Bulletin of the History of Medicine 72.3 (1998) 389-433

Intentional Ignorance: A History of Blind Assessment and Placebo Controls in Medicine

Ted J. Kaptchuk *

Introduction

Blind assessment is considered a critical component of modern medical research methodology. Nonetheless, attention to this practice in the historiography of human experimentation has varied from nil to minimal, occupying at most a few short paragraphs in a relatively few articles and books. Such brief, perfunctory accounts as do exist portray it as having been adopted after World War II following, at most, a few precursors. In addition, the details of these accounts are surprisingly inconsistent, and the rationale for blind assessment is presented as self-evident and devoid of social or historical dimensions. [1] The history of masked **[End Page 389]** assessment seems veiled in obscurity, with the implication that this method was not available until well into the twentieth century, when an eternal transhistorical scientific verity somehow became obvious to researchers. The aura of objectivity and neutrality attached to blind assessment itself may have benefited from this absence of a past.

In fact, however, blind assessment has been a continuous and complex scientific and social enterprise for more than two hundred years. The testing of human subjects under conditions of intentional ignorance has taken numerous forms. The simplest method was to use blindfolds or curtains so that the patient (and/or the experimenter) was unaware of the exact nature or timing of the intervention. The veil helped to eliminate the threat of imagination and bias, and sometimes it served as insurance against fraud and trickery. From its very inception, blind investigation also utilized a decoy or dummy intervention (such as a placebo or sham device) that allowed researchers to observe the effects of the appearance of intervention. [2] Blind assessment could also be "double-blind" **[End Page 390]** (to use the modern phrase), so that both the patient and the experimenter were unaware of whether the treatment was ritual alone or a ritual that included the experimental therapy. In any of its forms, and in all of them taken together, blind investigation helped medical science isolate "hard" knowledge and material causality

from the contamination of mental delusion, enthusiastic bias, or even calculated deceit.

This essay examines the unacknowledged saga of a research methodology that placed a higher value on information derived from people who were kept "ignorant" during an assessment. It concerns the study of living human beings. Most of the story is medical, but it also touches related research with human subjects in psychology and pharmacology, research that needed the guarantee of concealment. Its leitmotif is noble conflict and intellectual doubt. Its intent is to wrest certainty out of the shadows of human awareness. The test is an ordeal of darkness. The unspoken countertheme of human passions is revealed in the details of circumstances and timing: When was it critical to distinguish material causality from its mere appearance? When was fraud or *suggestion* through other sensory or mental pathways a lurking threat? When were seeing and feeling not to be believed? When did objective truth need to be protected from the contamination of the mind?

I will trace the development of blind assessment by examining the most representative episodes of its use up until its widespread acceptance in biomedical research. The story can be roughly divided into five phases. Blind assessment began in the late eighteenth century as a tool for detecting fraud in a campaign mounted by elite mainstream scientists and physicians to challenge the suspected delusions or charlatanism of unconventional medicine. It demarcated orthodox medicine from what was considered deviant healing. In the second phase, beginning in the mid-nineteenth century, blind assessment became a research tool within various medical communities. Often this utilization was a defensive adaptation by proponents of irregular healing, but some conventional physicians in other situations also employed it for polemical medical display. The third phase began in the late nineteenth century when experimental psychologists tried to separate the private and unquantifiable components of the mind from the objective and quantifiable components of sensation and perception. Psychical researchers sought to bolster shaky scientific identity with concealed experimentation. Most critically, blind assessment also became a decisive vehicle by which neurologists and psychiatrists could demarcate the newly stumbled-upon semilegitimate domain of *suggestion* from that of material causality. The rationale of this phase eventually penetrated aspects of pharmacology research performed by physiologists and psychologists, creating a distinct fourth phase. **[End Page 391]**

Blind assessment's most recent phase began in the 1930s when researchers perceived its value in designing no-treatment control groups in clinical trials. When blind assessment was later understood to be a valuable partner of the new randomized controlled trial (RCT) methodology introduced after World War II, its final triumph was guaranteed. With great speed, blind assessment and placebo controls became a moral imperative for an emerging clinical research agenda. The threat of contaminated evidence, unexamined bias, and subjective perceptions, originally thought to cling only to deviant practices, suddenly was internalized and introjected into authoritative scientific medicine itself. Unblinded evidence, whether presented by mainstream physician or by unconventional healer, was now considered suspiciously, with the intrinsic value of a folktale or an anecdote. In the decades after World War II, the "new method" of blind assessment, as an integral component of the RCT methodology, became a critical, routinized, self-evident, and normative procedure for the scientific assessment of

efficacy. [3] Clinical medicine's scientific and rhetorical claims to certainty now included assessment "in the dark." The "newness" of the method (despite its earlier availability and its long, deliberate neglect by most researchers) allowed blind assessment to appear to be a simple improvement for the use of any conscientious researcher.

In addition to the historical account, this essay also provides an overlooked example of the social dimensions of research activity and of the way in which the process of determining legitimacy often "involves [unexamined] prior agreements about what is to count as admissible evidence." [4] The history of "intentional ignorance" in research illustrates the formidable social considerations and a priori assumptions involved in what John Harley Warner has described as the constant redefinition of science in medicine. [5] The adoption of masked assessment in human research not only was an example of medicine "becoming more scientific," but also confirms that "one cannot distinguish purely technical **[End Page 392]** aspects of ideas from their role as political strategies in the competition for resources. . . . Ideas are judged not only for their truth value but also for their utility in discipline building." [6] At each point in its development, blind assessment was the product of a wide variety of social forces contending for scientific truth. But this search was also a struggle for moral authority, persuasive rhetoric, and the power that legitimacy confers.

Finally, this inquiry traces a trail of overlooked activity in human research through two centuries of history. Each episode deserves a tome: each act had the drama of winner and vanquished; each defeat was avenged. But sensitivity to each scene has here been sacrificed in order to document the continuity and history of this method. The reader's indulgence is requested for this emphasis on breadth over depth.

Blind Assessment and Unorthodox Healing

Blind assessment first emerged in history as a deterrent against unconventional healers. Such healers were automatically treated with suspicion and distrust when their claims disregarded normative science. Understandably, orthodox physicians suspected poor judgment and illusion, bias and overenthusiasm, or even quackery and fraud. Blind assessment was a prominent feature in the medical tug-of-war with these "irregulars" and became the ultimate tool to test and demonstrate that these practitioners were "unscientific" and selling bogus goods. [7] Unconventional healers sometimes responded by adopting the method as a touchstone for proving their own claims of efficacy. In its early development, blind **[End Page 393]** assessment was not articulated into a systematic methodology, but was regularly dusted off and used to either combat potential quackery or gain access to scientific respectability. Especially in the initial phases, it tended to be used in an ad hoc manner to demarcate or proclaim legitimacy whenever the boundary conflict between conventional and deviant became especially contentious. [8] The three most important conflicts that utilized blind assessment were mesmerism, perkinism, and homeopathy.

Mesmerism

As far as I can ascertain, the first series of blind assessments and sham interventions for the purpose of scientific appraisal was aimed at mesmerism, the most popular and threatening unconventional healing system to appear in the late eighteenth century. [9] Franz Anton Mesmer (1734-1815) claimed to have discovered a new healing "fluid" in nature, **[End Page 394]** analogous to gravitation, which he called "animal magnetism." [10] As a result of the enormous attention and notoriety surrounding Mesmer's methods, Louis XVI appointed a commission of inquiry consisting of members of the Academy of Sciences and the Academy of Medicine. Benjamin Franklin (1706-90), American scientist and minister plenipotentiary, headed the distinguished commission of scientists and physicians. [11]

The commission limited its investigation to determining whether the purported effects of animal magnetism were due to any "real" force. [12] "It was the duty of the commissioners to confine themselves to arguments purely physical, that is, to the momentaneous [*sic*] effects of the fluid upon the animal frame, excluding from these effects all the illusions which might mix with them." [13] Were the observed effects of mesmerism due to the contamination of the mind (and "all the illusions that mix within them"), or were they present without the influence of human awareness? [14] Between March and June 1784, in a series of numerous minitrials on single subjects, the commissioners removed genuine "knowledge" of treatment. Either blindfolds or decoy procedures were adopted to enforce this disembodiment of the mind. **[End Page 395]**

The first blindfold experiment was performed at Benjamin Franklin's house. A series of women selected by the cooperating mesmerist as "good subjects" were physically blindfolded, with bandages (so that they "could no longer know anything respecting the conduct of the experiment" [15]), and asked to locate where the mesmeric energy was being directed. It was observed that

> while the woman was permitted to see the operation, she placed her sensations precisely in the part towards which it was directed; that on the other hand, when she did not see the operation, she placed them at hazard, and in parts very distant from those which were the object of magnetism. It was natural to conclude that these sensations, real or pretended, were determined by the imagination. [16]

In another series of experiments, women patients were deceived by the scientists into believing that they were receiving mesmerism from an adjoining room through a paper curtain over a door. The "knowledge" of intervention produced the sensations. When they received mesmeric treatment but were not told they were being mesmerized (they were supposedly waiting), nothing happened. Many other experiments were performed, and each test led to the same conclusion: blinding could eliminate the effects of mesmerism, and sham worked as well as "real" mesmerism. [17]

The short form of the conclusion was clear: "This agent, this fluid has no existence," and any effects were due to "imagination." [18] The perceived effect of mesmerism was a result

of illusions created by the human mind. Intentional ignorance allowed the commission to disentangle the delusion of perception from "real" effects. [19] **[End Page 396]**

Although it was well publicized, the Royal Commission's negative conclusion set back mesmerism's popularity only temporarily. Throughout the nineteenth century, the magnetic movement continued with oscillating periods of strength and weakness. Medical doctors and, more consequentially, lay practitioners continued to minister to the sick. Public exhibitions of "higher" mesmeric phenomena--such as the diagnosis of disease in unknown persons, reading material in sealed boxes, clairvoyance, and precognition--dramatically fueled public fascination. [20] Debunkers and advocates alike quickly adopted the new blind assessment method to prove their points of view, and it became intrinsic to the entire controversy surrounding the nineteenth-century medical and extramedical mesmeric movement. In cloistered academic laboratories and on stages before hundreds, magnetic healers and itinerant entertainers were challenged to cure, detect, or perform wondrous feats with practitioners and/or subjects blindfolded. A cottage industry of blind assessment developed. [21] **[End Page 397]**

Blind assessment using concealment or sham treatment became routine in the medical investigations of mesmeric healing. Again, both negative and positive outcomes were extensively reported in professional medical journals and the popular press. Both sides of the dispute adopted the strategy of blind assessment and argued that any evidence supporting the opponent could be attributed to imperfect or unfair experimental conditions or fraud.

Out of these many trials and performances, at least one particular blind assessment of mesmerism had an important direct influence on the next significant development in the history of the technique. Also, it demonstrates a good example of a not-uncommon positive outcome. This series of experiments on magnetic healing took place between October and December 1820, at the request of students at the Hôtel Dieu in Paris. [22] The physician-in-chief, H. M. Husson (1772-1853), invited a well-known mesmerist to treat a seventeen-year-old woman who was exhausted with menstrual troubles and constant vomiting and had been totally refractory to treatment for eight months. She began to improve from the first magnetic session. But Husson wanted to know if she could be magnetized without "being aware" of the procedure. Was the magnetism "real" or "imagined"? He adopted an unambiguous form of intentional ignorance: the mesmerist (now called the "magnetist") was secretly put into a black cabinet (*un cabinet noir*), which was then securely locked and kept separate from the main room with a thick partition. [23] (This could be called history's first black-box experiment!) After telling the patient that the magnetist might not arrive that day, Dr. Husson dropped a pair of scissors, which was the prearranged signal for the magnetist to "emit" his magnetic fluid. Three minutes later, her usual elapsed time, the patient fell into a somnambulistic trance.

A week later the highly skeptical Professor Joseph C. A. Récamier (1774-56), who later inherited René Laennec's professorship, asked to supervise a repetition of the experiment on his own terms. The prearranged signal was for Récamier to ask the patient whether "she digested **[End Page 398]** meat." [24] Three minutes after the signal,

the patient again entered a somnambulistic state. Later experiments were performed in which the experimenters made feigned signals as if the magnetism were about to begin as it had in earlier sessions (such as with the dropping of the scissors), but in these experiments the patient was not affected. The magnetist also was able, without the patient's knowledge, to induce her into a somnambulistic state at times when no treatment took place at the hospital. Husson left before the experimental phase was complete, and his successor dismissed the unconventional healer because of the scandalous nature of the activity. The patient's sickness then returned. Another magnetist was secretly brought into the hospital, and she was discharged "in a rather satisfying state" on 20 January 1821. [25] Armand Trousseau (1801-67), who initiated the use of blind assessment in homeopathy at the same hospital thirteen years later (see below), was Récamier's most famous student. [26]

Perkinism

The second example of a legitimacy dispute was a minor skirmish compared to the earlier conflict surrounding mesmerism and the later conflict surrounding homeopathy. Nonetheless, the episode adds completeness to the story. It began not long after Franklin's original debunking of mesmerism, when a Connecticut physician, Dr. Elisha Perkins (1741-99), invented a "tractor" containing two metal rods that he thought conducted accumulated pathogenic "electroid" fluid (related to galvanic electricity) away from the body. [27] In 1799, after the device's British debut, Dr. John Haygarth (1740-1827) and colleagues, explicitly inspired by the **[End Page 399]** French investigators, decided to use what would now be called a single-blind experiment using a sham device. Their methodology was to "prepare a pair of false, exactly to resemble the true, tractors. Let the secret be kept inviolable. . . . Let the efficacy of both be impartially tried." [28] Five patients at the General Hospital in Bath, and an additional ten patients who were treated by collaborators at the Bristol Infirmary, were significantly relieved of pain or paralysis by both the wooden and metal tractors. Blind assessment therefore sidelined the tractor to the status of treatment for "the most illiterate peasant . . . [or] even the wildest savage." [29]

Homeopathy

The next battle using the armament of blind assessment began in 1834 at the Hôtel Dieu under Armand Trousseau. One of Trousseau's medical concerns was that "some honorable men and friends seriously occupied themselves with this novelty [homeopathy]." [30] Devised in Germany by Samuel Hahnemann (1755-1843), homeopathy had quickly become the most controversial and serious challenge to regular medicine in the nineteenth century. Enormously popular, it became an integral and **[End Page 400]** prominent component of the health scene in Europe and the United States. [31]

Hahnemann espoused the belief that whatever the symptom-complex a substance caused in a healthy person, a disease with a similar symptom configuration could be cured by small amounts of the same substances. *Similia similibus curentur*--like cures

like. Another major proposition of homeopathy was that the more dilute a substance (if prepared by a series of shakings called "succussion"), the more "spiritual vital essence" was released and therefore the more potent a medicine was created: less became more. Eventually Hahnemann and his followers were treating people with material so diluted as not likely to contain even a molecule of substance. Remedies had become dematerialized spiritual forces. Despite mainstream antagonism, Hahnemann attracted many disciples and followers.

Most of the orthodox professional elite refused to "debase [themselves] . . . by attempting to determine experimentally such enormous absurdities." [32] Some prominent medical leaders, however, actually undertook experimental administration--that is, testing by the then-customary procedure of giving the remedies in an open-label manner and seeing what happened. [33] (There was little idea of control groups in medical research at the time.) A group of Trousseau's students urged him to become involved in the homeopathic debates and together adopted the more rigorous strategy of blind assessment using placebo therapy. As one of the students said, the climate of "mistrust [and] the disbelief that **[End Page 401]** they [homeopathic investigations] can bring about . . . [required experiments] with all the necessary guarantees." [34] Double standards notwithstanding, intentional ignorance was selectively applied to this controversial healing system to clear the confusion caused by "the spontaneous course of most natural diseases." [35] For Trousseau and his colleagues this meant the deceptive administration of bread pills (*des pilules d'amidon* or *la mie de pain*). [36] These blind assessments using an inert substance are the earliest I have been able to uncover in which blind assessment was used to study a claim of drug efficacy.

