VANGUARD OF AN EVIDENCE-BASED SOCIETY:
THE EXPERIMENTAL MANDATE OF THE NO CHILD LEFT BEHIND ACT

full paper by Tytus Wilam

Department of Sociology, University of Chicago

Why evidence-based policy? Why the rapid ascendence of experimental development economics? Why increasing focus on What Works in philanthropic and governmental funding of NGOs? Where did hierarchies of evidence, meta-analyses and the "gold standard" of the RCT come from? Evidence-based policy and practice movement is making a strong argument that we should live in an "evidence-based society" where policy is based on best available quantitative evidence. This vision is already shaping policies and education of future policy makers ("Harris Brochure" 2020); yet we don't know where this overt emphasis on using good evidence, suggesting that earlier expert advice was not based on evidence, came from and how it gained its present prominence. This paper attempts to begin filling this gap.

Evidence-based policy arose simultaneously in, among others, social work, education, criminal policy. Since the timing of the rises of evidence-based thinking is correlated across applied disciplines, the rises either share a cause or are one and the same event, i.e., have caused one another. If they share a cause, then explaining any one of the events would give us insight into the mechanism(s) behind all events. If, on the other hand, the rise of evidence-based thinking across domains is a singular event, then we should try to begin with a central part and then build our understanding of the whole event outwards.

In this paper I explain why evidence-based thinking arose in one domain at one time—George W. Bush's early childhood reading teaching—with the hope learning something about the evidence-based movement in general. We have reasons to believe that education policy under George W. Bush is central. First, it is in relation to education, and specifically reading, that the phrase "scientifically based research" first appears in the congressional record, and it did so in the run-up to NCLB. Second, some of the institutions we associate with the evidence-based movement—e.g., Campbell Collaboration, What Works Clearinghouse—first arose in education

around the time of Bush's presidency. Third, teaching reading was extremely high profile politically at the time, more so than climate change and smoking. Fourth, Bush policies were created for education, but created experimental mandates and funding structures that promoted evidence-based research more broadly, with spillovers to other domains. Fifth, "evidence based policy" becomes mainstream shortly after NCLB (Erickson and Gutierrez 2002, Graham 2002, Mosteller and Boruch 2002, Paige 2002, Slavin 2002, Slavin 2003), with the NCLB contributing to that increase in popularity.

I focus on a relatively narrow question: why did the No Child Left Behind Act (NCLB) of 2001 conditioned school funding on the use of instructional materials backed by "scientifically based research" and why did the Educational Sciences Reform Act (ESRA) of 2002 require that funded research be conducted according to "scientifically based research standards"?

Other approaches and their limitations

The questions I raise have been raised before but previous answers are limited in scope, and, I argue, fall short in one of three ways: they ignore knowledge, they ignore politics, or they are not explanatory. First, many accounts of the NCLB and ESRA do not take knowledge into account. The passage of the act is presented exclusively in terms of political factors (a successful coalition for outcomes-based education of business, civil rights activists, and parts of the parties in Rhodes 2012, personal connection of Bush to reading guru Reid Lyon in Teske 2009, p.118). These explanations do not engage the "supply side" of where policy ideas originated, neglect the agency of researchers, and explain only the more salient accountability part of NCLB, while neglecting the experimental mandate. Those studies are cynical; they do not believe in an

autonomous domain of knowledge. With science studies I will instead treat knowledge as something that is produced and supported by networks pf human and non-human actants (Latour 1987), and that in turn supports and makes possible the existence of other networks of actants. Second, the accounts of the Reading Wars often present a story of progressive growth of knowledge about reading (e.g., Kim 2008, Manna and Petril 2008): with newer, better science we got newer, better laws. And so, we read that after the findings of the National Reading Panel (NRP) showed what science really knows about reading, the legislators realized that and begun asking for interventions grounded in "scientifically based reading research." These accounts: (a) tell the story from the perspective of the victors and thereby violate the symmetry principle (Latour 1987), (b) treat science as unproblematically cumulative, without mentioning the work of synthesizing that goes into making science cumulative, and (c) fail to notice the political aspects of knowledge production, such as the intentions of persons funding the science. This is also the framing accepted by persons writing within the evidence-based movement. Third, some accounts tell a history of "evidence-based" but stop short of providing an explanation. We learn about the legislative history of the phrase "scientifically-based research" (Towne 2003, Manna and Petril 2008), but we do not learn why it appeared at that, and not another, time and place; alternatively, we get histories that begin where this paper ends (e.g., Baron 2018).

In this paper I do not focus on the significance of the evidence-based policy and practice movement or the What Works movement. I do not speak on how their rise affected the functioning of schools and human service organizations, funding structures, and governance. The merits of their rise are debated elsewhere by observers of particular domains (e.g., Mosley et al. 2019 for social work), methodologists (e.g., Heckman 2020), philosophers (e.g., Cartwright 2012), and by actors themselves (e.g., current debates in the Arnold foundation), who

increasingly reflect on the approaches they promote. Here, I instead try to contribute by providing an explanation of their rise, without debating their merits. Neither do I focus on a "deep history of neoliberalism." An attentive reader will notice that I resolutely stick to actions and knowings of particular groups. I do so on purpose, sharing science studies' conviction that the appearance of social facts is created from complex alliances (Latour 1987), with the hope that I produce a somewhat satisfactory answer to my narrow question. It would be possible to place my account of the rise of the evidence-based policy and practice movement within a grander narrative, but I do not do so here.

