiments. Evaluating Public Programs with Experimental Methods*. London: Sage.
- Osborne, T. and Rose, N. (1999) 'Do the Social Sciences Create Phenomena? The Example of Public Opinion Research', *British Journal of Sociology* 50: 367–96.
- Parsons, T. (1952) *The Social System*. London: Routledge.
- Pearson, E. (1990) *'Student': A Statistical Biography of William Sealy Gosset*. Oxford: Clarendon Press.
- Peele, S. (1985) *The Meaning of Addiction. Compulsive Experience and its Interpretation*. Lexington, MA: Lexington Books.
- Peele, S. (1990) 'Addiction as a Cultural Concept', *Annals of the New York Academy of the Sciences* 602: 205–20.
- Porter, T. M. (1993) 'Statistics and the Politics of Objectivity', *Revue de Synthèse* 4: 87–101.
- Porter, T. M. (1995) *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Reinarman, C. and Levine, H. G., eds (1997) *Crack in America. Demon Drugs and Social Justice*. Berkeley: University of California Press.
- Rose, N. (1996) *Inventing our Selves. Psychology, Power and Personhood*. Cambridge, MA: Cambridge University Press.
- Ross, Dorothy (1991) *The Origins of American Social Science*. New York: Cambridge University Press.
- Rossi, P. H. and Lyall, K. C., eds (1976) *Reforming Public Welfare: A Critique of the Negative Income Tax Experiment*. New York: Russell Sage Foundation.
- Rucci, A. J. and Tweney, R. D. (1980) 'Analysis of Variance and the "Second Discipline of Scientific Psychology". A Historical Account', *Psychological Bulletin* 87: 166–84.
- Schellings, Ron, Alphons, G. H., Kessels, Roy Geurs and Sturmans, Ferdinand (1996) 'Overwegingen bij de Opzet van een Gerandomiseerd Heroïne-Verstrekkingsexperiment' [Deliberations Concerning the Design of a Randomized Heroin-Maintenance Experiment], *Tijdschrift voor Sociale Gezondheidszorg [Journal of Social Health Care]* 74: 376–82.

- Stone, Deborah (1998) *Policy Paradox. The Art of Political Decision Making*. New York: Norton.
- Swierstra, Koert E. (1990) *Drugscarrières [Drug Careers]*. PhD thesis. University of Groningen.
- Swierstra, Koert E. and Janssen, Otto (1986) *Heroïnegebruikers in Nederland. Dl. II: De Reproductie van Heroïnegebruik onder Nieuwe Lichtingen [Heroin Users in the Netherlands. Part II: The Reproduction of Heroin Use among New Shifts]*. Groningen: Rijksuniversiteit Groningen.
- Tweede Kamer [House of Parliament] (1996) *Drugsbeleid. Verslag van een Algemene Vergadering [Drug Policy. Report of a General Meeting]* 24-077, number 35. The Hague: SDU.
- Tweede Kamer der Staten Generaal [House of Parliament] (1997) *Drugsbeleid. Verslag van een Algemeen Overleg [Drug Policy. Report of General Consultations]* 24-077, number 60. The Hague: SDU.
- Tweede Kamer der Staten Generaal [House of Parliament] (1999) *Drugsbeleid. Verslag van een Schriftelijk Overleg [Drug Policy. Report of Written Consultations]* 24-077, number 70. The Hague: SDU.
- Van de Port, Pepijn (1998) 'Voor Oriëntalen een Stimulant. Een Antropologische Kritiek op het Medisch Verslavingsparadigma' [For Orientals a Stimulant. An Anthropological Critique on the Medical Addiction Paradigm], *Psychologie & Maatschappij [Psychology & Society]* 22: 364–72.
- Van den Brink, Wouter, Hendriks, Vincent M., Blanken, Peter and van Ree, Jan M. (2000) 'Het Nederlandse onderzoek naar de Effectiviteit van Heroïne op Medisch Voorschrift; Achtergrond, Onderzoeksopzet en Eerste Ervaringen' [The Dutch Research Project Investigating the Effectiveness of Heroin on Medical Prescription; Background, Design and First Experiences], *Nederlands Tijdschrift voor de Geneeskunde [Netherlands Medical Journal]* 144: 108–12.
- Williams, Walter (1971) *Social Policy Research and Analysis. The Experience in the Federal Social Agencies*. New York: Elsevier.
- Yoshioka, Alan (1998) 'Use of Randomization in the Medical Research Council's Clinical Trial of Streptomycin in Pulmonary Tuberculosis in the 1940s', *British Medical Journal* 317: 1220–3.
- Zinberg, Norman E. (1984) *Drug, Set and Setting. The Basis for Controlled Intoxicant Use*. New Haven, CT: Yale University Press.

BIOGRAPHICAL NOTE

TRUDY DEHUE is Professor of Theory and History of Psychology at the University of Groningen, the Netherlands. Her publications include *Changing the Rules* (Cambridge University Press, 1997) and articles on the history of methodology in the human sciences.

Address: Heymans Institute of Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, the Netherlands. Tel: 31 50 363 63 40. Fax: 31 50 363 63 04. [email: t.dehue@ppsw.rug.nl]