Following customary procedures of the time, a comparison arm (in this case, a treatment group using a genuine homeopathic remedy) was absent. The conclusion was to be based on the experimenter's expert judgment. The test consisted of a series of at least ten patients receiving treatment that they were led to believe was homeopathic. The sham included the full-blown ritual: the pills were administered with all the "sacramental words from Hahnemann," [37] and the treatment and follow-ups were "extreme and extraordinarily detailed . . . [going] through not only the symptoms but also the most fleeting feeling that could be felt." [38] Trousseau and his students believed that the observed results were due to natural history and imagination; they concluded that homeopathy was not "more active than the most inert substances," [39] and that its therapeutic effects were "nonexistent." [40]

The nineteenth century witnessed other blind assessments of homeo-pathy's efficacy. As the skeptical physician John Forbes (1783-1861) said, the claims of homeopathy warranted "the necessity of insisting on this extreme degree of evidence." [41] Some of these trials were an improvement (from a modern perspective) over Trousseau's model. Some actually **[End Page 402]** utilized concurrent arms with a genuine homeopathic remedy and a placebo; others involved a simultaneous masked comparison of homeopathic and orthodox treatment. [42] In both experimental designs, patients were placed under conditions of ignorance.

The most rigorous of such experiments that I have been able to find was a cooperative venture between homeopaths and orthodox physicians sponsored by the Milwaukee Academy of Medicine in 1879-80. [43] In this trial, which could be described in modern terms as "double-blind," [44] both patients and experimenters were blind to whether the treatment was a genuine homeopathic remedy or a sugar pill. (Homeopathic dilutions are routinely administered by means of sugar pills, making an identical-looking sham treatment easy to disguise.) The experiment utilized a set of several commonly used homeopathic remedies that the homeopath could determine in advance. Each vial of remedy was then matched with an identical-appearing decoy vial containing sugar pills. **[End Page 403]** The homeopathic physician was supposed to select an appropriate remedy for a chronic patient and give one of the matched real or sham remedies (they could later administer the other vial). Then the homeopath had to decide which of the two matched vials was the verum. The minister and professor of mental and moral philosophy who supervised the blinding reported:

> Great pains have been taken to exclude entirely the possibility of guessing the medicated vials, instead of discovering them by scientific experiment. Nothing has been permitted to indicate a difference in the vials tested, or to make it possible for any experimenter to detect in any way the reasons for choosing one number rather than another, of all the vials numbered to contain the medicated pellets. [45]

The trial had unanticipated recruitment problems, and the results were inconclusive. [46]

Blind Assessment as a Research Tool in the Nineteenth Century

Blind assessment was well known by the middle of the nineteenth century. The earlier confrontations were well publicized in the professional and popular press. By the middle of the century, at least two medical communities saw advantages in adopting the procedure to investigate some of their intraprofessional agendas. In a quirk of history, the homeopathic medical profession became the first group of researchers to routinely adopt blind assessment in their internal evaluation of homeopathic remedies. Also enlisting the method were the mandarin medical leaders of the therapeutic nihilist movement.

Homeopathic Provings

The homeopathic internalization of blind assessment began in 1842 in Vienna, then a hotbed of the "new medicine" with a group of self-described "scientific" homeopaths. [47] One of the first modifications these scientific homeopaths proposed was a refinement of Hahnemann's original **[End Page 404]** idea that one needed to understand what symptoms a substance provoked in healthy people in order to match them to the symptom-constellation of patients. Hahnemann had developed a method called "proving" (a transliteration of the German *Prüfung,* meaning test or assay), which tested substances on healthy volunteers. [48] These scientific homeopaths were worried about imagined symptoms during provings if the participants knew the substance's identity.

Accordingly, they adopted the blind precautions then in vogue for magnetic experiments and for challenges to homeopathy's claims of effectiveness. [49] Volunteers were given substances whose identity was secret. The first such masked assessment, which took place from 1 November 1842 to 10 January 1843, **[End Page 405]** employed a group of fifteen volunteers (mostly physicians) who "used a tincture of koloquinte [*Colocynthis germanica*] prepared according to Hahnemann's instructions, and . . . almost all of the provers did not know which medications they were proving." [50] Subsequently, this Viennese circle of provers took more rigorous measures where "no one knew what substance they were taking." [51] By 1857, they had performed at least eighteen such provings. [52] Blind assessment continued to be used by homeopaths, and by 1900 most homeopaths had routinely adopted some sort of concealment procedure in provings. [53] **[End Page 406]**

Therapeutic Nihilism

While blind assessment before the end of the nineteenth century was mostly confined to peripheral areas of healing and science, placebo assessments also took place within the inner sanctum of orthodox medicine. Exponents of the mid-nineteenth-century elite medical movement of "therapeutic nihilism," which believed that no treatment at all was as good as or better than routine therapeutics, adopted the tool to demean prevailing practice. [54]

The earliest sham assessment for an orthodox intervention that I have found was performed by Austin Flint (1812-62), one of the most prominent American medical leaders of his time (his name is still memorialized in the Austin Flint heart murmur). His last research effort, performed at Bellevue Hospital Medical College and published posthumously, assessed whether the prevailing drugs for rheumatism had any effect on the outcome of the "natural course" of disease. Thirteen patients "were placed on the use of a placebo which consisted, in nearly all the cases, of the tincture of quassia, very largely diluted. This was given regularly, and became well known in my wards as the *placeboic remedy* for rheumatism." [55] Following the prevalent custom, the experiment lacked a comparison arm and the interpretation relied on the experience and wisdom of the senior clinician. After giving the details of each case, Flint concluded that nature took its own course and the disease was self-limited. The orthodox medical treatment was a concurrent event that usurped the credit due "nature."

At the same time that therapeutic nihilism was influencing America, Guy's Hospital in London was also a center for the movement and another site for an early intraprofessional polemical challenge. William Withey Gull (1816-90), Guy's leading practitioner, undertook to demonstrate that the prevailing medical treatment for rheumatic fever had only the illusion of efficacy. Twenty-one rheumatic fever patients ("with no selection") were treated "for the most part by mint water" thought by the physicians to be pharmacologically neutral. [56] Patients were led to believe **[End Page 407]** that it was a real intervention. The patients seemed to do well (again, there were no concurrent controls). The conclusion of this sham case series was that the "natural course of the

disease had more to do with the result than the remedy." [57] Similar trials in German-speaking countries may have preceded these Anglo-American efforts. [58]

Blind Assessment and Mental Phenomena at the Fin de Siècle

In the late nineteenth century, in its attempt to objectify and quantify the apparently private contents of the mind, psychology generated its own legacy of blind assessment, which eventually had an impact on medical and scientific methodology. At the same time, blind assessment was also used by psychical researchers to establish their scientific credentials while seeking to detect controversial mental powers. The most important relevant events of this period, however, concerned the fin de siècle debates over hypnotism-suggestion. For the first time, at least in France, blind assessment moved to the very center of a "conventional" medical controversy, when well-publicized ordeals of darkness sought to prove that the modus operandi of hypnotism was either mental psychology or physiological neurology. Three episodes of blind assessment investigating potential mental processes are described below.

The Peirce and Jastrow Experiments

One of the earliest uses of blind assessment in psychology was a study by Charles Sanders Peirce (1839-1914), one of the founders of American pragmatism, and his student Joseph Jastrow (1863-1944), who became president of the American Psychological Association. This experiment has already been well described by the historian-philosopher Ian Hacking, but it is worth retelling. [59] Peirce and Jastrow were interested in an earlier psychological experiment to determine what was the smallest **[End Page 408]** discernible difference in sensation. [60] Between December 1883 and April 1884 they improved on the original design by using a "screen," so that the person who was trying to feel the slight differences of weight could not tell whether the weights were being increased or decreased. This American team was "fully on . . . guard against unconsciously received indications." [61] This use of blinding and randomization is probably the first time that scientists studying a mainstream question (as opposed to a marginal and deviant issue) self-consciously saw value in their remaining ignorant. In this case, using a veiled assessment was less a means of fraud detection and more a guarantee of accurate observation. Exactly where Peirce and Jastrow's idea for this blinding came from is unclear. One possibility is that Peirce understood the problem of errors of observation and systematic bias from his earlier work in geodesy and astronomy. [62] But it is also conceivable that their interest came from the unconventional domain: both scientists were also actively involved in the furor concerning spiritualism and psychical research (see below). [63] **[End Page 409]**

Telepathy and Psychical Research

The word *telepathy* (meaning "thought transference") was invented in England in 1882 by Frederic W. H. Myers (1842-1901), one of the founding fathers of the Society for Psychical Research (SPR). [64] The aim of the new society, founded in the same year,

was to scientifically investigate the phenomena of spiritualism, which was raging through the second half of the nineteenth century. [65] Spiritualism claimed that the dead could communicate with the living, thereby providing empirical and unassailable evidence of a soul that survives the death of the body. [66] Spirits (through mediums) diagnosed and even prescribed medical treatment. [67] The SPR considered the possibility that such communication (if it could be demonstrated to be accurate) was not the result of discarnate spirits, but rather the product of a subtle "natural" mental capacity that might belong to people generally, or perhaps only to some "sensitives." They wanted to remove the fringe and occult taint associated with spiritualist phenomena. Accusations of delusion and fraud quickly challenged the SPR to establish a semblance of scientific skepticism.

The earliest use of blind assessment in psychical research seems to have been initiated by the French research physiologist Charles Richet (1850-1935), whose career culminated with the 1913 Nobel Prize for his work on anaphylaxis. His work on psychical research has been well **[End Page 410]** described, again by Ian Hacking, but a recapitulation is fitting. [68] Beginning in 1884, Richet was concerned with whether a person could draw a card at random from a deck of cards and then, with concentration, communicate this card to another person. Although he was worried about the possibility of "trickery" and any "tell-tale signs either in the movement of the eyes or in facial expressions," [69] and he was likely aware of earlier mesmerism experiments, he did not adopt a blindfold method in his very first card trials. Later, he decided to keep the subject "hidden behind a screen" (*caché derrière un écran*) to avoid any taint of fraudulent maneuvers. [70] From this point on, blinding quickly became an essential feature of psychical research, as did Richet's random selection methods (*au hasard*), which he used as an additional precaution to ensure concealment. [71] When university-sanctioned psychical and parapsychology research centers were opened in the early twentieth century, blind assessment and early forms of randomization were also an integral component of their research protocols. [72] **[End Page 411]**

The Conflict between Salpêtrière and Nancy over Hypnotism and Suggestion

The mostly French debates concerning hypnotism-suggestion produced a flurry of blind assessment experiments to determine whether the observed effects were due to "objective" material agency or to suggestion. In many ways, this affair represents the migration of a "sanitized" neomesmeric controversy into the inner sanctum of orthodox medicine. [73] *Hypnotism* was James Braid's refurbished term for a "psychological" mesmerism lacking an occult or vitalist component. It had mostly languished on the fringe of medicine until it was single-handedly rescued from mainstream oblivion by Jean-Martin Charcot (1825-93), the founder of modern clinical neurology at the Salpêtrière in Paris. Charcot came to see hypnotism (or "Braidism") as an abnormal physiological phenomenon allied to hysteria and describable in neurological terms. [74] He believed he had discovered in the hypnotic state objective and mechanical indices with definable stages.

Others, especially Hippolyte Bernheim (1840-1919) of Nancy, radically disagreed with Charcot and shared views closer to Braid's. Instead of objective stages they saw hypnotism as suggestion, which Bernheim characterized as "the influence exerted by an idea . . . received by the mind" **[End Page 412]** that could translate into action, sensation, or movement. [75] The early debate revolved around Bernheim's contention that suggestion was responsible for Charcot's patient's ability to perform on cue and the contention of Charcot's followers that Bernheim's patients were not victims of genuine neurological *grand hypnotisme.*

Blind assessment entered the trenches when Charcot's disciples joined the fracas. The concern was not therapeutic efficacy, but rather the establishment of neurological and/or psychological facts. In 1885 two of Charcot's pupils, Alfred Binet (1857-1911--destined to be famous for his innovative work in psychological testing) and Charles Féré (1852-1907--later known for his work on criminality), provided additional experimental evidence of the material and objective nature of hypnotism. Studying an earlier reported phenomenon of "hypnotic transfer," Binet and Féré, by using magnets, were able to transfer various hypnotic phenomena, such as unilateral hallucinations (i.e., hallucinations that were visible in only one eye) from one side of the body to the other. In order to rule out suggestion, they used "a magnet hidden under cloth" (*un aimant a été dissimulé sous un linge*) and performed sham maneuvers. [76] These experiments, as far as I can determine, were the Charcot team's first attempt at "intentional ignorance" in this controversy. Bernheim countered that the magnets were empowered by subtle visual and auditory clues. [77] When more rigorous caution was used (such as not talking, and staying out of view), he argued, Binet and Féré's experiments were not reproducible. He also used false suggestive cues (saying he was **[End Page 413]** magnetizing a leg instead of an arm), and instead of magnets used "a pencil, a bottle, or a piece of paper," and obtained equally good results. [78]

In 1886 Joseph Babinski (1857-1932), of the famous reflex, took hypnotic transference one step further and with magnets was able to transport hypnotic symptoms (e.g., paralyses, contractures) from one somnambulistic subject to another. In his experiments, Babinski took what he considered "all the precautions necessary to keep subjects in complete ignorance," which included separate rooms for induction and covering the first subject with a veil to totally hide the body. [79] Again, charges of inadequate attention to subtle cues were made. [80]

A significant number of other blind assessments took place during the suggestion conflict. [81] Some concerned more outrageous claims, such as the "neomesmeric" proposition that drug effects could be transmitted at a distance of several feet. The report of one such trial on distant drug effects contains what may be the first explicit description of what is now called a double-blind assessment for a conventional drug (albeit in an unusual, nonphysical delivery vehicle!): an experimental mishap occurred when the experimenter confused the test drug and unintentionally (along with the test subject) became "ignorant." [82] **[End Page 414]**

Blind Assessment in Turn-of-the-Century Physiology and

Pharmacology

The hypnotism-suggestion debate generated enormous scientific and public awareness (at least on the European continent) of the potential power of suggestion. This recognition gradually influenced and promoted a perception of the necessity to adopt blind assessment with placebo controls when investigating the effects of stimulants and other substances on human beings.