Structure & summary

I begin with the NCLB Act and its place in US educational policy. I notice that the political battle over the act was fought over the accountability provisions, with much less attention placed on the experimental mandate, which, nonetheless, is frequently mentioned in the act. Then, with the aid of a timeline, I pinpoint the moment when educational documents began using the phrase "scientifically based research." I discuss Robert Sweet, the person who came up with the phrase in 1997, and place the experimental mandate into the context of the long-standing Republican affinity for teaching reading via phonics. I provide additional background on the Reading Wars—the long-running American debate about teaching reading. We cannot understand reading legislation without the Reading Wars which shaped how politicians and researchers thought about reading in the 90s, created an enduring alliance between conservatives and phonics methods, and, through high-profile reforms in California and Texas, helped make education the foremost policy issue in the Bush-Gore election. Then I notice that there is continuity in conservative championing of phonics instruction, but that their method of arguing

has changed: e.g., in the 80s, as Reagan-era director of the National Institute of Education (NIE) Robert Sweet packed peer review panels to fund more phonics research, while in the 90s he turned into a champion of scientific objectivity. This is because the "evidence base" has changed. I show that NCLB legislation based its conception of "scientifically based research" on an influential report by the National Reading Panel (NRP), which was convened by National Institute of Child's Health and Development in 1997. The NRP is the latest in a series of reviews of reading research. With an analysis of the five most influential summaries of research on reading used by the federal government since the 60s I show that there has been a change in the late 90s towards the use of meta-analyses, or quantitative research syntheses. The NRP was influential, because it has been used to fill in the empty legislative mandate for "scientifically based research" and because the work on it has shaped the text of NCLB Act itself. I then show that the turn towards quantitative synthesis in the NRP occurred because a group of persons which included medical doctors, statisticians and social scientists were deliberately trying to emulate evidence based medicine (EBM) in social policy. They pursued a vision of an "evidence based society", an idea articulated by Adrian F.M. Smith in 1996. I then show that EBM was an attempt by scientists to have jurisdiction over the medical profession. EBM relied on research syntheses, which were new tools developed in the social sciences to deal with the post WWII growth in science. Evidence-based society took a long route to social policy through medicine, even though the methods of systematic review of research it drew on first arose in the social sciences, because medical researchers' methodology effectively took over the divided field of education research.

The No Child Left Behind Act

In the 90s both parties moved towards a consensus. With Clinton's Improving America's Schools Act (IASA) there is a push from the Democratic party away from increased spending towards increased standards. IASA and Goals 2000 don't succeed, however, because of weak enforcement and low political priority after Republicans take control of congress in 1995. In the late 90s there emerges a broad coalition supporting roughly similar policy goals of educational reform through increased accountability. This coalition has Democrats, Republicans, and the business community and civil rights activists but alienates teacher unions and some educational conservatives (Rhodes 2012). In the Bush v. Gore election education is the top issue and after he is elected, George W. Bush signed the No Child Left Behind Act (NCLB). NCLB strengthened the accountability provisions of IASA and linked some federal funding to the use of instructional materials backed by "scientifically based research," i.e., by quantitative and experimental or quasi-experimental research. The phrase scientifically based research appears in the text of the act over 100 times.

The experimental mandate of NCLB was not something that all coalition partners favored. Rhodes shows that the NCLB Act was supported by a "civil rights activists", that is, groups like the Citizen's Commission on Civil Rights and the Education Trust. Rhodes gives as an example a statement by the Citizen's Commission on Civil Rights, which said that "many of the proposals [in George W. Bush's NCLB plan] for standards, assessments, and accountability build upon or duplicate the 1994 reforms [that is, the IASA]. We note that these reforms were broadly supported by civil rights and advocacy organizations, parents and educators, the previous Administration, and, significantly, a bipartisan agreement of the Congress" (Citizen's Commission 2001, p.4). What is important for my purposes is that the emphasis on "scientifically based education" did not come from the civil rights group; research is seldom

discussed, and never systematically, by the Citizen's Commission in the "Analysis of President George W. Bush's Education Plan" (Citizen's Commission 2001) and in "Making Schools Work for Children in Poverty". Similarly, research syntheses were not used by the Education Trust, a think tank with an activist mission. Similarly in the business lobby discussed by Rhodes. The coalition around NCLB Act was gathered around standards and accountability, not around the experimental mandate; the accountability provisions of the NCLB Act largely follow IASA, the Bill Clinton ESEA reauthorization, but the requirements for the use of experimental evidence are new. How can we explain the introduction of an experimental mandate?

I have constructed a rough timeline of events and introduced bills. This timeline, which puts texts that use the term "scientifically based research" or "evidence-based research" in bold, shows the rapid ascension of "scientifically-based research" in congress around the year 1998.

[Timeline 1; all timelines, definitions, and figures are at the end of the paper]

The legal definition of the term has changed slightly (see Definitions 1-5). It first replaced previous formulations that were also meant to target experimental, quantitative research in the context of reading legislation sponsored by Republicans. It was then incorporated into the Democratic "Public Education Reinvestment, Reinvention, and Responsibility Act" of 1999 and eventually became the backbone of the NCLB Act of 2001, where it was used to tie federal funding to instructional programs that were "scientifically based, and the Educational Sciences Reform Act of 2002, where it was used to prescribe what type of research was to be funded by the Institute of Education Sciences (IES).