Brown-Séquard's Testicular Extract

The trail of blind assessment from neurology and psychiatry to physiology and pharmacology, as best I can uncover, began on 1 June 1889. Charles Édouard Brown-Séquard (1817-94), one of the most prominent scientists of his time, stunned the Société de Biologie in Paris by announcing that subcutaneous injections of animal testicles could rejuvenate physical and mental health. Brown-Séquard ended his original announcement with an unusual caveat for a research physiologist: with the Charcot-Bernheim furor in the background, he was forced to remark that he did not want to discuss to what extent "autosuggestion, without hypnosis" (*une sorte d'auto-suggestion, sans hypnotisation*) could explain his outcome, and he dismissed the possibility as remote. [83] Despite all the attention to blind assessment in the hypnotism-suggestion literature, Brown-Séquard followed the "mainstream" convention of his science and did not consider blind assessment helpful in physiology or medicine. Rather, he advocated the old system of reliance on experienced investigators and called on other "physiologists advanced in age" to perform similar self-experiments to confirm his results with case histories. [84] **[End Page 415]**

For most of the medical and scientific profession the debate on the extract was framed with the need to produce verified case histories. [85] But the fear of delusion began to be raised, especially as testicular extract became an international popular sensation. [86] Segments of professional medicine even began to raise the likelihood of "suggestion and the influence of the imagination." [87] A handful of French researchers considered blind assessment. Less than two months after Brown-Séquard's original announcement, Dr. M. G. Variot at the Hôtel Dieu performed the first independent confirmation of the new therapy: he gave three elderly men the injections, telling them that they were receiving "fortifying" (*fortifiante*) injections, and obtained results identical to Brown-Séquard's. [88] Subsequently, to address the suspicion of suggestion, Variot gave two other patients injections of water with the same instruction, with no effect; these two individuals were later injected with the extract and underwent youthful changes. [89] These two patients and the blind test of the extract are, to my knowledge, the first instance of a substance advocated **[End Page 416]** by a "mainstream" research scientist having undergone an assessment under conditions of ignorance. Again, intentional ignorance was selectively utilized for potentially outrageous claims. [90]

Testing Stimulants in Germanic Europe

Regarding the Brown-Séquard extract, German-speaking physiologists saw the necessity of blind assessment more urgently than did their French counterparts. In this part of Europe, the scientific concern with hypnotism and suggestion quickly became better established than in France. [91] By 1894, Auguste Forel (1848-1931), director of the famed Burghölzli Psychiatric Hospital in Zurich and well known as a brain anatomist and entomologist, made an early plea for "a concerted scientific-logical activity . . . to eliminate . . . the suggestive aspects [of a treatment] from other aspects in the healing consequences of therapy." [92] He was especially critical of the "senile-erotic ideas" of testicular extract, whose purported effects he considered to be due to "powerful suggestive factors." [93]

In 1894, Dr. Fritz Pregl at the Institute of Physiology at the University of Graz (Austria) was one of the early scientists to take up Forel's challenge and perform an experiment on the effects of testicular extract on the work performance of two students. Pregl adopted precautions to ensure that "the influence of suggestion" had been "eliminated or [was] present in equal quantities in the test persons." [94] The experiment supported the testicular extract claims as being beyond the effect of suggestion alone. **[End Page 417]**

The interest in suggestion generally, and the Brown-Séquard testicular extract debate in particular, seemed to generate a more widespread movement by German physiologists toward blind assessment of various substances on healthy subjects. By 1895, a series of papers had begun to appear in German that had placebo controls in the evaluation of such substances as cola, caffeine, cocaine, mate, alcohol, and tea. For example, in 1895 one of the earliest such physiologists, Waclaw Sobiera´nski, a Polish-born docent at the University of Marburg (Germany), explained how he preceded his experimental assays of caffeine and cocaine with saline injections. [95] He also criticized his colleagues for neglecting suggestion. He reported that he routinely used bread pills and saline injections to control for "autosuggestion," which "played a large role in healing," and that "much of the healing properties of chemical medicine is in reality attributable to autosuggestion." [96] He argued that in "order to keep psychological influence to a minimum, all subjects needed to be kept unaware of the experimental substance," kept ignorant that a deception was possible, and that all smells and appearances that could identify pills or injections needed to be "masked." [97] Placebo controls had become a method, for at least some mainstream scientists, of investigating **[End Page 418]** potential "orthodox" claims for the action of substances on human subjects.

Pharmacological Testing in English-Speaking Countries

Soon after the turn of the century, the experimental psychologist and anthropologist W. H. R. Rivers (1864-1922), of Cambridge University, explicitly adopted his German colleagues' method for testing stimulants and other substances. Reflecting the prevailing Anglo-American lack of interest in the Continental preoccupation with "suggestion," Rivers stated that his concern was less suggestion and more the arousal caused by "the interest and excitement" in the experiment. [98] His experiments were "carried out with the use of control mixtures which have usually been wholly indistinguishable from those

containing the active substance." [99]

Rivers's notions of blinding in pharmacological research seem to have been introduced to the United States--with only minimal effect--by such figures as H. L. Hollingsworth (1880-1956) at Columbia University, Torald Sollmann (1874-1965) of the American Medical Association's Council of Pharmacy and Chemistry, and David Macht (1882-1961) at Johns Hopkins University. All three were aware of Rivers's work and the preceding German work, though "suggestion" never seems to be mentioned in their writings. In 1912, Hollingsworth replicated Rivers's method with caffeine. [100] In 1913, a physician under Sollmann's direct supervision compared the effects of "natural" and "synthetic" sodium salicylate (two active agents) by having a group of physicians blindly administer boxes with "only a serial number." [101] The rationale for the blinding was the issue of "bias," described in terms that related to the concern with **[End Page 419]** variances in systematic observations of astronomy. [102] In 1916 Macht, following the introspective tradition of experimental psychology, performed a series of experiments on himself and two students studying the precise quantity of analgesia produced by opium alkaloids. Again, to avoid "bias," he followed Rivers's methods and adopted "as controls normal saline and other inactive substances [which] were often substituted in place of the drug without the subject's knowledge." [103]

German Scientists Incorporate Blind Assessment into Clinical Medicine

Eventually, the German concern about suggestion moved into the arena of clinical research. The first large-scale comparative clinical trial of a conventional medical treatment involving a blind sham assessment that I have been able to uncover is a forgotten experiment from the end of the diphtheria antitoxin debate. Performed by Adolf Bingel (1879-ca. 1950), a physician at a district hospital in Braunschweig, near Hanover, the experiment sought to demonstrate that the diphtheria antitoxin, newly accepted in the profession, was in fact no better than a sham treatment. [104] Between 1911 and 1914, Bingel assigned 937 patients alternately **[End Page 420]** to either diphtheria antitoxin serum or normal horse serum (as a sham). [105] All patients and participating physicians (except Bingel) were unaware of which treatment was genuine; the trial would probably be considered "double-blind" in today's lexicon. [106]

The German interest in suggestion and the necessity of blind evaluation for pharmacological substances reached its culmination in the work of Paul Martini (1889-1964), a Berlin medical researcher. As far as I can determine, his sixty-nine-page manual, *Methodenlehre der therapeutischen Untersuchung* (The methodology of therapeutic investigation), is medicine's first monograph to meticulously describe a methodology for controlled experimentation in clinical drug investigation. Martini's rules were designed to exclude the confounding effects of suggestion. [107] He implemented **[End Page 421]** his procedure in "an exaggeratedly strict fashion" in a series of experiments on drugs that treated acute symptoms (especially for angina pain) in primarily single-subject crossover studies, which would now be called "single-blind." [108]

With Martini's students carrying forward his methods, blind assessment became more frequent in 1930s Germany (though it was still far from common), with researchers using sham assessment to control for what they called "autosuggestion" [109] and the "suggestibility, the expectations of the patient and . . . the physician's personality." [110]

Blind Assessment and the Anglo-American Design of Controlled Trials

The motivation for blind assessment in Anglo-American clinical research developed mainly as an issue of experimental design and initially had little to do with the Continental concern with suggestion. [111] For Anglo-American researchers, it was technical organizational problems that were the driving force. Later, new rationales (unknowingly borrowed from the unconventional wars of an earlier era) would be deployed, but the originally structural issues were fundamental. This Anglo-American adoption of blind assessment can be described in three stages.

Concurrent Controls and Placebos

At the beginning of the twentieth century, the broad community of clinical researchers had only begun to understand the importance of concurrent controls and a systematic method of matching units of a trial **[End Page 422]** in living systems. [112] When researchers began to enlist patients (as opposed to soldiers and volunteers) in a no-treatment, "natural history," comparison arm of a clinical trial, they were faced with a serious problem: in the words of a famous researcher, "one cannot invite . . . [patients] for a trial, [and] obviously keep half as 'guinea-pigs,' and then hope for co-operation and good records." [113] An obvious no-treatment arm became a recruitment and retention nightmare. Patients would demand "real" treatment or seek out cointerventions. For Anglo-American clinical researchers, the initial adoption of a placebo sham in an experiment was an architectural device to create a viable and camouflaged concurrent no-treatment arm in a clinical trial. (The term *clinical trial* itself was probably coined as late as 1931.) [114] Because informed consent was still not an ethical norm, this dummy treatment could easily become a legitimate concurrent control.

The first Anglo-American experiment using placebos to provide coherence for a no-treatment group appears to have been the Michigan tuberculosis trial of 1926-31. (This trial is one of the most cited as "first blind assessment" in conventional histories of medical research.) At a tuberculosis sanatorium in Northville, near Detroit, researchers seem to have been neither aware of any earlier examples of blind assessment nor concerned with suggestion. Rather, this team felt that previous investigators could not distinguish "the notorious tendency of the disease to fluctuate naturally" from changes induced by the substance or procedure. [115] Explicitly stating that the common procedure of relying on a **[End Page 423]** series of case histories was unreliable, the team adopted the emerging notion of concurrent controls. Therefore, twelve patients received sanocrysin (sodium-gold-thiosulphate, 37.4% gold) and twelve meticulously matched

patients received "intravenous injections of distilled water." [116] Two of the three authors knew which group received the gold solution, as did the ward nurse, making the trial what today would be called single-blind. [117]

The second Anglo-American trial using a placebo in a no-treatment group was apparently an effort at the Cardiac Department of London Hospital, beginning in 1930. This experiment involved ninety ambulatory patients with angina pain. Seemingly unaware of Martini's concurrent work on angina, the investigators deceptively administered placebos to give the "no-treatment" comparison group in the crossover design a defined structure. A variety of sixteen drugs (including nitrates, narcotics, digitalis, and belladonna) was given, interspersed with periods of placebo administration. Like the Michigan group, these researchers were not concerned with suggestion, but rather wished to control for "the natural variations in the severity of the symptoms" that a comparison with an untreated group could detect. [118] No tested drug in this trial was better than the placebo. Several other blind assessments occurred in the 1930s that used a sham to masquerade as treatment in a no-treatment group. [119] **[End Page 424]**

The American Recognition of an "Active" Sham Treatment

Harry Gold (1899-1973) and a team of colleagues at Cornell University Medical School published a paper in 1937 that, especially in retrospect, introduced a shift in the American (and eventually British) understanding of the necessity for blind assessment. [120] For the first time in the English-language research literature a placebo treatment had implications beyond just disguising a no-treatment group: it could also have a quantitative medical outcome, in addition to producing information about the natural history of a disease. [121] For the first time, American researchers conceded that "ineffective" therapy could have confounding effects on expert judgment even if a physician "knew" the spontaneous course of the disease.

Gold's group tested whether the methyl xanthines (theobromine and aminophylline) were helpful for angina patients. The experiment was similiar in design and outcome to the London angina experiment mentioned earlier (although Gold claimed that it was independently conceived). [122] The interpretation had one "minor" but crucial difference. **[End Page 425]** Gold mostly repeated the normal explanation of "spontaneous variations" as the cause of the equal improvement in the experimental and placebo arms; but at the end of these standard remarks, Gold's team introduced as subsidiary points an entirely new explanation for the effectiveness of sham intervention. These points included "confidence aroused in the treatment," the "encouragement afforded by any new procedure," and "a change of the medical advisor." [123] Although Gold seemed to studiously avoid the continental word *suggestion,* operationally the European justification for a placebo in blind assessment had crept into the English-language research literature. [124] Gold and his colleagues became important proponents of blind assessment after World War II, **[End Page 426]** but Gold's impact before the war was minimal and evidence of his influence can be found in only a few experiments and discussions. [125]

The Randomized Controlled Trial

The potential of intentional ignorance to preserve the integrity of concurrent controls or to combat delusion, bias, or other psychological factors convinced only a few pre-World War II Anglo-American researchers of the necessity for treating patients "in the dark." Until well past World War II, clinical research consisted overwhelmingly of open-labeled comparative trials. [126] The impetus for the swift adoption of blind assessment into a post-World War II mantra of modern experimentation was intimately tied to the rapid introduction, acceptance, and wholesale assimilation of the fully randomized research design of R. A. Fisher (1890-1962). Blind assessment was perceived as the corollary and companion of a new methodology that could make medicine into a full-fledged "hard" science.

Fisher's insights developed at the agricultural experimental station in Rothamsted, England. They began to have a decisive impact on experimental design with his second statistical book, *The Design of Experiments* (1935), which emphasized the importance of randomization for allowing **[End Page 427]** a measure of uncertainty and variability. [127] For Fisher, randomization allowed estimates of error and tests of significance "to be fully valid" because "the very same causes [variability and chance] that produce our real error shall also contribute the materials for computing an estimate of it." [128] Fisher's method allowed the use of probabilistic statistics, and the very indeterminateness of variability inherent in biological phenomena, to force nature to yield a mathematically precise answer. This measure of uncertainty itself allowed for causal inference.

For the medical elite struggling to make medicine more of a science, Fisher's ideas offered a method with the appearance of scientific exactitude that could imitate the determinism and objectivity of the laboratory model and mimic the mechanical and mathematical precision of hard science. [129] Clinical research could free itself from human judgment and resemble the precision of the laboratory. The perceived scientific rigor of the RCT could legitimately stake a claim to much of the terrain that had previously clung to medicine's "art." The RCT offered "the dream of the scientist who arrives at new knowledge by a completely mechanized procedure." [130]

The new methodology of assessing efficacy, standardized by statisticians and predetermined protocols, would take "a multitude of decisions out of the hands of participating investigators . . . [and] remove a series of opportunities for clinicians to frustrate" and thereby confound experiments. [131] As the historian Harry Marks has noted, for the researcher, "the improvements in experimental method offered by statisticians represented an elegant technical fix for a host of previously insoluble organizational and social problems." [132] The masked RCT could significantly **[End Page 428]** realign the power relationships between "art" and "science" in medicine, as it was itself a product of this transformation. [133]

Fisher's original justification for randomization was purely statistical. As mentioned earlier, randomization was a technical statistical maneuver that permitted valid null-hypothesis testing and empowered causal inference.

But stochastic statistics were insufficient justification for most physicians. Clinicians resisted random allocation. Few physicians wanted to randomly assign patients to treatment, forgo the individualization of therapy, and withhold new, promising therapies. Randomization threatened to curtail the autonomy of practitioners and the undisciplined nature of old-style clinical researchers. Austin Bradford Hill (1897-1991), architect of the 1948 British streptomycin clinical trial (generally considered the first genuinely randomized clinical trial in history), many years later confessed that he had "deliberately left out the words 'randomization' and 'random sampling numbers' at that time because . . . I might have scared them [collaborating physicians] off." [134] Hill had to counter arguments from physicians who thought something was wrong with treating patients as so many "bricks in a column," and who feared the "elimination of the responsibility of the doctor to get the individual back to health." [135] Even as late as 1955, *JAMA* could print appeals that the evaluation of drugs should be "[kept] in the hands of the general practitioner and not . . . [be] on the basis of experiments." [136]

The medical reformers who wanted to move clinical research toward the Fisherian model found insufficient support from their colleagues. Mathematical theories did not impress most physicians. [137] An additional justification specifically tailored to clinical medicine was needed. Suddenly, **[End Page 429]** elite physicians were talking about "the fallibility of human judgement in general and of clinical . . . judgement in particular." [138] Researchers began to justify randomization in terms of needing to protect treatment arms from overzealous advocates. Previously, the taint and accusations of bias, prejudice, overenthusiasm, credulity, and delusion were reserved for deviant healers; now, what was once a fringe threat was internalized. Even the judgments of the most senior clinicians concerning the efficacy of new therapeutics were suspect. "Bias" now haunted medicine. A "placebo effect" tied to older ideas of suggestion and expectation was recognized in Anglo-American biomedicine. [139] Harry Gold's concern for the psychological impact of inert interventions suddenly became relevant. [140] Advocates of the RCT forcibly argued for expanding "hard" science into the domain of clinical research, and their new reasons were "secondhand" ones that had once pertained only to marginal medicine. Medical researchers grafted the earlier logic of intentional ignorance onto the polemic for randomization.

The new "medicalized" justification for the RCT quickly inspired both theoretical and practical reasons for also adopting blind assessment in mainstream research. On the theoretical level, the new rationale for the RCT in mainstream research was identical to the old argument for blind assessment at the fringe: randomization and blinding both shielded patients and researchers from the contamination of "knowledge." [141] Now even biomedical physicians were susceptible to bias, suggestion, and **[End Page 430]** delusion. Intellectually, blind assessment became the perfect and necessary corollary for the RCT.

On a practical level, it was realized that the RCT benefited from blind assessment. One cannot easily make random assignments to a no-treatment group when it is clear the treatment is a dummy. Physicians and patients both tended to fudge compliance. As

discussed earlier, a few pre-World War II researchers had already realized that no-treatment groups were difficult to maintain. It was also difficult just to keep the assignment and assessment personnel from becoming aware of the treatment group, unless everyone was fully blinded. [142] It was understood that "blinding is [genuinely] only possible when randomization is employed," and that randomization is only successful under blind conditions. [143] Randomization needed blind assessment. Blindness became the darkness enforcing randomization. Medical researchers were now giving new meaning, relevance, and urgency to such old refrains as "whereas I was blind, now I see." [144]

Blinding together with the RCT became the "gold standard" of science in clinical medicine. Blinding insured compliance with randomization and enhanced the elimination of bias that randomization promised for medicine. Blind assessment moved from the fringe of medicine to its very core. Suddenly, there was a "relatively new method of . . . [the] blindfold tests" that needed to be adopted by biomedicine to insure objectivity and rigor. [145] Clinical medicine needed to rely on uncontaminated evidence. By repeating enough times that blind assessment was new, most observers came to believe that intentional ignorance in research **[End Page 431]** was indeed a recent advance. [146] Medical science managed to re-create its modus operandi and its own history.

Conclusion

Modern research methodologies, including blind assessment, obviously have a complex, "context-bound" social history. The adoption of blind assessment in medicine has had as much to do with shifting political, moral, and rhetorical agendas and technical research design issues as with scientific standards of evidence. The project of using blind assessment as a tool to demarcate the boundary between material causality and mere belief has been an enterprise with far-reaching epistemological ramifications. But blind assessment has also been a vehicle to confer social authority and moral legitimacy. Intentional ignorance began as a method to challenge the "bogus" claims of unconventional medicine; some unorthodox practitioners adopted it in self-defense. At times, some nineteenth-century iconoclastic conventional medical leaders found it valuable in their polemics. Later, veiled procedures moved into psychology, psychic research, neurology, psychiatry, and pharmacology. Although the method had been available since the late eighteenth century, conventional medicine perceived its value only when other pressing considerations were at hand. At each stage of this movement, the motivation for intentional ignorance was distinct and included scientific and extra-scientific dimensions.

The history of concealed assessment has been hidden from both researchers and historians. Perhaps part of the reason for this shadowy past is the intense fervor and absolute authority with which modern biomedicine advocates it (at least when its use is possible). To use Ian Hacking's phrase, the justification is "self-authenticating." [147] Concealed history augments the appearance of an obvious transcendent truth. Questions are discouraged. "It becomes less something molded by interests, **[End Page 432]** and more an unquestioned resource upon which any interest must draw, if it ever

hopes for the accolade of objectivity." [148] History disturbs the veneer of eternal validity. Perhaps an examination of blind assessment, along with other research methodologies, needs to be continued into the present. What are the scientific and extrascientific assumptions that underpin these modern methodologies? [149] Can an examination of such questions make our contemporary methods of demarcation between fact and fiction a less simple but richer process that reveals even more of the "light of truth"?

Ted J. Kaptchuk is an Instructor in Medicine at Harvard Medical School and the Associate Director of the Center for Alternative Medicine Research at Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02115 (e-mail: tkaptchu@bidmc.harvard.edu). Much of his previous work concerned East Asian medicine, and one of his current interests is research methodology and the construction of evidence.

* I wish to thank several individuals for their help on this manuscript: Velia Wortman for guidance, encouragement, and intellectual persistence; Jonathan McDonagh for research assistance; David Eisenberg for advice and support; Helmut Kiene, Irving Kirsch, Klaus Linde, Harry Kaptchuk, Patrick Mehr, Jacek Czyzewski, Adriane Fugh-Berman, Howard Brody, Herman Engelbart for archival and research help; and Janet Walzer, Debora Fischer, and Sonya Keller for editorial assistance. I also want to acknowledge the insightful suggestions provided by the five anonymous reviewers of this paper. My research was partially supported by grants from the John E. Fetzer Institute and NIH grant U24 AR43441.

Notes

1. A review of even the best sources on the history of medical research reveals little, inaccurate, and contradictory information. For example, one source devotes five paragraphs to masked assessment and traces the first double-blind trial to 1931: Abraham M. Lilienfeld, "*Ceteris paribus:* The Evolution of the Clinical Trial," *Bull. Hist. Med.,* 1982, *56:* 1-18. No direct mention of blind assessment can be found in J. P. Bull, "The Historical Development of Clinical Therapeutic Trials," *J. Chron. Dis.,* 1959, *10* (3): 218-48. A third source has three sentences on blind assessment and traces the first double-blind trial to 1937: Michael D. Rawlins, "Development of a Rational Practice of Therapeutics," *Brit. Med. J.,* 1990, *301:* 729-33. A fourth has four paragraphs on blind assessment and placebo controls and mentions the first use as 1933: John H. Gaddum, "Clinical Pharmacology," *Proc. Roy. Soc. Med.,* 1954, *47:* 195-204.

Textbooks on clinical trials usually have an introductory section on history, and again blind assessment is given only perfunctory attention. For example, one text has two paragraphs on placebo controls describing the Perkins tractor experiment of 1799 and the mint-water experiment of 1863: Christopher J. Bulpitt, *Randomized Controlled Clinical Trials* (The Hague: Martinus Nijhoff, 1983). Two paragraphs on the history of blind assessment and a chart with two pre-World War II episodes can be found in Curtis L. Meinert and Susan Tonascia, *Clinical Trials: Design, Conduct, and Analysis* (New York: Oxford University Press, 1986).

A similar pattern emerges in histories of research. For example, only a passing mention of blind assessment is found in J. Rosser Matthews, *Quantification and the Quest for Medical Certainty* (Princeton: Princeton University Press, 1995). One sentence on the double-blind method can be found in Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton: Princeton University Press, 1995).

2. It should be noted that this essay deals with the placebo under experimental conditions, and not with its history as a therapy. For a summary of the issue of the placebo in clinical practice, see: Ted J. Kaptchuk, "Powerful Placebo: The Dark Side of the Randomized Controlled Trial," *Lancet,* 1998, *354:* 1722-25; cf. Arthur K. Shapiro, "Semantics of the Placebo, " *Psychiatric Quart.,* 1968, *42*(4), 653-95. Additionally, the overwhelming majority of experiments described in this essay took place before the notion of "informed consent" had been developed. Most experiments were conducted on patients who unknowingly received sham interventions. For a discussion of the history of informed consent, see Ruth R. Faden and Tom L. Beauchamp in collaboration with Nancy M. P. King, *A History and Theory of Informed Consent* (New York: Oxford University Press, 1986); Susan E. Lederer, *Subjected to Science: Human Experimentation in America before the Second World War* (Baltimore: Johns Hopkins University Press, 1995).

3. Several excellent recent histories directly concern the development of the randomized controlled-trial method in medicine. See, for example, Harry M. Marks, *The Progress of Experiment: Science and Therapeutic Reform in the United States, 1900-1990* (Cambridge: Cambridge University Press, 1997); and Marcia Lynn Meldrum, "'Departures from the Design': The Randomized Clinical Trial in Historical Context, 1946-1970," Ph.D. dissertation, State University of New York at Stony Brook, 1994 (Ann Arbor, Mich.: University Microfilms, 1997).

4. Kurt Danziger, *Constructing the Subject: Historical Origins of Psychological Research* (Cambridge: Cambridge University Press, 1994), p. 3.

5. John Harley Warner, "Science in Medicine," *Osiris,* 2d ser., 1985, *1:* 37-58, quotation on p. 52. This account could also be considered an example of "the ongoing narrative reconstruction of scientific practice" that simultaneously reconfigures the past while it sets future research priorities. Joseph Rouse, *Engaging Science: How to Understand Its Practice Philosophically* (Ithaca, N.Y.: Cornell University Press, 1996), p. 176.

6. Robert E. Kohler, *From Medical Chemistry to Biochemistry: The Making of a Biomedical Discipline* (Cambridge: Cambridge University Press, 1982), p. 6. A similar sentiment is even more emphatically expressed by Richard Lewontin in the context of other scientific enterprises: "the origins of differences in required rigor are not always easy to discern. . . . The quality of evidence itself is tailored to fit ideological demand" (Richard C. Lewontin, "Facts and the Fictitious in Natural Sciences," in *Questions of Evidence: Proof, Practice, and Persuasion across the Disciplines,* ed. James Chandler, Arnold I. Davidson, and Harry Harootunian [Chicago: University of Chicago Press, 1994], pp. 489, 491).

7. There are many excellent discussions on the emergence of the modern orthodox-

versus-sectarian medical disputes, including William F. Bynum and Roy Porter, eds., *Medical Fringe and Medical Orthodoxy, 1750-1850* (London: Croom Helm, 1987); Roy Porter, *Health for Sale: Quackery in England, 1660-1850* (Manchester: Manchester University Press, 1989); Roger Cooter, ed., *Studies in the History of Alternative Medicine* (New York: St. Martin's Press, 1988); Alison Klairmont Lingo, "Empirics and Charlatans in Early Modern France: The Genesis of the Classification of the 'Other' in Medical Practice," *J. Soc. Hist.,* 1986, *19:* 583-604.

8. The word *boundary* here follows one of Thomas Gieryn's definitions, where it is meant to "exclude[s] rivals . . . by defining them as outsiders with labels such as 'pseudo,' 'deviant,' or 'amateur'" (Thomas F. Gieryn, "Boundary-Work and the Demarcation of Science from Non-science: Strains and Interests in Professional Ideologies of Scientists," *Amer. Sociol. Rev.,* 1983, *48:* 781-95, quotation on p. 792). Cf. Roy Wallis, ed., *On the Margins of Science: The Social Construction of Rejected Knowledge* (Keele, UK: University of Keele, 1979).

9. For a discussion of the popularity of mesmerism and the entire mesmeric phenomenon, see Robert Darnton, *Mesmerism and the End of the Enlightenment* (Cambridge: Harvard University Press, 1968). It is possible to find "precursor" incidents of blind assessment in the pre-modern era. Generally they were used to demonstrate medical virtuosity, as opposed to squeezing veracity out of the distortions of the imagination, For example, Muslim and medieval western medicine both featured a well-known story in which a famous physician claimed to be able to perform diagnosis through a string tied to the patient's radial pulse. In a test, the Sultan deceptively gave the physician a string tied to the tibial pulse of a cow and asked the physician to diagnose the "modest" woman in the adjoining room. When the physician said that the creature needed grass he was considered truly proficient (Reuben B. Amber and A. M. Babey-Brooke, *The Pulse in Occident and Orient: Its Philosophy and Practice in India, China, Iran, and the West* [New York: Santa Barbara Press, 1966], p. 2).

Another example from seventeenth-century France involved Sir Kenelm Digby (1603-65), who reported an accidental blind assessment confirming his "sympathetic natural magic" method of treating wounds. Digby inadvertently discontinued the treatment, while at the exact same moment the patient (in another part of the room) suffered an acute exacerbation of his pain. The incident was reported as a secondary corroboration of Digby's method rather than an important procedure to unshackle material science from the delusions of the mind (Kenelm Digby, *Of the Sympathetic Powder: A Discourse in a Solemn Assembly at Montpellier, Made in French, by Sir Kenelm Digby, Knight, 1657* [London: John Williams, 1669], pp. 148-50). Only with the advent of mesmerism did material science perceive the need to overthrow the notion that medical events are accurately accessible to sensory awareness and the mind.

10. See George Bloch, trans., *Mesmerism: A Translation of the Original Scientific and Medical Writings of F.A. Mesmer* (Los Altos, Calif.: William Kaufmann, 1980), p. 25.

11. Franklin had already conducted a series of observations on the use of electricity in paralytic cases in Pennsylvania since at least 1757. These experiments were not blind. See Benjamin Franklin, "Letter XIX. to John Pringle, M.D. and F.R.C.S., December 21,

1757," in *Benjamin Franklin's Experiments: A New Edition of Franklin's Experiments and Observations on Electricity,* ed. I. Bernard Cohen (Cambridge: Harvard University Press, 1941), p. 347. For further discussion of Franklin's relationship to mesmerism and of how the commission's research strategy may have been conceived, see Denis I. Duveen and Herbert S. Klickstein, "Benjamin Franklin (1706-1790) and Antoine Laurent Lavoisier (1743-1794), Part II: Joint Investigations," *Ann. Sci.,* 1955, *11* (4): 271-308.

12. Mesmer wanted a prospective comparative experiment for any disease but venereal ones: see Claude-Anne Lopez, "Franklin and Mesmer: An Encounter," *Yale J. Biol. Med.,* 1993, *66:* 325-31, quotation on p. 327; cf. Geoffrey Sutton, "Electric Medicine and Mesmerism," *Isis,* 1981, *72:* 375-92, quotation on p. 387. When Mesmer's offer was refused he declined to cooperate with the commission, and one of his principal disciples, Charles d'Eslon (1739-86), physician-in-ordinary to the king's brother, performed the treatments during the experiments.

13. Benjamin Franklin, Majault, Le Roy, Sallin, Jean-Sylvain Bailly, D'Arcet, De Bory, Joseph-Ignace Guillotin, and Antoine Laurent Lavoisier, *Report of Dr. Benjamin Franklin, and Other Commissioners, Charged by the King of France, with the Examination of Animal Magnetism, as Now Practiced in Paris,* trans. William Godwin (London: J. Johnson, 1785), p. 38. (Whenever possible, I have used a published English translation of original source material; in other cases the translations are my own.)

14. Ibid.

15. Ibid., p. 56.

16. Ibid., p. 58.

17. Sham or decoy assessments were also part of the experimental strategy. For example, a twelve-year-old boy subject, selected by the mesmerist, was led up to five trees, one of which had been mesmerized, in Franklin's garden. Previously, the boy had routinely fainted in the presence of a mesmerized tree. This time he had his eyes covered with bandages so that there would be "no communication" between him and the mesmerist; he passed out and needed to be carried out of the garden when he embraced the wrong tree (ibid., p. 67). At another session, this time at Lavoisier's house, a patient was mesmerized by plain water (when told it was "mesmerized" water), but had no sensations from genuinely treated water (ibid., p. 73).

18. Ibid., p. 97. The commission actually concluded that the effects of mesmeric fluid were due to "compression [the touching that happened during sessions], imagination and imitation" (ibid., p. 97). But imagination was generally considered the chief cause: ibid., p. 102.

19. There was another commission appointed by the Royal Society of Medicine that was also supposed to investigate mesmerism. Its report, which was even more adversarial in tone, was issued a few days after the Academy of Sciences' and was never as widely circulated. One of this commission's members, Antoine Laurent de Jussieu (1748-1836)--a distinguished botanist and physician--wrote a separate report dissenting from the

majority. He described a blind assessment in which the patient was unaware she was being magnetized in a crowded room where a mingling of researchers' bodies created a shield. De Jussieu claimed that magnetism occurred even under ignorant conditions. De Jussieu's report is in Alexandre J. F. Bertrand, *Du magnétisme animal en France* (Paris: J. B. Baillière, 1826), pp. 151-210. Critics countered that the precautions to insure blindness were inadequate: see Claude Burdin and Frédéric Dubois, *Histoire académique du magnétisme animal accompagnée de notes et de remarques critiques sur toutes les observations et expériences faites jusqu'à ce jour* (Paris: J. B. Baillière, 1841), pp. 160-65.