Conservative politics and scientifically based reading research

We know who is responsible for inventing the term "scientifically based reading research." It was coined by Robert Sweet, a professional staff member for the majority members of the House Education and Workforce Committee, who spent several months consulting with university-based researchers, mostly cognitive psychologists, to come up with a definition of educational research which was used in the Reading Excellence Act (personal communication of Robert Sweet and Lisa Towne, Eisenhart and Towne 2003). Sweet introduced a new term but not a new concept; earlier versions of the bill, see Timeline 1, used the term "reliable, replicable research" (see Definitions 1 and 2 at the end of the paper). Both acts explicitly mention phonics research as an example of what the term of art is supposed to cover, with phrases to the effect that "reliable, replicable research" should be conducted on phonics and other topics. The revision that added "scientifically based reading research" has also narrowed down the definition of reading, no longer counting "the ability to think critically about the meaning, message, and aesthetic value of the text" ("Reading Excellence Act (105 H.R. 2614 RFS Referred to Senate Committee)", 1997) as part of the definition of reading and removed a definition of "reading readiness" from the text of the bill altogether.

Robert Sweet has been involved in politics of education since the 70s, first as a teacher, then as an activist. He was active in the Reagan campaign and, after his election, started working as the deputy director of the National Institute of Education (NIE). After his superior at the NIE, Edward Curran, was asked to resign over a letter he wrote to the president asking him to abolish the agency (Zodhiates 1988, p.106). Michigan historian of education Maris Vinovskis testified in in a statement to the senate budget committee in 1997 that Sweet and his predecessor at NIE "were accused of politicizing NIE by changing research topics to reflect a right-wing agenda,

ignoring the peer review system, firing NIE professionals, hiring unqualified employees (some of whom had been active in Republican politics) and ignoring congressional mandates." (Congressional record 1997, p. 202). Under Sweet's tenure at the NIE, 159 employees left the institute and 26 were hired to replace them. Of those hired, most were Reagan's loyalists active on the far-right wing of the republican party without appropriate qualifications or research experience (Department of Education Weekly 1982). Sweet also disregarded the peer review system for the Unsolicited Proposals grant competition at NIE; reviewers were chosen based on their party or religious affiliation, not based on their demonstrated research ability, and Sweet, then the acting director of the NIE, retained the final choice over the proposals chosen (this has been paralleled in the rollout of the Reading First portion of the NCLB, when the Bush Department of Education "failed to follow its own guidance for conducting the peer review process [Office of Inspector General 2006]). Sweet's choices fully disregarded recommendations by NIE staff. Some topics were eliminated altogether, and these included studies of social development and behavior and proposals on desegregation (Sweet's list of approved proposals from August 1982 as cited in Zodhiates 1988, p. 134). What is interesting for the purposes of this paper, out of the eight proposals that won, five mentioned reading or writing, which represented a marked increase from the previous focus on equality towards a new focus on so-called "fundamental skills". That the 1998 Republican legislative proposal for the "Reading Excellence Act" singled out phonics as an exemplar of "scientifically based reading research" shows the durability of the conservative affinity with phonics and a rhetoric of teaching skills. In 1983 Robert Sweet explained that "public schools have become the laboratory for behavioral scientists. "We have spent too much time analyzing the psyche of kids and too little time in trying to develop the intellect" (Sweet cited in Zodhiates 1988, p.108) and that "parents want to

get back to fundamental skills, not the social change of the '60s" (Sweet cited in Zodhiates 1988, p.108). At the same time, Sweet had little belief in social sciences' ability to bolster his position. He said, "from my experience, you can't have research in social science that isn't political. The federal government shouldn't get involved in educational research. You shouldn't give discretionary money at the federal level because it will be abused. My fundamental premise is that social science isn't research; it doesn't deal with hard data (…) [The problem with government sponsorship of the social sciences is that] you give money to find the answers you want" (Sweet cited in Zodhiates 1988, p.111-2). Above I illustrated that Sweet favored one side of the reading wars, the phonics side, but that he did not speak of conclusive evidence that his side was the more scientific. In fact, he faced active resistance from the researchers at the NIE. Twenty years later Sweet, who under Reagan meddled with the mechanisms of science, would be the foremost champion of deferring to the scientific consensus.

We cannot understand reading legislation without looking at the Reading Wars

We have seen above that Republicans played a key role in introducing "scientifically based" or "evidence based" practices in education. They introduced "scientifically based reading research" into the legislation for the first time and then it was their administration that signed the NCLB and the Education Sciences Reform Acts. Moreover, the republican-controlled House and Senate asked for syntheses o reading research which bore "evidence-based" in their title (see timeline 1). Why did Republicans champion evidence-based practices? Why did they emerge in reading and not in some other policy domain?

The Reading Wars are a long-running debate about the proper way to teach children how to read. It has been mostly confined to the US, although it has had some resonance around the

globe. The Wars are closely tied to the Science of Reading, which originated in the late 19th century and rose to prominence in the interwar period in the US. The precise beginning of the Wars is difficult to specify for there always has been some degree of debate about early childhood education. A book that is often cited as an instigator is a 1955 polemic by Rudolf Flesch who argued that poor reading abilities of American children came from the lack of phonics instruction in primary schools (what is more important here than the history is the fact that already in the 50s the debate was going on and the positions were fleshed out). In *Why Johnny Can't Read* he virulently attacked a method of teaching reading that taught children to first learn how to quickly recognize whole words. This method, known as "look-say", was one of a string of instructional approaches that focused on the meaning of the text. Instead, Flesch promoted phonics instruction which begun by teaching children the relationship between letters and sounds and then instructed them how to blend those sounds into unknown words. The poles of the ensuing debate have been summed up thusly, "Do children learn better with a beginning method that stresses meaning or with one that stresses learning the code?" (Chall 1967, p.75). In the following decades the Wars were fought between proponents of whole language and phonics-based methods of literacy instruction.