20. For excellent discussions of these phenomena see Alan Gauld, *A History of Hypnotism* (Cambridge: Cambridge University Press, 1992); Frank Podmore, *Mesmerism and Christian Science: A Short History of Mental Healing* (London: Methuen, 1909); Robert C. Fuller, *Mesmerism and the American Cure of Souls* (Philadelphia: University of Pennsylvania Press, 1982).

21. The most encyclopedic study of mesmeric phenomena is Eric J. Dingwall, ed., *Abnormal Hypnotic Phenomena: A Survey of Nineteenth-Century Cases,* 4 vols. (London: Churchill, 1967-68). This monumental work meticulously documents scores of episodes of masked mesmeric scientific assessments and entertainment theater throughout Europe and North and South America. Theater performances routinely used blindfolds to show that remarkable feats were possible and as proof of the absence of "trickery." Demonstrations of eyeless sight and forms of clairvoyance where the eyes were bandaged and stuffed with cotton wads attracted both popular and scientific interest. For scientific investigations, more rigorous forms of blinding were adopted. For example, in 1838 John Elliotson (1791-1868) employed a "well-contrived brown paper cap, which completely precluded vision," for his testing of magnectic subjects ("Faculties of Elizabeth O'key," *Lancet,* 1838, *2:* 873-77, quotation on p. 875). John Kearsley Mitchell (1798-1858--father of neurologist Silas Weir Mitchell [1829-1914]) used a thick "doubled shawl, through which I could not see the slightest ray of light" (John Kearsley Mitchell, *Five Essays* [Philadelphia: Lippincott, 1859], p. 165). James Braid (1795-1860) put patients in a "dark closet" in his experiments on the "odic force" (James Braid, *Magic, Witchcraft, Animal Magnetism, Hypnotism and Electro-Biology: Being a Digest of the Latest Views of the Author on These Subjects* [London: John Churchill, 1852], p. 27).

22. Pierre Foissac, *Rapports et discussions de l'Académie Royale de Médecine sur le magnétisme animal recueillis par un sténographe, et publiés, avec des notes explicatives* (Paris: J. B. Baillière, 1833), pp. 272-79. Also see Alfred Binet and Charles Féré, *Animal Magnetism* (London: Kegan Paul, Trench, 1888), p. 33.

23. Foissac, *Rapports et discussions* (n. 22), p. 275.

24. Ibid., p. 276.

25. Ibid., p. 279.

26. Other royal and scientific commissions were created because of public pressure. The most famous of these was that appointed by the Royal Academy of Medicine in 1831. Its

report, which included blind assessments with concealed magnetic healers in adjoining rooms and masked diagnostic clairvoyants, was ultimately positive and stunned the Academy. It concluded that while the phenomenon of magnetic somnambulism was capable of being "feigned and furnish[ed] to quackery the means of deception . . . [it] has been produced in circumstances, in which the persons magnetised could not see or were ignorant of the means employed to occasion it" (A Committee of the Medical Section of the French Royal Academy of Sciences, *Report on the Experiments on Animal Magnetism,* trans. John C. Colquhoun [Edinburgh: Robert Cadell, 1833], pp. 194-95). The commission only established what it considered to be the veracity of the phenomena; it felt that it did not have sufficient patients for a long enough period to decide how to evaluate magnetism's therapeutic effects.

27. See Eric T. Carlson and Meribeth M. Simpson, "Perkinism vs. Mesmerism," *J .Hist. Behav. Sci.,* 1970, *6:* 16-24; Jacques M. Quen, "Elisha Perkins, Physician, Nostrum-vendor, or Charlatan?" *Bull. Hist. Med.,* 1963, *37:* 159-66; idem, "Case Studies in Nineteenth-Century Scientific Rejection: Mesmerism, Perkinism, and Acupuncture," *J. Hist. Behav. Sci.,* 1975, *11:* 149-56. Perkins received the first U.S. government patent for a medical device, won the enthusiastic support of Supreme Court Chief Justice Oliver Ellsworth and of Nathan Smith (founder of Yale Medical School), and even sold a set of "metallic tractors" to George Washington. He died in 1799, while trying to demonstrate the tractor's and his other potions' effectiveness against the New York yellow fever epidemic. His death, however, did not stop his son from bringing the device to England, where it received enormous attention.

28. John Haygarth, *Of the Imagination, as a Cause and as a Cure of Disorders of the Body; Exemplified by Fictitious Tractors and Epidemical Convulsions* (Bath: R. Cruttwell, 1801), p. 2.

29. Ibid., p. 41.

30. Armand Trousseau and Henri Gouraud, "Répertoire clinique: Expériences homoeopathiques [*sic*; this Germanic spelling is often used] tentées à l'Hôtel-Dieu de Paris," *Journal des Connaissances Médico-Chirurgicales,* 1834, *8:* 238-41, quotation on p. 239. Trousseau, of the eponymous spasm, is also generally remembered for pioneering the use of tracheotomy and intubation in medicine, and for his *Traité de thérapeutique et de matière médicale.* He was acutely aware of the magnetic debate, as is evident from his signed entry on magnetism in the *Dictionnaire de Médecine* (Paris: Bechet, Librairie de la Faculté de Médecine, 1833), pp. 11-25. For a description of the tenor of the homeopathic debates in Paris at the time, see Armand Trousseau, Henri Gouraud, and J. Lebaudy, "Correspondance médicale," *Journal des Connaissances Médico-Chirurgicales,* 1833, *1:* 141-42. Homeopathy was very much a physician-based practice with an upper-class clientele. See Olivier Faure, *Le débat autour de l'homéopathie en France, 1830-1870: Évidence et arrière-plans* (Lyon: Centre Pierre Léon, Bioron, 1990).

31. On Europe, see Martin Dinges, ed., *Weltgeschichte der Homöopathie: Länder-Schulen-Heilkundige* (Munich: C. H. Beck, 1996); Phillip A. Nicholls, *Homoeopathy and the Medical Profession* (London: Croom Helm, 1988). On the United States, see Martin

Kaufman, *Homeopathy in America: The Rise and Fall of a Medical Heresy* (Baltimore: Johns Hopkins Press, 1971); William G. Rothstein, *American Physicians in the Nineteenth Century: From Sects to Science* (1972; Baltimore: Johns Hopkins University Press, 1985).

32. Jean Baptiste Bouillaud, "Rapport sur l'homoeopathie," *Bulletin Général de Thérapeutique,* 1835, *8:* 158-59, quotation on p. 159. It may not be a coincidence that the most venomous opposition to homeopathy in Paris came from people like J. B. Bouillaud (1796-1881) who were considered relentless in their advocacy of bloodletting, while Gabriel Andral (see below), who was allied with Pierre C. A. Louis (1787-1872) in urging less bleeding, was willing to test homeopathy. For a discussion of some of the parallel medical debates during this period, see Erwin H. Ackerknecht, *Medicine at the Paris Hospital, 1794-1848* (Baltimore: Johns Hopkins Press, 1967).

33. The most important such experimental administration was performed by Gabriel Andral (1797-1876) at the Pitié on at least eighty-nine patients: Gabriel Andral, "Expériences homéopathiques faites par M Andral à l'Hôpital de la Pitié," *Bulletin Général de Thérapeutique,* 1834, *5:* 318-22. Also see M. Lisfranc, "Discussion sur l'homoeopathie," *Gazette Médicale de Paris,* 2d ser., 1835, *3:* 189-90.

34. D. M. P. Pigeaux, "Étonnantes vertus homoeopathiques de la mie de pain: Expériences faites à l'Hôtel-Dieu," *Bulletin Général de Thérapeutique Médicale et Chirurgicale,* 1834, *6:* 128-31, quotation on p. 128. Pigeaux was a student of Trousseau and performed his experiment under his teacher's supervision. This report was published slightly before Trousseau and Gouraud's report (n. 30) and included overlapping data on some of the same patients; the two are clearly linked and should be considered a single experiment.

35. Trousseau and Gouraud, "Expériences homoeopathiques" (n. 30), p. 239.

36. Trousseau and Gouraud, "Expériences homoeopathiques" (n. 30) called them "pilules d'amidon" and Pigeaux (n. 34) called them "la mie de pain." The pills in both experiments were made in the same lot by the same pharmacist and included gum arabic as a binder.

37. Pigeaux, "Étonnantes vertus" (n. 34), p. 128.

38. Trousseau and Gouraud, "Expériences homoeopathiques" (n. 30), p. 239.

39. Ibid., p. 241.

40. Pigeaux, "Étonnantes vertus" (n. 34), p. 131.

41. John Forbes, "Homoeopathy, Allopathy and 'Young Physic,'" *Brit. & For. Med. Rev.,* 1846, *21:* 225-65, quotation on p. 251.

42. Forbes claimed that "several [experiments] have been made in the German hospitals" that involved "two sets of parallel cases of disease, the one treated homoeopathically, the other treated *apparently* in the same manner"--but with fictitious

globules in lieu of the real globules of homeopathy (ibid., pp. 239-40) (italics in original). Forbes himself performed a comparative homeopathic efficacy experiment using a sham-bread-pill arm: "Many years ago . . . we had occasion to treat an epidemic diarrhea of considerable violence but not dangerous. . . . [W]e put half of our remaining patients on a course of orthodox physic, and half on homoeopathic doses of flour . . . in the shape of bread-pills; and it puzzled us sadly to say which was the most successful treatment" (ibid., p. 249). A blind assessment of homeopathy using a bread pill seems to have taken place in St. Petersburg in 1834: see Otto Prokop and Ludwig Prokop, *Homöopathie und Wissenschaft: Eine Kritik des Systems* (Stuttgart: Ferdinand Enke, 1957), p. 22. Lisle used homeopathic and sham preparations in experiments that actually focused on the power of bread pills; his experiments had no concurrent controls: E. Lisle, "Feuilleton de l'homoeopathie orthodoxe," *L'Union Médicale,* 1861, *128:* 11-72.

43. This trial was not an isolated event; there was much research activity into homeopathy in the American Midwest at the time. For example, a large-scale open-label comparative trial with more than five thousand patients took place at Chicago's Cook County Hospital between 1881 and 1887: see Kaufman, *Homeopathy* (n. 31), pp. 150-51. For a related discussion of homeopathy's involvement with statistical comparative methods, see James H. Cassedy, *American Medicine and Statistical Thinking, 1800-1860* (Cambridge: Harvard University Press, 1984), pp. 124-30.

44. One can find many experiments in the mesmeric tradition that can also be described as "double-blind." For example, in 1818 a series of "stomach-seeing" experiments (the subjects read playing cards or written texts in a darkened room through their bellies) took place in Langenberg, Germany; the written material was put in opaque envelopes at other locations, and both the experimenter and the subject were unaware of the content: see Liselotte Moser, "Hypnotism in Germany," in Dingwall, *Abnormal Hypnotic Phenomena* (n. 21), 2: 136-48. See n. 49 below for a different type of homeopathic experiment on human subjects that also utilized a double-blind design at an even earlier stage of homeopathic history.

45. Samuel Potter and Eugene F. Storke, "Final Report of the Milwaukee Test of the Thirtieth Dilution," *Homeopathic Times: A Monthly Journal of Medicine, Surgery, and the Collateral Sciences,* 1880, *7* (12): 280-81.

46. For a full account of the experiment see Ted J. Kaptchuk, "Early Use of Blind Assessment in a Homoeopathic Scientific Experiment," *Brit. Homoeopathic J.,* 1997, *86:* 49-50.

47. At the same time, these physicians were also active participants in the pioneering innovations in pathological anatomy and physical diagnosis unfolding in orthodox medicine in Vienna. A few of these "half-homeopaths," as their "purist" homeopathic cousins called them, even had appointments at the Allgemeines Krankenhaus: see Hannelore Petry, "Die Wiener Homöopathie, 1842-1849" (Ph.D. diss., University of Mainz, 1954). For an English-language source containing some references to the homeopathic presence in Vienna's medical circles at this time, see Erna Lesky, *The Vienna Medical School of the Nineteenth Century* (Baltimore: Johns Hopkins University Press, 1976). These scientific homeopaths wanted to combine the best of both worlds

and use the most rigorous science to improve upon Hahnemann's vision. For a discussion of the battles between "scientific" homeopathy and "purist" homeopathy, see Anthony Campbell, *The Two Faces of Homeopathy* (London: Robert Hale, 1984).

48. Hahnemann initially had used toxological reports. His later provings were open-label and relied on the investigator's integrity to insure accuracy. See Samuel Hahnemann, *Organon of Medicine* (1921; New Delhi: B. Jain, 1980), pp. 209-10 (this is the 6th ed. of the *Organon,* which was published posthumously). Also see Franz Hartmann, "Hahnemann's Union for Proving Remedies," in Richard Haehl, *Samuel Hahnemann: Life and Work,* vol. 2, trans. Marie L. Wheeler (New Delhi: B. Jain, 1992). Thousands of symptoms were recorded in homeopathic tomes that dwarfed in size any kind of conventional medical text.

49. The impetus for this internal homeopathic self-correction may have also come from external confrontation directed not at homeopathy's medical efficacy, but toward the internal validity of its dilution and proving claims. In fact, the earliest such investigation I have found is also the earliest "double-blind" trial of a "substance" I have uncovered. This trial, which had the ambience of theater as much as of science, was organized by a journalist in Nuremberg beginning on 4 February 1835. See George Löhner, *Die homöopathischen Kochsalzversuche zu Nürnberg, Mit einem Anhang: Ein Beispiel homöopathischer Heilart* (Nuremberg, 1835): Common salt was first prepared according to Hahnemann's method of "potentiation" (Hahnemann believed that some inert substances such as salt became extremely powerful through the process of dilution and shaking). Then, fifty bottles were filled with potentized salt and another fifty with distilled snow, which served as a dummy control. The numerically coded contents were placed in a sealed envelope and the bottles were carefully "mixed up" ("gemischt," "gemengt" [p. 15]) to further conceal their identity. Fifty-five participants then received numbered vials. On 12 March at the Red Rooster Inn ("Gasthaus zum rothen Hahn" [p. 6]), the fifty participants who completed the study mostly reported that they had noticed nothing (nineteen had been taking homeopathic salt, and twenty-three taking snow water); of the other eight subjects, a few in each group had either cold symptoms or lower abdominal discomfort. Homeopaths criticized the trial because the participants had not followed the proper diet for a proving. The report explicitly stated that "the decisive *punctum saliens* [was]: to prevent the individual test persons from knowing when they are receiving certain homeopathic medications or certain nonmedicated trial substances. Even the person preparing and distributing the doses may not know, as in our experiments, what this [vial] or the other [vial] may contain" (pp. 23-24).

Other such challenges to homeopathy's internal validity also took place (e.g., the Milwaukee investigation described above actually included a second experiment with a similiar design). Oliver Wendell Holmes's famous denunciation of homeopathy also mentioned such a discussion of research methodology in Paris: Oliver Wendell Holmes, *Homoeópathy [sic] and Its Kindred Delusions: Two Lectures Delivered before the Boston Society for the Diffusion of Useful Knowledge* (Boston: William D. Ticknor, 1842), p. 44.

50. Philipp Anton Watzke, "Materialien zu einem physiologischen Umbau der Hahnemann'schen Arzneimittellehre. I: Die Koloquinte," *Österreichische Zeitschrift für Homöeopathie,* 1844, *1:* 1-151, quotation on p. 41.

51. Philipp Anton Watzke, "Wirkung des Kochsalzes im gesunden menschlichen und thierischen Körper--Unfreiwillige physiologische Kochsalz-prüfungen," *Österreichische Zeitschrift für Homöeopathie,* 1849, *4:* 13-129, quotation on p. 125.

52. Edith Heischkel, "Arzneimittelversuche in ärztlichen Vereinen um die Mitte des 19. Jahrhunderts," *Hippokrates: Zeitschrift für praktische Heilkunde,* 1955, *26:* 536-39. Also see Petry, "Die Wiener Homöopathie" (n. 47).

53. For example, between 1901 and 1903 the "scientific" camp conducted a proving coordinated at eleven centers through the Boston University School of Medicine (which was then homeopathic). This experiment, which probably was the largest proving ever performed, adopted placebo controls, which were "inert solutions [that] so resemble the tincture or dilutions to be employed . . . in dose, taste and color, that [the subject] will be unable to discriminate between the blank and the medicine" (Howard P. Bellows, *The Test Drug-Proving of the "O. O. & L. Society": A Re-Proving of Belladonna* [Boston: O. O. & L. Society, 1906], p. 25). Even the "purist" homeopaths adopted blind assessment. For example, the dean of the so-called classical uncompromising school of homeopathy, James Tyler Kent (1846-1916), spoke of blind assessment as a routine procedure in his turn-of-the-century writings: "The provers do not know what they are taking" (James Tyler Kent, *Lectures on Homeopathic Philosophy* [1900; Berkeley: North Atlantic, 1972], p. 185).

54. Although these experiments were often performed by prominent practitioners, they seem to have had little influence on ordinary practice and have generally been forgotten. Cf. Oswei Temkin, "Historical Aspects of Drug Therapy," in *Drugs in Our Society,* ed. Paul Talalay (Baltimore: Johns Hopkins Press, 1964), pp. 3-16.

55. Austin Flint, "A Contribution Toward the Natural History of Articular Rheumatism; Consisting of a Report of Thirteen Cases Treated Solely with Palliative Measures," *Amer. J. Med. Sci.,* 1863, *46:* 17-36, quotation on p. 21 (italics in original).

56. Henry G. Sutton, "Cases of Rheumatic Fever, Treated for the Most Part by Mint Water. Collected from the Clinical Books of Dr. Gull, with Some Remarks on the Natural History of that Disease," *Guy's Hospital Report,* 1865, *11:* 292-428, quotation on p. 392.

57. Ibid.

58. Josef Skoda (1805-81) and Ferdinand von Hebra (1816-80), the Viennese therapeutic nihilists, are reported to have performed "feigning treatment in some cases in order to demonstrate to [their] own satisfaction that [patients] could get well of themselves" (Fielding H. Garrison, *An Introduction to the History of Medicine* [1913; Philadelphia: Saunders, 1968], p. 434). Similiar events are implied by other historians: see Erwin H. Ackerknecht, *A Short History of Medicine* (Baltimore: Johns Hopkins University Press, 1982), p. 155; Max Neuburger, *The Doctrine of the Healing Power of Nature Throughout the Course of Time* (New York: New York Homeopathic College, 1933), p. 177.

59. Ian Hacking, "Telepathy: Origins of Randomization in Experimental Design," *Isis,* 1988, *79:* 427-51. Also see Stephen M. Stigler, *The History of Statistics: The Measurement of Uncertainty before 1900* (Cambridge: Harvard University Press, 1986), p. 253.

60. Peirce and Jastrow were interested in quantifying the relationship between physical stimuli and the mental experience of those stimuli, a field of research that had earlier received the hopeful name of *psychophysics* from Gustav Fechner (1801-87). They were improving on Fechner's earlier unblinded experiment. An even earlier series of blind assessments in psychophysics is described in Trudy Dehue, "Deception, Efficiency, and Random Groups: Psychology and the Gradual Origination of Random Group Design," *Isis*, 1997, *88:* 653-73.

61. Charles Sanders Peirce and Joseph Jastrow, "On Small Differences of Sensation," *Mem. Nat. Acad. Sci.,* 1884, *3* (1): 75-83, quotation on p. 79. Also see Stephen M. Stigler, "Mathematical Statistics in the Early States," *Ann. Statist.,* 1978, *6:* 239-65. To further prevent bias, Peirce and Jastrow also used a mathematical randomization scheme with playing cards to select the direction of their tests (whether to increase or decrease the weights).

62. From the beginning of the nineteenth century astronomers were acutely aware of what was called the *personal equation,* which represented a tendency for different observers to have a persistent and systematic variance in their observations. See Simon Schaffer, "Astronomers Mark Time: Discipline and the Personal Equation," *Science in Context,* 1988, *2* (1): 115-45.

63. A direct linkage to psychical research can easily be argued. For example, Peirce was recruited by his friend William James (1842-1910) to be a founding member of the American Society for Psychical Research (ASPR) in 1884, and the unremitting skeptic Jastrow was an original member of the ASPR's Scientific Advisory Council. They were both immersed in the telepathy debates, which affected their conventional work (and vice versa). In fact, the original publication of their study concluded by stating that their methodology had "highly important practical bearings" for the plausibility of telepathy, which "ought to be fully studied by the psychologist and assiduously cultivated by everyman" (Peirce and Jastrow, "Small Differences of Sensation" [n. 61], p. 83). Blinding very gradually became standard in some types of psychology investigations. An early example includes the adoption of "screens" reported in C. E. Seashore, "Measurements of Illusions and Hallucinations in Normal Life," *Stud. Yale Psychol. Lab.,* 1895, *3:* 1-67, quotation on p. 6.

64. The practice of adopting a new legitimate name for an unconventional phenomenon is a recurrent theme in the history of unconventional science. For example, the word *psychic* itself was invented in 1856 by Robert Hare (1781-1858)--a chemist at the University of Pennsylvania, and the inventor of the oxyhydrogen blowpipe--to dissociate his research on spiritualism from the taint of quackery: see James McClenon, *Deviant Science: The Case of Parapsychology* (Philadelphia: University of Pennsylvania Press, 1984), p. 5.

65. The SPR sought to "naturalize the supernatural by inserting into that framework" the methods and goals of scientific research (Janet Oppenheim, *The Other World: Spiritualism and Psychical Research in England, 1850-1914* [Cambridge: Cambridge University Press, 1985], pp. 152-53).

66. Making a clear distinction between "higher" mesmerism and spiritualism can be problematic. Spiritualism "surpassed" higher mesmerism but was also a direct continuation of the movement. See J. Stillson Judah, *The History and Philosophy of the Metaphysical Movements in America* (Philadelphia: Westminster, 1967), pp. 51-56; Fuller, *Mesmerism* (n. 20), pp. 69-104; R. Laurence Moore, *In Search of White Crows: Spiritualism, Parapsychology, and American Culture* (New York: Oxford University Press, 1977), pp. 9-11.

67. See R. Laurence Moore, "The Occult Connection? Mormonism, Christian Science, and Spiritualism," in *The Occult in America: New Historical Perspectives,* ed. Howard Kerr and Charles L. Crow (Urbana: University of Illinois Press, 1986), pp. 135-61. "Clairvoyant physicians" were not uncommon. See John Patrick Deveney, *Paschal Beverly Randolph: A Nineteenth-Century Black American Spiritualist, Rosicrucian, and Sex Magician* (Albany: State University of New York Press, 1997), p. 25.

68. Hacking, "Telepathy" (n. 59). For additional background, see Stewart Wolf, *Brain, Mind, and Medicine: Charles Richet and the Origins of Physiological Psychology* (New Brunswick, N.J.: Transaction, 1993).

69. Charles R. Richet, "La suggestion mentale et le calcul des probabilités," *Revue Philosophique,* 1884, *18:* 609-74, quotation on p. 635.

70. Ibid., p. 652. With time, Richet's blinding methods became more rigorous and he used envelopes and then double envelopes: Charles R. Richet, "Relation de diverses expériences sur la transmission mentale, la lucidité et autres phénomènes non explicables par les données scientifiques actuelles," *Proc. Soc. Psychical Res.,* 1888, *5:* 18-168.

71. Hacking, "Telepathy" (n. 59). Hacking, in his perceptive article on the subject of early randomization, does not mention an earlier, more primitive, mesmeric and hypnotism tradition in relation to this methodology. For example, in 1846 James Braid, in his tests of the "odic force" of Karl von Reichenbach (1786-1869), used what would now be called quasi-randomization methods to turn real or sham electromagnets on or off with "no regular order in the experiments" (James Braid, *The Power of the Mind Over the Body* [1846], reprinted in *Foundations of Hypnosis: From Mesmer to Freud,* ed. M. M. Tinterow [Springfield, Ill.: C. C. Thomas, 1970], p. 333). Other examples are easy to find. Also, homeopathy experiments attempted to use some method of "mixing" in their methodology (see n. 49). It should be noted that after Richet's experiments, the Society of Psychical Research quickly adopted blind assessment and randomization in its experiments. By 1889, telepathy experiments routinely selected numbers or cards "*drawn at random*" as a further precaution against subtle cues ("Messrs. Hansen and Lehmann on the Telepathic Problem," *J. Soc. Psychical Res.,* 1889, *9:* 113-30, quotation on p. 119 [italics in original]).

72. Hacking, "Telepathy" (n. 59). Also see McClenon, *Deviant Science* (n. 64). One of the earliest such efforts was at the Division of Psychical Research at Stanford University. Between 1912 and 1917 an avowed skeptic-scientist, John Edgar Coover (1872-1938), performed more than ten thousand trials on more than two hundred subjects, using cards selected from a deck. These experiments utilized a method of randomization, and always "the reagent sat with his back toward the experimenter, and in the experimental interval he closed his eyes" (John Edgar Coover, *Experiments in Psychical Research at Leland Stanford Junior University* [Stanford, Calif.: Stanford University Press, 1917], p. 54). Also see Seymour H. Mauskopf and Michael R. McVaugh, *The Elusive Science: Origins of Experimental Psychical Research* (Baltimore: Johns Hopkins University Press, 1980).

73. The fact that it is hard to determine when the mesmeric issues ceased to be unconventional and became an agenda within an intraorthodox debate exemplifies the idea that the boundary between conventional and irregular medicine is not necessarily sharp or fixed. In fact, preceding the hypnotism debates, Charcot was interested in "metallotherapy" (the utilization of magnets for curative purposes), which in many ways was an extension of mesmerism. Additionally, as in all the mesmeric debates, metallotherapy significantly involved blind assessment. For example, Dr. Landouzy at La Charité Hôpital in Paris, in cooperation with Charcot, investigated magnetic effects while blindfolding ("bander les yeux") his patients (L. Landouzy, "Relation d'un cas de léthargie provoquée par l'application d'un aimant," *Progrès Médical,* 1879, *7:* 60-62, quotation on p. 61). The English medical literature also had numerous reports of experiments on such therapy utilizing blinding and wooden decoys; e.g., A. Hughes Bennett, "Case of Complete Anaethesia of the Right and Partial Anaesthesia of the Left Side.--Experiments on Metalloscopy and Metallotheraphy," *Brit. Med. J.,* 1878, *2:* 759-61, quotation on p. 759. For a discussion of this entire episode and its effect on the later Salpêtrière-Nancy debate, see Anne Harrington, "Metals and Magnets in Medicine: Hysteria, Hypnosis, and Medical Culture in *fin-de-siècle* Paris," *Psychol. Med.,* 1988, *18:* 21-38.

74. See Christopher Goetz, Michel Bonduelle, and Toby Gelfand, *Charcot: Constructing Neurology* (New York: Oxford University Press, 1995).

75. Hippolyte Bernheim, *Suggestive Therapeutics: A Treatise on the Nature and Uses of Hypnotism,* trans. Christian A. Herter (New York: Putnam, 1897), p. 125. This is the English translation of *De la suggestion, et de ses applications à la thérapeutique* (Paris: O. Doin, 1886).

76. Alfred Binet and Charles Féré, "L'hypnotisme chez les hystériques," *Revue Philosophique,* 1885, *19:* 1-25, quotation on p. 4.

77. Bernheim attacked the experiments as not having sufficient precautions to exclude suggestion: Bernheim, *Suggestive Therapeutics* (n. 75), pp. 91-104. The Society of Psychical Research also criticized these experiments as having insufficient precautions to exclude suggestion: see Frederic W. H. Myers, "Report of the General Meeting," *J. Soc. Psychical Res.,* 1886, *2:* 443-55. The harshest contemporary critique of the experiments was provided by the Belgian psychology professor J. L. R. Delboeuf (1831-96) of the University of Liège, who after his visit to the Salpêtrière reported that the

experimenters announced "aloud what was going to happen," and that the magnet was a visible "heavy horseshoe" casually drawn from the pocket (Joseph R. L. Delboeuf, *Le magnétisme animal à propos d'une visite à l'École de Nancy* [Paris: Ancienne Librairie Germain Baillière, 1889], pp. 7-8). Delboeuf also replicated these experiments with the additional precaution of using both "false and true magnets [*avec de faux et avec de vrais aimants*]" and, like Bernheim, "without any magnets at all" (ibid., p. 19). (See p. 414.)

78. Bernheim, *Suggestive Therapeutics* (n. 75), p. 93. Even before the suggestion debates, Bernheim was involved in at least one blind assessment utilizing a sham "magnetic field" therapy that had claims akin to metallotherapy: see Robert C. Hillman, "A Scientific Study of Mystery: The Role of the Medical and Popular Press in the Nancy-Salpêtrière Controversy on Hypnotism," *Bull. Hist. Med.,* 1965, *39:* 163-82, quotation on pp. 169-70.

79. Joseph F. F. Babinski, "Recherches servant à établir que certaines manifestations hystériques peuvent être transférées d'un sujet à un autre sous l'influence de l'aimant," *Revue Philosophique,* 1886, *22:* 697-700, quotation on p. 700.

80. Myers, "General Meeting" (n. 77), provided an excellent firsthand description of the Babinski experiments. The weakness of the blindness was apparent, and Myers recommended sham magnets and sham metals as a necessary but missing component of any replication of Babinski's work. Myers also reported that Charcot was comfortable just repeating the experiment in his presence without troubling to use any intervening screen.

81. Additional cases can be found in Dingwall, *Abnormal Hypnotic Phenomena* (n. 21), and Gauld, *History of Hypnotism* (n. 20).

82. This episode began with a series of blind assessments performed in 1885 at the medical school of Rochefort under the direction of two professors. They were able to demonstrate "magnetically" transmitted drug effects: see Henri J. H. Bourru and Prosper F. Burot, *La suggestion mentale et l'action à distance des substances toxiques et médicamenteuses* (Paris: J. B. Baillière, 1887). The mishap occurred during a "nonpartisan" replication performed at the School of Naval Medical Officers in Rochefort under the direction of a Dr. Duplouy. A. T. Myers (1851-94), Frederic Myers's brother, described the incident: "Another gentleman during Dr. Duplouy's investigation made an experiment which was rendered especially important by a mistake. He had two similiar bottles in his pocket, both wrapped in paper; one contained cantharides, the other valerian; he chose the one which he thought contained cantharides and held it up to the patient; to his surprise the results which accompany valerian followed, and then he found that he had made a mistake and was holding the bottle containing valerian" (Arthur Thomas Myers, "On the Action of Drugs at a Distance," *J. Soc. Psychical Res.,* 1885, *2:* 58-62, quotation on p. 61).

A commission of the Academy of Medicine headed by Georges Dujardin-Beaumetz replicated versions of these experiments in what seems an intentional double-blind manner. Tubes with genuine drugs were matched with identical empty tubes; the tubes

were numbered and then covered with paper so that the experimenter could not tell whether substance or sham was being tested: Georges Dujardin-Beaumetz, "Sur l'action des médicaments à distance," *Bulletin Général de Thérapeutique Médicale et Chirurgicale,* 1888, *114:* 241-61. Most of the other independent replications seem to have been single-blind; e.g., J. Voisin, "Suggestion, auto-suggestion et vivacité du souvenir dans le sommeil hypnotique.--Action des médicaments à distance," *Revue de l'Hypnotisme,* 1888, *2:* 209-11.