The Reading Wars created polarized positions, with conservatives championing phonics for a variety of reasons. The affinity of phonics instruction and conservative politics has been explained by cold war pressure for fundamental skills education, desire to undermine teacher unions and teacher autonomy, dislike for progressive educational ideals such as holistic and child-centered education, dislike for the books assigned in meaning-oriented teaching process, an ideological bundling of meaning approaches with other progressive approaches, and a "jurisdictional battle" over the teaching process by progressive teacher unions and conservative

contenders. In the civil rights era, the battle was often about equality against “fundamental skills”, with “fundamental skills” being championed by politicians wanting a neutral language in which to oppose civil rights. Overall, there has been stability in conservatives’ championing of phonics, although at times it has been used to argue for decreased spending (Reagan) and at others it was a vehicle for federal expansion (George W. Bush). In the literacy debates the consequences of settling the scientific issues were immediate—how should children be taught. And because the debate concerned the young, every parent was involved, and the stakes were high. The debate was not confined to academia and quickly reached national politics, where stable polarized positions were created, with progressives more favorable towards the meaning-based approaches, and conservatives favoring the skills-based approaches.

Reading Wars bring reading instruction to the forefront of national debate in the 90s, via California and Texas

In late 80s there California reorients its literacy curriculum around meaning-based approaches. According to Timothy Shanahan, a member of the National Reading Panel, “[w]hole language influenced policies translated into a ban on the use of state money to purchase spelling books (whole language proponents opposed spellers)” (Shanahan 2006, p. 78). Instead, the California language arts framework argued that children construct knowledge on their own. The enthusiasm was short lived. In 1995 the results of the 1992 first wave of the state-level National Assessment of Educational Progress (NAEP) were released; the assessment survey placed California near the bottom of the country, with declines in reading ability documented across socioeconomic groups. In wake of the California fiasco, there is a wave of phonics reform

around the country. Between 1994 and 1997, phonics bills were introduced in 18 states, with California leading the way (Paterson 2000).

In 1994 George W. Bush was elected as governor of Texas. With backing of the business community (Sandy Kress of the Texas Business and Education Coalition shaped Bush's Texas educational strategy [Rhodes 2012, p.138]; Pyle 2005) he pursued an accountability-based education reform, which then serves as a plank of his presidential campaign. It is also then that Bush forms ties with Reid Lyon of the NICHD who at the time led the NICHD studies of phonics, which later spurred the National Reading Panel.

Moreover, an explanation of the NCLB Act needs to note the increasing problematization of education in American politics. This was in part due to the debates surrounding the Reading Wars, but also to sense of threat of foreign competition and increased availability of surveys showing poor achievements of American children. The 1985 *Nation at Risk* report was instrumental in bringing education, and especially education of disadvantaged students, to national attention.

Change in the available "best evidence" on reading

Non-experts, including politicians, encounter science in summaries of research, which serve the crucial role of summarizing big science to its practitioners and producing a unified "scientific consensus" to external audiences; reviews also allow science to lend credence to one voice which weighs in on matters of concern. There are different types of documents that can be used to summarize the current state of a science. Those documents have been used both internally, to represent science to itself, and externally, to represent science to the rest of the world. The former concern the fact that science is itself a social process, with actors in time

(which entails processes of becoming and socialization) and in interaction (which entails the achievement of cooperation by actors in different positions and different tacit and discursive knowledge). The latter concerns the fact that science is an adversarial but presents itself as cumulative and consensual. The review of research can take different forms, depending on the method used to arrive at the conclusions. Most common forms are either deference to expert judgment, where the authority of the review comes from the authority of its author or authors, be it individual narrative synthesis or a consensus panel of experts, or a systematic review of published research based on explicit selection criteria, e.g., only including peer-reviewed research with experimental study designs, which are then summarized either narratively, through vote counting, or with a quantitative synthesis. A quantitative synthesis of the findings, also known as a meta-analysis, is a type of systematic review which combines the results from multiple studies to arrive at novel statistical results. Reviews of research can be published as stand-alone articles, commissioned reports, as literature review section of research papers, or, and this is the genre that has received most attention from science studies, in textbooks.

The importance of the one such synthesis, the National Reading Panel's *Teaching Children to Read* (NRP), should not be understated. Its subtitle is *An Evidence-Based Assessment of the Scientific Literature on Reading and its Implication for Reading Instruction* and it is the first report on education to bear "evidence-based" in its title. The report has been widely promoted during the NCLB rollout and its findings have been used to substantively fill in what research is and what is not "scientifically based"; most notably, the Department of Education based its interpretation of the "scientifically based research" requirements of the Reading First program on the findings of the NRP (Office of Inspector General 2006). Moreover, the report directly comes out of the congressional debates about literacy that went on at the same time that

the term “scientifically based reading research” was introduced into legislation; we have reasons to suspect that the report influenced the legislators a long time before it was published. The puzzle about NRP is that there have been many reports before it, but they did not have the same kind of direct influence on policy. As a quantitative synthesis, the NRP is also different in methodology from the reports that preceded it. Why did this influential report break with the tradition of presenting research that came before it, and why was it so influential?

Summaries of reading research have changed in the period under study with federal research on early reading coming to rely more on quantitative syntheses and experimental evidence. Some see this change as sufficient to explain the experimental mandate: as the research got better its voice grew stronger in the congressional debates. Below I discuss the research syntheses available to policymakers. I have chosen the texts in two ways. First, they are widely cited in the literature as the defining reviews of research on reading produced at their time. Second, they cite one another; each report claims to be building on the findings from the previous ones. The texts are: *Learning to Read: The Great Debate* (Chall 1967), *Becoming a Nation of Readers* (Anderson 1985), *Beginning to Read: Thinking and Learning About Print* (Adams 1990), *Preventing Reading Difficulties in Young Children* (National Research Council 1998), and *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Literature on Reading and its Implication for Reading Instruction* (National Reading Panel 2001).