83. Charles Édouard Brown-Séquard, "Des effets produits chez l'homme par des injections sous-cutanées d'un liquide retiré des testicules frais de cobaye et de chien," *Comptes Rendus de la Société de Biologie,* 1889, *41:* 419.

84. Ibid.

85. For example, the *Brit. Med. J.*'s report of Brown-Séquard's observations stated that "they [Brown-Séquard's observations] would require to be rigidly tested and fully confirmed by other self-experimenters before they were likely to meet with general acceptance" (*Brit. Med. J.,* 1889, *1:* 1416); no mention is made of the need to make blind assessments. Throughout the debate, Brown-Séquard's defense comprised long recitals of the effects of the extracts on hundreds of cases of such diseases as cancer and tuberculosis. See Charles Édouard Brown-Séquard, "On a New Therapeutic Method Consisting in the Use of Organic Liquids Extracted from Glands and other Organs," *Brit. Med. J.,* 1893, *1:* 1212-14. He never seemed to discuss blind assessment, with the exception of the one passing reference discussed below (see n. 89).

86. The popular press spoke of the "fountain of perpetual youth," "elixir of youth," the "Alchemist's Dream," and said that a "cult of injection spread like wildfire" (James M. D. Olmsted, *Charles-Édouard Brown-Séquard: A Nineteenth-Century Neurologist and Endocrinologist* [Baltimore: Johns Hopkins Press, 1946], pp. 210-11). Also see Michael J. Aminoff, *Brown-Séquard: A Visionary of Science* (New York: Raven Press, 1993); Merriley Borell, "Brown-Séquard's Organotherapy and Its Appearance in America at the End of the Nineteenth Century," *Bull. Hist. Med.,* 1976, *50:* 309-20.

87. Editorial, "Animal Extracts as Therapeutic Agents," *Brit. Med. J.,* 1893, *1:* 1279.

88. M. G. Variot, "Trois expériences sur l'action physiologique du suc testiculaire injecté sous la peau, suivant la méthode de M. Brown-Séquard," *Comptes Rendus de la Société de Biologie,* 1889, *41:* 451-54.

89. Brown-Séquard reported on the Variot experiments and the subsequent single episode of blind assessment in Charles Édouard Brown-Séquard, "Remarques à l'occasion du travail de M. Variot, sur les injections de liquide testiculaire chez l'homme," *Comptes Rendus de la Société de Biologie,* 1889, *41:* 454-55, quotation on p. 455. (Also see idem, "The Effects Produced on Man by Subcutaneous Injections of a Liquid Obtained from the Testicules of Animals," *Lancet,* 1889, *2:* 105-7, which also provides an English summary of Brown-Séquard's original French report.) This is the only mention of blind assessment by Brown-Séquard that I have been able to find.

90. At least one other blind assessment on Brown-Séquard's testicular extract was performed in France. It was a single-subject cross-over design (extract, then water, then extract): see Charles Éloy, *La méthode de Brown-Séquard* (Paris: J. B. Baillière, 1893), p. 47.

91. Gauld, *History of Hypnotism* (n. 20), p. 345.

92. Auguste Forel, "Das Verhältnis gewisser therapeutischer Methoden zur Suggestion," *Zeitschrift für Hypnotismus, Suggestionstherapie, Suggestionslehre und verwandte psychologische Forschungen,* 1893/94, *2:* 385-90, quotation on p. 390. The article is based on a lecture presented in 1894 at the 60th Meeting of the German Biological Researchers and Doctors in Vienna. Forel visited Bernheim in 1887 and was considered an accomplished hypnotist. See Henri F. Ellenberger, *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry* (New York: Basic Books, 1970), p. 88.

93. Forel, "Verhältnis gewisser therapeutischer Methoden" (n. 92), p. 388. Forel specifically stated that "one speaks of results obtained without the patient's kowledge; but how can a patient not notice an injection? One should also make comparisons with injections of other substances" (ibid.).

94. Fritz Pregl, "Zwei weitere ergographische Versuchsreihen über die Wirkung orchitischen Extraktes," *Archiv für die gesamte Physiologie,* 1896, *62:* 379-99, quotation on p. 387. After an initial baseline assessment, one student received the extract and another received sham glycerin. Later, in a crossover manner, they received the oppoosite substances. Pregl states that the blinding precautions included mixing salt with the glycerin so that "its injection . . . more or less produced the same burning and pressure sensation" as the verum (ibid., p. 385). As a further safeguard, syringes were filled in another room that the subjects could not enter, and subjects were treated separately (ibid., p. 386).

95. Waclaw Sobiera´nski, "Über den Einfluss des pharmakologischen Mittels auf die Muskelkraft der Menschen," *Centralblatt für Physiologie,* 1896, *5:* 126-27. His research papers do not use the word *blind,* but that is clearly his intention. This particular paper studied the effects of cocaine and caffeine on muscle strength as measured by "exhaustion curves" using ergography.

96. Waclaw Sobiera´nski, "O wplywie ´srodków farmakologicznych na sile me´sniowa ludzi," *Gazeta Lekarska (Warsaw),* 2d ser., 1896, *16:* 86-95, quotation on p. 90. A summary of his research efforts was reported in a Polish-language lecture in Warsaw on 24 September 1895, though most of his scientific papers were written in German.

97. Ibid., p. 89. Even when blind assessment was acknowledged as important, its implementation was erratic. For example, in 1899 a Hanover military officer and professor named Schumburg was interested in rumors that cola kept French soldiers marching while their horses and mules were too tired to even eat. In his experiments, soldiers were given either a sham substance or cola-extract on different days; for unexplained reasons, however, when he evaluated caffeine, tea, mate, and alcohol he used no sham intervention: Wilhelm A. E. F. Schumburg, "Ueber die Bedeutung von Kola, Caffee, Thee,

Mate und Alkohol für die Leistung der Muskeln," *Archiv für Anatomie und Physiologie,* 1899, *5:* 289-313, quotation on p. 293. Schumburg's other experiments include Wilhelm A. E. F. Schumburg, "Ueber die Bedeutung des Zuckers für die Leistungsfähigkeit des Menschen," *Zeitschrift für diätetische und physikalische Therapie,* 1899, *2:* 185-88. The German interest in blind assessment received a boost with the discovery of subliminal stimuli in 1908; see Gerd Gigerenzer, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Krüger, *The Empire of Chance: How Probability Changed Science and Everyday Life* (Cambridge: Cambridge University Press, 1989), p. 87. The psychophysics movement may have also contributed to the German pharmacological interest in blind assessment. See Dehue, "Deception" (n. 60).

98. William H. R. Rivers, *The Influence of Alcohol and Other Drugs on Fatigue: The Croonian Lectures Delivered at the Royal College of Physicians in 1906* (London: Edward Arnold, 1908), p. 19.

99. Ibid., p. 20.

100. Harry Levi Hollinsworth, "The Influence of Caffein on Mental and Motor Efficiency," *Arch. Psychol.,* 1912, *22:* 1-166.

101. Albion W. Hewlett, "Clinical Effects of 'Natural' and 'Synthetic' Sodium Salicylate," *JAMA,* 1913, *61:* 319-21, quotation on p. 319. Of the 82 physicians recruited, only 27 reported back on 230 separate observations. The word *blind* does not appear in the paper. Hewlett (1874-1925) was a professor at the University of Michigan Medical School.

102. Hewlett described the rationale for the blinding as related to "the personal equations of different observers [and] the tendency to bias" (ibid., p. 319). (Also see n. 62.) In general, when the European concern for blind assessment was translated into English, the Continental preoccupation with suggestion was omitted. This can be seen in the writings of Sollmann, who was German-born and did postgraduate work in Germany. When he made an early English-language plea for blind assessment he spoke only of natural history: even "the best type of clinical reports . . . lack one important essential, namely, an adequate control of the natural course of the disease. . . . Since this cannot be controlled directly, it must be compensated indirectly. . . . The . . . method consists in the attempt to distinguish unknown preparations by their effects--the method that might be called . . . the 'blind test'" (Torald Sollmann, "The Crucial Test of Therapeutic Evidence," *JAMA,* 1917, *69:* 198-99, quotation on p. 199). Sollmann's call seems to have been mainly ignored, and it is unclear to what extent he acted on his own prescription. Harry Marks, *Progress of Experiments* (n. 3), p. 36, cites at least one other Sollmann blind assessment.

103. David I. Macht, N. B. Herman, and Charles S. Levy, "A Quantitative Study of the Analgesia Produced by Opium Alkaloids, Individually and in Combination with Each Other, in Normal Man," *J. Pharm. Exp. Ther.,* 1916, *8:* 1-37, quotation on p. 7.

104. The introduction of diphtheria antitoxin in 1894 was accompanied by much hope and also doubt. Inexact identification of the illness, difficulty with the timing of the serum's

administration, and general utilization of historic or other-site comparisions when the virulence of the disease varied from season to season and place to place, made evaluation difficult and engendered lack of enthusiasm, dissension, and open skepticism. See Rothstein, *American Physicians* (n. 31); Evelynn Maxine Hammonds, "The Search for Perfect Control: A Social History of Diphtheria, 1880-1930" (Ph.D. diss., Harvard University, 1993).

105. The actual design of the experiment was more complex. In 1911, Bingel admitted only adults into the trial. In 1912, he lowered the age of the subjects, and by 1913 he was allocating patients regardless of age or severity of disease. Eventually, 90 percent of his patients were children. See Adolf Bingel, "Über Behandlung der Diphtherie mit gewöhnlichem Pferdeserum," *Deutsches Archiv für klinische Medizin,* 1918, *125:* 284-332.

106. Bingel's exact words were: "in order to make as objective a test as possible . . . [I asked for evaluations from] the attending physicians . . . without explaining to them the nature of the test serum" (ibid., p. 288). Bingel explicitly called his procedure a "blind" method (his quotes) and recommended its adoption by his colleagues (ibid.). He continued to perform experiments with diphtheria antitoxin even during World War II: Adolf Bingel,"Wirkt das Diphtherieheilserum bei der menschlichen Diphtheriekrankheit spezifisch durch seinen Antitoxingehalt oder unspezifisch?" *Deutsche medizinische Wochenschrift,* 1949, *74:* 101-3; idem, "Zur umstrittenen Wirkung des Di.-antitoxins beim Menschen," ibid., 1950, *47:* 1585-87. He also received explicit support for his position from another clinical trial performed on "more than 450 patients" in which "treatment was alternatively either genuine serum antitoxin or 'empty' horse serum [*Pferde-Leer-Serum*]" (A. Hottinger and D. Töpfer, "Über den Wert der Serumtherapie bei Diphtherie, insbesondere bei der malignen, toxischen Form," *Zeitschrift für Kinderheilkunde,* 1933, *54:* 505-40, quotation on p. 513). Not surprisingly, Bingel's work has been entirely forgotten, while the famous 1898 open-label diphtheria experiment conducted by Johannes A. G. Fibiger (1867-1928) at Blegdam's Hospital in Copenhagen with 488 patients is a more commemorated episode in the official histories of diphtheria and of clinical trials. (It is worth speculating that the "peculiar" results that some blind assessments produced may have contributed to the resistance to the method. Warner has pointed out a similiar predicament when orthodox medicine resisted the "numerical method . . . as a revealer of therapeutic truth" because it could present homeopathy in a favorable light [John Harley Warner, *The Therapeutic Perspective: Medical Practice, Knowledge, and Identity in America, 1820-1885* (Cambridge: Harvard University Press, 1986), pp. 202-3].)

107. His methods included rules to "exclude [*Ausschaltung,* switch off] suggestive or other irrelevant [*unsachlicher*] factors in the blinded test [*unwissentliche Versuchsanordnung*]. . . . The medications must be given to the patient in a shape or wrapping that does not permit recognition of their special character or purpose, they must be camouflaged. . . . Even during the preobservation period . . . [one must use] a fake medication treatment using inert substances" (Paul Martini, *Methodenlehre der therapeutischen Untersuchung* [Berlin: Julius Springer, 1932], p. 8).

108. E.g., Paul Martini, "Klinische Untersuchung des sog. Herzhormone bei Angina

pectoris," *Deutsche medizinische Wochenschrift,* 1932, *58:* 569-72, quotation on p. 570. Interestingly enough, in 1938-39 Martini applied the blind assessment method to a series of historic déjà-vu assessments of homeopathic remedies; e.g., Paul Martini, L. Bruckmer, Karl Dominicus, A. Schulte, and A. Stegemann, "Homöopathische Arzneimittel--Nachprüf-ungen," *Naunyn Schmeidebergs Archiv für experimentelle Pathologie,* 1939, *191:* 141-71.

109. A. Krumeich, "Klinische Prüfung der Wirkung von Arzneimitteln auf den erhöhten Blutdruck," *Deutsches Archiv für klinische Medizin,* 1933, *173:* 527-40, quotation on p. 527.

110. R. Schwenk, "Über den Wert des Histidins bei der Behandlung des Ulcus ventriculi und duodeni," *Deutsches Archiv für klinische Medizin,* 1941, *189:* 139-58, quotation on p. 139.

111. Generally speaking, American and British mainstream physicians lagged behind their Continental colleagues in conceding any ground to suggestion, hypnotism, or nonmaterial agency. An example would be the late recognition that the British Medical Association and the American Medical Association accorded hypnotism: cf. Subcommittee of the Psychological Medicine Group Committee of the British Medical Association, "Medical Use of Hypnosis," *Brit. Med. J.,* Suppl., 1955, *1:* 190-93; Council on Mental Health, "Medical Use of Hypnosis," *JAMA,* 1958, *168:* 186-89. Also see n. 101.

112. Although there were early examples (e.g., the scurvy trial conducted by James Lind [1716-1794]), widespread attention to the issue began only with Louis Pasteur's adoption of concurrent controls in his anthrax experiments on animals (1881). Fibiger's 1898 diphtheria experiment (see n. 105) is often credited as the first carefully performed experiment on humans, where treatment was allocated impartially on entrance (by alternative assignment). But the principle did not begin to take hold for more than another generation. See Alvan R. Feinstein, *Clinical Epidemiology: The Architecture of Clinical Research* (Philadelphia: Saunders, 1985), pp. 685-86; and cf. Lilienfeld, "*Ceteris paribus*" (n. 1). The emerging perception that random assignment could eliminate bias in treatment groups (which began to have an effect in medical research with the pioneering work of Major Greenwood [1880-1949] and Udny Yule [1871-1951] on typhoid and cholera) also helped ignite an awareness of the need for equivalent control groups.

113. Austin Bradford Hill, "The Clinical Trial," *Brit. Med. Bull.,* 1951, *7* (4): 278-82, quotation on p. 281.

114. "Clinical Trials of New Remedies," *Lancet,* 1931, *2:* 304.

115. James Burns Amberson, B. T. McMahon, and Max Pinner, "A Clinical Trial of Sanocrysin in Pulmonary Tuberculosis," *Amer. Rev. Tuberc.,* 1931, *24:* 401-35, quotation on p. 429.

116. Ibid., p. 406. Treatment allotment was by "flip of the coin," and neither the word *blind* nor the word *placebo* appeared in the paper.

117. Lilienfeld's reading ("*Ceteris paribus*" [n. 1], p. 17) of the paper makes him consider this trial "double-blind," which seems plausible.

118. William Evans and Clifford Hoyle, "The Comparative Value of Drugs Used in the Continuous Treatment of Angina Pectoris," *Quart. J. Med.,* n.s., 1933, *26:* 311-38, quotation on p. 336. The report emphasized the spontaneous variation of the disease as the reason to adopt placebo control but did mention "mental suggestion" as a possible explanation for the positive results obtained through use of the placebo (ibid., p. 335). This paper used the word *placebo,* which at the time was very unusual in a research report.