Major reviews, often commissioned by the federal government, combined diverse strands of evidence. Always it was implied that the evidence used was the best available evidence. In practice, however, the texts combined diverse sources of knowledge and proceeded via narrative and argument, not quantitative synthesis. Often the texts incorporated a systematic review in a

narrative synthesis. For example, Chall (1967) used a "vote counting" method to count how many studies supported code-based and how many meaning-based reading instruction. Adams (1990) influential *Beginning to read* demonstrates phonics to be a valid approach draws on a quantitative research synthesis (p.46) by Pflaum, Walberg, Karegianes, and Rasher (1980) and two large experimental program evaluations in the US (1964-1967, Bond and Dykstra 1967) and 1970s Follow Through. The text does not undertake a synthesis of research but discusses an external quantitative research synthesis. The text presents the quantitative research synthesis as a novelty and combines it with other methods, highlighting above the inability of quantitative synthesis to discriminate between different instructional practices: "within each [evaluated method of instruction], there tend to be at least some aspects that are uniquely beneficial as compared to the standard methods. The challenge left to us is that of identifying the pearls" (Adams 1990, p. 47).

With the National Research Council report, we see a major rhetorical shift with the introduction of "criteria of selection" sections in papers which often emphasize experimental and quasi-experimental designs in an explicit statement of the hierarchy of evidence used. The 1998 report from the National Research Council, titled *Preventing Reading Difficulties in Young Children* (National Research Council 1998), has a section subtitled "selection criteria" where it is explained that the report took into account "converging evidence from experimental investigations, correlational studies, nonequivalent control-group studies, and various other quasi-experimental designs and multivariate correlational designs" (National Research Council 1998, p.178). This is a major rhetorical shift; previous reports did not have "criteria of selection" sections. Those sections often emphasize experimental and quasi-experimental designs in an explicit statement of the hierarchy of evidence used. The NRC report notes the tradeoff between

external and internal validity, with correlational studies being more generalizable but less robust. The report is a consensus report, which means that methods of selection were chosen to reflect the composition of the panel, without additional authority being delegated to certain methods.

The 2001 National Reading Panel is a major departure from previous reports. It is a systematic review only, with stringent selection criteria including only experimental and quasi-experimental studies of highest quality and reanalyzes all previous findings. The first chapter is on methodology and is based on Harris Cooper's and Larry Hedges' textbook on quantitative research syntheses (Cooper and Hedges 1994) and Harris Cooper has been consulted by the panel and is thanked in the acknowledgments.

The importance of medical doctors and the analogy to medicine in the NRP

With the NRP, the NICHD created a federally sponsored report on reading that was be a quantitative synthesis and not a consensus report. One of the panel members said that the decision to include only experimental and quasi-experimental studies was "urged by the NICHD at the first panel meeting" and later accepted by the panel overwhelmingly composed of university researchers (Yatvin 2002). The drive towards meta-analysis came from the NICHD, not from Congress which requested the report, and this is the probable reason why the report took three years instead of the one year initially allotted by congress. The emphasis on experimental method is even more visible in the summary of the report, which even the leading member of the panel admits was misleading about the claims about the comprehensiveness of the report (Shanahan 2003).

The panel moved in the direction of quantitative synthesis because it was orchestrated by medical doctors and dominated by university reading researchers. When we look at the selection process to the NRP, we see a significant departure from what came before. Most significantly,

there are direct connections to medicine. The person who led the NICHD, the agency that commissioned the NRP, was dr. Duane Alexander, a medical doctor. He was responsible for the selection of the chair of the panel, and he chose a person with little connection to education research, Dr. Langenberg, a physicist. Together they chose the remaining panelists: 12 university professors (8 reading researchers, 2 administrators, one medical doctor, and one teacher educator), one parent, one teacher, and one school principal. The panel was orchestrated by medical doctors and dominated by university reading researchers. It is interesting to point out that the same time a similar panel was convened in the UK to guide its literacy programs. The UK literacy task force consisted of an equal mix of teachers and university researchers, with no members that were leaders of the field. The UK task force produced a report that was similar to those produced in the US before the NRP; it was a review of research that drew on meta-analyses but meta-analyses did not dominate (Beard 1999).

At the same time the persons leading the NICHD framed reading as a public health problem. “I think that it is important to point out that our intensive research efforts in reading development and disorders is motivated to a great extent by our seeing difficulties learning to read as not only an educational problem, but also a major public health issue,” (Congress Appropriations Committee for Fiscal Year 1998) said dr. Duane Alexander, director of NICHD, while presenting the NICHD research that led to the NRP. His collaborator and Bush’s “reading czar”, dr. Reid Lyon, also made the analogy to medicine when he argued that not using data to drive teaching practice “is tantamount to malpractice” (Why Children Can’t Read 1997, p.9).