119. The most important of these experiments was a series of large-scale trials concerning various treatments and preventive measures for the common cold, performed on student subjects under Harold S. Diehl (1891-1973) at the University of Minnesota. The reports included Harold S. Diehl, "Medicinal Treatment of the Common Cold," *JAMA,* 1933, *101:* 2042-49; Harold S. Diehl, A. B. Baker, and Donald W. Cowan, "Cold Vaccines: An Evaluation Based on a Controlled Study," *JAMA,* 1938, *111:* 1168-73; and Donald W. Cowan, Harold S. Diehl, and A. B. Baker, "Vitamins for the Prevention of Colds," *JAMA,* 1942, *120:* 1268-71. The placebo controls were designed to show "how much improvement should be considered as due to spontaneous recovery" (Diehl, "Medicinal Treatment" [n. 119], p. 2044). The 1933 experiment trial was "double-blind": "The ratings were made by me and independently by another physician without either of us knowing what medication had been given to the person making the report" (ibid., p. 2043). Other pre-World War II examples of placebo controls used to shape a no-treatment group include Ben Z. Rappaport, Michael Zeller, and Emanuel Padnos, "Ragweed Oral Pollen Therapy Compared with Oral Placebo," *JAMA,* 1940, *115:* 25-27; and George V. LeRoy, "The Effectiveness of the Xanthine Drugs in the Treatment of Angina Pectoris," *JAMA,* 1941, *116:* 921-25 (both of these experiments took place in 1939).

120. It seems that as late as 1933, Gold and his team were still unfamiliar with and had never implemented an experiment with blind assessment. See, e.g., Harold L. Otto, Harry Gold, and Charles R. Messeloff, "Studies on Digitalis in Ambulatory Patients with Cardiac Disease," *Arch. Intern. Med.,* 1933, *52:* 725-38. In 1935, Gold co-authored a study using the method of the "blind test" (quotation marks in the original) comparing two active substances (Ella M. Hediger and Harry Gold, "U.S.P. Ether from Large Drums and Ether from Small Cans Labeled 'For Anesthesia'," *JAMA,* 1935, *104:* 2244-48, quotation on p. 2245.) "Those administering the anesthetics . . . were unaware of the source of the ether and identified the specimens in terms of code numbers." (ibid.)

121. There were already a few eloquent English statements from medical brahmins that acknowledged "suggestion" and the fallout of the Salpêtrière-Nancy debate. In his last literary work, Charcot himself conceded defeat and spoke of "a confidence, a credulity [and] receptivity of suggestion" as being responsible for the healings that occur at famous religious shrines (Jean-Martin Charcot, "The Faith-Cure," *New Rev.,* 1893, *8:* 18-31, quotation on p. 19; this appeared in French as "La foi qui guérit," *Archives de Neurologie,* 1893, *25:* 72-87). Osler repeated Charcot's argument in discussing religious

healing: William Osler, "The Faith That Heals," *Brit. Med. J.,* 1910, *1:* 1470-72. But such sentiments were generally restricted to grandiose discussions of the "art" of medicine and were not seen as having research relevance.

122. Harry Gold, Nathaniel T. Kwit, and Harold Otto, "The Xanthines (Theobromine and Aminophylline) in the Treatment of Cardiac Pain," *JAMA,* 1937, *108:* 2173-79, quotation on p. 2178.

123. Ibid., p. 2177. Actually, Gold's experiment had one other crucial difference from the London experiment. Gold was also concerned with physician bias, and his experiment could almost be considered double-blind: "in a further attempt to eliminate the possibility of bias, the questioner usually refrained from informing himself as to the agent that had been issued until after the patient's appraisal" (ibid., p. 2175). The study seems to have begun with a single-blind design and evolved into a double-blind one. It should be noted that it was Gold and his colleagues who in 1950 seem to have been the first to use the now-established phrase *double-blind test:* see Theodore Greiner, Harry Gold, McKeen Cattel, Janet Travell, Hyman Bakst, Seymour Rinzler, et al., "A Method for the Evaluation of the Effects of Drugs on Cardiac Pain in Patients with Angina on Effort," *Amer. J .Med.,* 1950, *9:* 143-55, quotation on p. 146.

124. While Gold's work eventually became critical for the acceptance of placebo controls among his peers in the biomedical community, the notion that the mind and belief could produce "medical" outcomes had already made inroads into the English-language research literature. In 1936, two single-blind clinical trials on ulcer disease used saline injections as controls. The positive results of the sham were explained as partly due to "psychic effects" (David J. Sandweiss, "Treatment of Gastroduodenal Ulcer with Histidine Monohydrochloride [Larostindin]," *JAMA,* 1936, *106:* 1452-59) or "suggestion" (C. A. Flood and C. R. Mullins, "Treatment of Peptic Ulcer by Means of Injections," *Am. J. Dig. Dis.,* 1936, *3:* 303-5). Also in 1930, a placebo was given to 40 hypertensive patients. This experiment had no active treatment. The positive outcomes were explained as "the suggestion inherent in any drug" (David Ayman, "An Evaluation of Therapeutic Results in Essential Hypertension," *JAMA,* 1930, *95* (4): 246-49, quotation on p. 249). Ayman explicitly states that this idea is a personal opinion. Additionally, the orthodox anti-quackery crusades of the 1920s often used sham controls and may have contributed to this increased mainstream recognition of the imagination and subconscious bias. For example, see discussions in: Anon., "Two Electronic Diagnoses: The Reactions of a Guinea-Pig and Sheep to the Reaction of Abrams," *JAMA,* 1922, *79* (27): 2244-48; Austin C. Lescarboura, "Our Abrams Verdict. The Electronic Reactions of Abrams and Electronic Medicine in General Found Utterly Worthless," *Scientific American,* 1924, *131* (3): 158-160, 220-22; Arthur K. Cramp, "Some Bald Facts: Professor Scholder Appeals to an Ancient Weakness," *Hygeia,* 1927, *5:* 497-99.

125. The most direct pre-World War II influence of Gold seems to have been on another New York team at Mt. Sinai Hospital, which replicated Gold's experiments on the xanthines; their report's analysis followed the exact wording of the Gold paper and explicitly mentioned "confidence," "encouragement," and the patient-physician relationship: Arthur M. Master, Harry L. Jaffe, and Simon Dack, "The Drug Treatment of Angina Pectoris Due to Coronary Artery Disease," *Amer. J. Med. Sci.,* 1939, *197:* 774-

82, quotation on p. 774. Another team, at Boston's Beth Israel Hospital (with connections to Harry Gold), adopted sham saline injections as a control in a rheumatoid and osteoarthritis experiment; again, their discussion saw the control as necessary for both "the tendency to natural remission in chronic arthritis . . . [and] the psychological effect of the injection itself" (Nathan Sidel and Maurice I. Abrams, "Treatment of Chronic Arthritis: Results of Vaccine Therapy with Saline Injections Used as Controls," *JAMA,* 1940, *114:* 1740-42, quotation on p. 1742. By 1942, Harold Diehl (mentioned in n. 118) had also explicitly adopted the idea that a control group could experience "psychologic effects" (Harold S. Diehl, "Abstract of Discussion," *JAMA,* 1942, *120:* 1270-71, quotation on p. 1271).

126. E.g., see Otho B. Ross, "Use of Controls in Medical Research," *JAMA,* 1951, *145:* 72-74, which presented a quantitative discussion of the lack of well-controlled comparative clinical studies in the most prestigious medical journals between January and June of 1950. Ross's criteria emphasized concurrent controls, and he introduced the idea of randomization. Although his two examples of well-controlled trials used a sham interventon for concurrent treatment, he did not explicitly mention blind assessment as a criterion for a good trial.

127. Much has been written on R. A. Fisher. Examples of different approaches to his influence include C. Radhakrishna Rao, "R. A. Fisher: The Founder of Modern Statistics," *Statist. Sci.,* 1992, *7:* 34-48; and F. Yates, "Sir Ronald Fisher and the Design of Experiments," *Biometrics,* 1964, *20:* 307-21.

128. R. A. Fisher, *The Design of Experiments* (Edinburgh: Oliver and Boyd, 1935), p. 47.

129. See Gerd Gigerenzer, "Probabilistic Thinking and the Fight against Subjectivity," in *The Probabilistic Revolution,* vol. 2: *Ideas in the Sciences,* ed. Gerd Gigerenzer and Mary S. Morgan (Cambridge: MIT Press, 1987), pp. 11-33.

130. Gigerenzer et al., *Empire of Chance* (n. 97), p. 211.

131. Harry M. Marks, "Notes for the Underground: The Social Organization of Therapeutic Research," in *Grand Rounds: One Hundred Years of Internal Medicine,* ed. Russell C. Maulitz and Diana E. Long (Philadelphia: University of Pennsylvania Press, 1988), pp. 319-20.

132. Ibid., p. 319. I am indebted to Dr. Marks for sharing a prepublication version of portions of *Progress of Experiment* (n. 3). Many of the insights in the RCT section of this paper are derived from this pioneering work, esp. pp. 136-63.

133. Cf. David Armstrong, "Clinical Sense and Clinical Science," *Soc. Sci. Med.,* 1977, *11:* 599-601.

134. Austin Bradford Hill, "Suspended Judgement: Memories of the British Streptomycin Trial in Tuberculosis. The First Randomized Clinical Trial," *Controlled Clin. Trials,* 1990, *11:* 77-79, quotation on p. 77.

135. Austin Bradford Hill, "The Clinical Trial," *N. Engl. J. Med.,* 1952, *247* (4): 113-19, quotation on p. 118. The article cites a source for this criticism that is not traceable.

136. Robert C. Batterman, "Appraisal of New Drugs" [Letter], *JAMA,* 1955, *158:* 1547.

137. Mainland spoke of researchers' "antagonism to statistics" and experimenters as "long resistant to statistical tests" (Donald Mainland, "The Use and Misuse of Statistics in Medical Publications," *Clin. Pharmacol. Therap.,* 1960, *1:* 411-22, quotations on pp. 411, 412). Reid described physicians as feeling that "nothing [was] . . . more depressing than . . . 'the repellent symbolism' of the mathematical statistician" (D. D. Reid, "Statistics in Clinical Research," *Ann. New York Acad. Sci.,* 1950, *52:* 931-34, quotation on p. 931). Also see Donald Mainland, "The Clinical Trial--Some Difficulties and Suggestions," *J. Chron. Dis.,* 1960, *11:* 484-96.

138. Reid, "Statistics" (n. 137), p. 933. Also see, e.g., Walter Modell and Raymond W. Houde, "Factors Influencing Clinical Evaluation of Drugs," *JAMA,* 1958, *167:* 2190-99. A description of an earlier scientific situation applies here. Suddenly, the physician was no longer safe from "forces working on him that would shift his utterances out of correspondence with reality" (Steven Shapin, *A Social History of Truth: Civility and Science in Seventeenth-Century England* [Chicago: University of Chicago Press, 1996], p. xxvii).

139. This transition in understanding the "placebo" is obvious in the medical literature but is rarely discussed. See Kaptchuk, "Powerful Placebo" (n. 2).

140. Harry Gold and his colleagues were instrumental in organizing two conferences that advocated blind assessment in order to control for this newly detected placebo effect: see Conferences on Therapy, "The Use of Placebos in Therapy," *New York J. Med.,* 1946, *46:* 1718-27; Conference on Therapy, "How to Evaluate a New Drug," *Amer. J. Med.,* 1954, *17:* 722-27.

141. Hacking points out that probability has had two distinct functions: an epistemological one having to do with credibility and with "assessing reasonable degrees of belief in propositions," and a statistical one having to do with "stochastic laws of chance" (Ian Hacking, *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction, and Statistical Inference* [Cambridge: Cambridge University Press, 1975], p. 12). A similiar dichotomy could be said to apply to randomization itself, while blind assessment would have only a credibility dimension.

142. Cf. Richard Doll, "Development of Controlled Trials in Preventive and Therapeutic Medicine," *J. Biosoc. Sci.,* 1991, *23:* 265-78. One can see Hill still struggling with assuring genuine randomization in his 1951 description of the 1948 streptomycyin trial. In this trial, only the assessment radiologists were blind to intervention. In order to ensure successful randomization Hill had to resort to enforced secrecy, which can be as difficult with medical personnel as it is with anyone. Hill stated that "the allocation of the patient to treatment or control is kept secret from the clinician until after . . . [the] patient's admission. Thus he can proceed to that decision . . . without any fear of bias" (Hill, "Clinical Trial" [n. 113], p. 280).

143. Stephen Senn, "A Personal View of Some Controversies in Allocating Treatment to Patients in Clinical Trials," *Statist. Med.,* 1995, *15:* 2667. Cf. Feinstein, *Clinical Epidemiology* (n. 112), p. 688; and Peter Armitage, "The Role of Randomization in Clinical Trials," *Statist. Med.,* 1982, *1:* 347.

144. Hill, "Clinical Trial" (n. 135), p. 117.

145. Eugene F. DuBois, "The President's Address," *Trans. Assoc. Amer. Phys..* 1939, *54:* 1-5, quotation on p. 5. DuBois was a close senior associate of Harry Gold at Cornell-New York Hospital; in his address he educated physicians regarding the availability and potential value of blind assessment, but he still was not advocating its universal adoption.

146. The words *blind* or *double-blind* or *double unknowns* are often kept in quotation marks in biomedical journals well into the mid-1950s to denote the novelty of the technique. E.g., see Louis Lasagna, John M. von Felsinger, and Henry K. Beecher, "Drug-Induced Mood Changes in Man," *JAMA,* 1955, *157:* 1006-20; Henry K. Beecher, "Appraisal of Drugs Intended to Alter Subjective Responses, Symptoms," ibid., *158:* 399-401.

147. Ian Hacking, "Statistical Language, Statistical Truth and Statistical Reason: The Self-Authentification of a Style of Scientific Reasoning," in *The Social Dimensions of Science,* ed. Ernan McMullin (Notre Dame: University of Notre Dame Press, 1992), pp. 130-57; Hacking describes this style as follows: "the truth is what we find out in such and such a way. We recognize it as truth because of how we find it out. And how do we know that the method is good? Because it gets at the truth" (p. 135).

148. Ibid., p. 132.

149. Obviously, many of the authors cited in this essay--such as Marks, Gigerenzer, Porter, and Matthews--are actively engaged in this critical analysis. Many other sociologists and historians could be mentioned. On the scientific side, many debates are also raging. Concerning blind assessment, it should be mentioned that there is already a small literature concerned with critically examining the unintended consequences of blind assessment and the a priori assumptions embedded in the methodology. Important examples of this literature include Irving Kirsch and Michael J. Rosadino, "Do Double-Blind Studies with Informed Consent Yield Externally Valid Results?" *Psychopharmacology,* 1993, *110:* 437-42; Irving Kirsch and Lynne J. Weixel, "Double-Blind versus Deceptive Administration of a Placebo," *Behav. Neurosci.,* 1988, *102:* 319-23; Mauro Moscucci, Louise Byrne, Michael Weintraub, and Christopher Cox, "Blinding, Unblinding, and the Placebo Effect: An Analysis of Patients' Guesses of Treatment Assignment in a Double-Blind Clinical Trial," *Clin. Pharmacol. Therap.,* 1987, *41:* 259-65; Sydnor B. Penick and Seymour Fisher, "Drug-Set Interaction: Psychological and Physiological Effects of Epinephrine under Differential Expectations," *Psychosom. Med.,* 1965, *27:* 177-82; J. H. Noseworth, G. C. Ebers, M. K. Vandervoort, R. E. Farquhar, E. Yetisir, and R. Roberts, "The Impact of Blinding on the Results of a Randomized, Placebo-Controlled Multiple Sclerosis Clinical Trial," *Neurology,* 1994, *44:* 16-20. In fact, reading this literature was the incentive for me to undertake this examination of the

history of blind assessment.

.