The term “scientifically-based research” appears in the congressional record once before it is introduced in the REA by Robert Sweet, and it is in connection to the National Reading Panel. It was used a single time by prof. Laura Fredrick in a hearing on “Why Children Can’t

Read" on September 3rd, 1997. It should come as no surprise that prof. Fredric used the term "scientifically-based research" in a statement that gave the recommendation to "establish an "Expert Reading Panel" that would synthesize research-based knowledge about reading and advise the Administration on a continuing research agenda in reading. I understand from the National Center to Improve the Tools of Educators that there is already a possibility of establishing such a panel" (Why Children Can't Read 1997, Fredrick p. 398). That same hearing also featured dr. Reid Lyon, dr. Duane Alexander's collaborator at the NICHD and later Bush's "reading czar" who promoted the idea of a "teacher that uses research-based information about reading development and about reading difficulties." Reid also called for a reading panel, "we need to develop a formal procedure to assess the current status of scientific research-based knowledge relevant to reading development, reading disorders, and the effectiveness of various approaches to teaching children to read" (Why Children Can't Read, p.118). Among the scientific witnesses summoned to congress there appears a conviction that science has arrived at a consensus about what works in early reading instruction but that the knowledge was not distributed to the teachers yet. If this is true, it would mean that the research synthesis undertaken by the NRP was not a method of producing new knowledge but of buttressing it and disseminating it widely. This focus on synthesis and dissemination is shared with EBM.

To understand the NRP, we need to understand how meta-analysis became important as a policy tool.

Calls for an evidence-based society

In the 90s many quantitative researchers were in thrall of the idea of an "evidence-based society."

The idea of an “evidence-based society” was articulated by prof. Adrian F.M. Smith in the 1996 presidential address to the British Royal Statistical Society. An evidence-based society would be such “in which decisions about matters of substance with significant potential social or personal implications are taken on the basis of the best available evidence, rather than on the basis of irrelevant evidence or no evidence at all.” Best available evidence means quantitative evidence, for an evidence-based society is such “in which informed quantitative reasoning is the dominant modality in public debate, as well as in the decision-making processes of government, business and individuals”, although “the nature of what constitutes evidence in any particular instance could be a matter for significant debate.” In the address, Smith went on to argue that those principles were already followed by the evidence-based medicine movement and that similar efforts could and should follow across the whole spectrum of public policy issues, and especially in education, because “what is so special about medicine?” Such efforts would require the creation of an “evidence base” in other areas then medicine, which would serve the dual purpose of creating databases for scientists and “communicating findings in a readily assimilable form to relevant practitioners.” Smith then called for “developing analogues of the Cochrane Collaboration, to provide suitable evidence bases in other areas besides medicine, with the aim of achieving a quantal shift in the quantitative maturity of public policy debates” (Smith 1996). Soon after he creators of the Cochrane Collaboration launched an and analogous effort to extend “evidence-based society” in areas of education, social work/services, crime and justice, and dissemination/technology; this was known as the Campbell Collaboration.

It is quite important to note the difference between medicine and teaching. Whereas in medicine, evidence-based guidelines synthesizing best practice proposed to supersede the expertise of each doctor, which was developed through engagement with the natural sciences, in

education there was much less of an agreement about what kind of expertise a teacher has and much less connection between teaching and theory. Evidence based medicine saw itself as being the next step after scientific medicine; policy based on scientifically based research saw itself as introducing scientific method to the classroom for the first time. In another influential address from 1996, which was titled "Teaching as a Research-Based Profession" and given to the British Teacher Training Agency, David Hargreaves argued that teaching should model itself as a profession on scientific medicine as it was institutionalized in the middle of the 20th century (Hargreaves 1996).

Rise of evidence based medicine in the 90s

Before I discuss the Campbell Collaboration and its ties to the NRP and NCLB, I want to point out a few facts about the development of evidence based medicine. For context, I begin with word frequency time-series plots produced by means of the Google Ngram Viewer. Those plots reliably tell us when a term appears in printed word; inferences about the significance and popularity of any given term, or any comparisons between terms, are not warranted because (a) the corpus of books need not be reflective of the population of ideas and (b) the corpus of books does not reflect the popularity or influence of each book. Figure 1 shows the frequency of the search term "evidence based *." The asterisk indicates a wildcard, which means that the search included the 10 most popular three-word phrases beginning with "evidence based". We see that "evidence based medicine" emerged earlier than all other terms, and that "evidence based policy" is the only phrase in the top 10 that is not related to medicine. Figure 2 shows that the phrases related to policy and reading research gained popularity after the year 1995, a half decade after "evidence based medicine". Figure 3 shows the invention of the term "meta-analysis" in the late 70s and its steady rise along with "randomized controlled trial", "RCT", and

"evidence based *". Here the wildcard is collapsed, which means that the frequencies of all "evidence based" phrases from Figure 1 were added.

[Figure 1]

[Figure 2]

[Figure 3]

Similar analysis using the Web of Science corpus yields similar results, with an early rise of meta-analyses, followed by evidence based medicine and evidence based practice, followed by evidence based policy.

Growth of quantity of experimental studies and invention of quantitative research synthesis as a prerequisite of EBM

Evidence based movement in medicine has required the development of a new knowledge tool, the research synthesis, which allowed for authoritative summaries of current best science. From the science studies perspective, research syntheses are a rhetorical form that allow for the settlement of controversies and creation of uncontested/uncontestable scientific "facts." It is important to note that research syntheses became a necessary genre only at a specific point in the growth of science, when the amount of research produced became difficult to take in by a single person (before that point textbooks were research syntheses, but they served a pedagogic, not summative, function). Second important development is when research syntheses became recognized as legitimate scientific output, capable of conferring prestige on its authors; this is closely connected with the rise of "quantitative research synthesis" or "meta-analysis". There are multiple related terms here that differ slightly in terms of content and the

period when they were used. The more general term is "systematic review", which refers to the collection and synthesis of all relevant studies on a specific topic, with relevance defined in a systematic manner. Other terms that have been used for a similar meaning are "research synthesis" (Mandel, 1936) and "critical review". Quantitative research synthesis and the more specific term "meta-analysis" has been introduced by Glass in a presidential address to and means the "statistical synthesis of the data from separate but similar, i.e., comparable studies, leading to a quantitative summary of the pooled results" (Last 2001, 114). When did research syntheses begin? Founder of the Cochrane Collaboration, Ian Chalmers (2002), notes several early calls for meta-analysis, including Gauss' and Laplace's methods in astronomy (Airy, 1861), Pearson's in vaccine evaluation (1904), and Fisher's in agriculture (1932). Chalmers notes that the methods were not widely used until in the 60s the post-war rapid growth of social science research created the need to combine results from hundreds of studies. Any full explanation of the rise of an “evidence based society” would have to explain the prior growth of available quantitative evidence, which I do not tackle in this essay. As research syntheses gained popularity, they also gained in prestige and begun to be treated as legitimate scientific output. This culminated in a set of guidelines for their production crystalizing in the late 70s and early 80s; this is also the moment when various "hierarchies of evidence" with randomized controlled trials on top becomes a stable part of the guidelines for writing syntheses.

The methods underlying evidence-based medicine first arose in experimental psychology and education. Meta-analysis was introduced in a presidential address to the American Educational Research Association (Glass 1976, Chalmers 2002, p. 24) also hints that there was a two-way connection between the social sciences and medicine. For example, the influential medical "Guidelines for reading literature reviews" (Guyatt 1988), which are often cited as the

originary article for evidence-based medicine, mention prior practices in the social sciences as inspiration. In the 90s, calls for dissemination of research syntheses in medicine intensified in high-profile medical publications. In 1993 the Cochrane Collaboration is founded by Ian Chalmers to conduct and disseminate syntheses. Changing communication technology seems to have played an important enabling function because maintaining an up-to-date database of research syntheses, rather than a periodically updated handbook, became viable. Evidence based medicine was hugely successful and soon after explicit calls for its imitation appeared in education.

Transplanting EBM into social policy

Campbell Collaboration is created by the same people who created the Collaboration. The main person responsible is dr. Ian Chalmers who got funding from Michael Peckham. As the director of the Research and Development Programme at the U.K. National Health Service Peckham had previously funded the early efforts of the Cochrane Collaboration. In 1998 he was the Director of the School of Public Policy at the University College London (UCL) and provided Chalmers with some funding for an analogue in other areas. Chalmers met with Fred Mosteller. Chalmers decided that the Collaboration must be centered in the US, where majority of relevant research is produced; he meets Robert Boruch at a Cochrane gathering in Baltimore and Boruch is decided to be the leader of the American side of the movement.

A report was issued by the President's Council of Advisors on Science and Technology (PCAST 1997), where the education panel consisted of industrialists, astronauts, and scientists, without much involvement from education researchers. National Academy of Education issued its own report titled “Recommendations Regarding Research Priorities: An Advisory Report” (Brown et al. 1999) was chaired by university-based education researchers. The National

Research Council issued a report titled “Improving Student Learning: A Strategic Plan for Education Research and Its Utilization” (NRC 1999); the NRC report “was undertaken largely because of the concern of Bruce Alberts, President of the National Academy of Sciences (NAS), that research has not had the kind of impact on education that is visible in medical practice, space exploration, and many other fields” (AERA 1999) and it called for greater use of research syntheses.

Also in 1999, at a Brookings Press Forum at Stanford called “Can We Make Education Policy on the Basis of Evidence?“ Paul Peterson, the director of the program on educational policy and governance at the John F. Kennedy School of Government at the Harvard University, explained that he and Howard Hyatt, co-organizer and professor of medicine at Harvard, “we felt that the kind of research that is characteristic of medical research, the randomized field trial, the use of the pill and placebo (…) that that is the best way of conducting a scientific experience to find out whether something works or not” (Brookings Press Forum 1999, http://stanford.edu/~hakuta/www/docs/brookings.html). Also present at the conference were university researchers, medical doctors, statisticians, and Peter Boruch, the director of the conceived Campbell Collaboration. What’s interesting, also at the conference Paul Peterson also called for educational research to become more like management consulting, where freshly minted MBAs disseminate best practices learnt in business schools, which are often based on experimental or quasi-experimental studies.

All this goes to show that in the late 90s there were many calls to make education more like medicine. These efforts were at first directed towards the dissemination of research, but later towards the production of a specific kind of quantitative research that could then be turned into best practices. At the same time the framing of education research as an applied, practical

science gained in popularity: the ultimate justification of research was that it would be useful once disseminated.

Evidence-based society took a long route to social policy through medicine, even though the methods of systematic review of research first arose in the social sciences.

Summary

Why then did the NCLB and ESRA feature the experimental mandate?

a. The Reading Wars consolidated Republican position on phonics instruction, so the Republican-controlled congress commissioned the NRP, which was known to favor results favorable to phonics.
b. The NRP went even further than the congressional mandate, because of the idea of an "evidence based society." Post WWII growth of science and of experimental studies conducted leads to the development of a new tool for synthesizing knowledge, the quantitative research synthesis. The quantitative research synthesis leads to evidence-based medicine, which enjoys spectacular success. Field of education remains low-status and faces persistent legitimacy crises, which lead to calls for the emulation of medicine in the social sciences, which leads to deliberate attempts to create evidence-based policy by leading figures of the EBM movement, statisticians, and social scientists. NICHD, an agency directed by doctors, responds to a congressional mandate to create an expert panel with a panel that exceeds the congressional mandate and is supposed to produce a quantitative research synthesis.
c. The Reading Wars in general, and the developments in California and Texas in particular, create a rising awareness of literacy as a problem, which leads to high-profile debates and a need for policy position of the Bush administration. There emerged a broad-based

coalition for standards-based reforms, which was not against the experimental mandate. Bush, who led the coalition, strongly supported the experimental mandate because of his personal ties to dr. Reid Lyon, an important figure in the NICHD and his "reading czar" from his time as governor of Texas. The NRP becomes the backbone of NCLB.

Conclusion: studying policy science, studying policy knowledge with tools from science studies

Reading Wars and other curricular debates could supplement science studies' current canonical examples: climate change denial, smoking, and vaccination skepticism. Those cases are important to study, but because of their historical significance and specific characteristics, they might not be well suited to answer general questions about politics of knowledge production and the use of knowledge for governance. Climate change denial, smoking, and vaccination skepticism are like the French Revolution to the studies of revolutions: important as a historical phenomenon, but ill-suited to be a stand-in for all other revolutions. Monika Krause (2021) argues that social scientists always study a concrete case, the French Revolution, while pursuing a more general epistemic target, all revolutions. Politics of knowledge about reading is relevant to multiple epistemic targets: how is this knowledge used to govern? What is the relationship between citizens and the knowledge used to govern them? How do decision makers create and use knowledge to govern? How is knowledge produced and disseminated in knowledge societies?

Studying policymaking can benefit from science studies because science—applied social science and the science of reading—plays a big role in how policymakers think and therefore understanding science qua science is important for to understand the policy process. Moreover, science studies are a suite of strategies for studying knowledge production, regardless of whether it is scientific knowledge, lay knowledge, or knowledge of an uncertain status, such as policy oriented social scientific knowledge. This emphasis on studying knowledge was not the case

with early sociology of science, which deliberately steered clear of questions about knowledge (Merton 1973). The question of knowledge was then left to the philosophers, who studied science as an exemplary case in epistemology (e.g., Popper 2005). Later, sociologists started studying the epistemic practices of sciences empirically, bringing ethnographic and ethnomethodological methods to bear on the study of the sites of knowledge production (Knorr-Cetina 1999, Latour 1987). Those methods can be extended to the policy arena.

The policymaker faces moral and epistemic problems and how he or she resolves both needs to be part of a full explanation of a policy decision. Moral problems are those that relate to ends and answer the question of what goals are worthy of pursuit and how they should be pursued. This type of problem is the only type of problem encountered by some posited decision makers, including those we encounter in some moral philosophy and some experimental research, which aims at person's moral faculties through highly abstracted situations that require purely moral judgments, e.g., is it permissible to sacrifice one person to save five in a trolley problem. The policymaker is not in a thought experiment and the decisions he must face are not like those in the trolley problem because the rational policymaker finds himself in a complex, open world where neither it is clear what courses of action are possible, nor what will be the consequences of each. This creates an epistemic, i.e., relating to knowledge, problem for the policymaker. To explain a policy decision, we need to consider the epistemic methods (ethno-methods) used by the policymakers to gain knowledge of, and imagine action in, their world.

One such method is the quantitative research synthesis. It should not be seen, however, as a method that is controlled by the politicians. Rather, meta-analysis is a 'policy device', to take Hirschman and Bearman's phrase, which are "the sociotechnical tools that allow policymakers to see the world in certain ways (like GDP, or the unemployment rate) or assist them in making

decisions (like cost–benefit analysis)" (Hirschman and Bearman 2013, p. 790). This has its own context of production in the scientific arena; it is a tool which helps scientists make their sciences cumulative and is promoted by quantitative scientists who use it to gain expert jurisdiction over social problems. Meta-analyses are successful and on the rise because they are "boundary objects" (Star and Griesemar 1989): useful to politicians but produced and maintained by scientists (though often with money that must be secured politically).

Timeline 1: Event history of NCLB (texts using the term "scientifically based research" are in bold)

1994:

Improving America's Schools Act (Democratic) reauthorization of ESEA
Goals 2000 Act (Democratic)

1996

March 26-27, Education Summit at IBM campus
"Reading Excellence Act" (Republican) introduced in House; Definition 1

1997

Senate asks NICHD to convene a panel on reading after NICHD presents its own findings
"Literacy: Why Children Can't Read?" hearing introduces the phrase "scientifically based research" to the congressional record

1998

"Education at a Crossroads" report
"Dollars to the Classroom" Bill (Republican)
"Straight A's Bill" (Republican)
The phrase "scientifically based reading research" is introduced to the **"Reading Excellence Act"** (Republican) by Robert Sweet; Definition 2

1999

"Can We Make Education Policy on the Basis of Evidence?" Brookings Press Forum conference at Stanford
"Public Education Reinvestment, Reinvention, and Responsibility Act" (Democratic); this bill uses Definition 2

2000

Castle Bill (later Education Sciences Reform Act) introduced in Congress; Definition 3
"Literacy Involves Families Together Act" (Republican)
"Education Opportunities to Protect and Invest in our Nation's Students Act" (R)

2001

National Reading Panel releases **"Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature"**
"No Child Left Behind Act" (Republican); Definition 4
"Use of Scientifically-Based Research Working Group Conference"
CEBP: Coalition for Evidence Based Policy launched
J-PAL launched
What Works Clearinghouse launched

2002

"Educational Sciences Reform Act" (Republican) creates the Institute of Education Sciences (IES); Definition 5

Definition 1: "Reliable, replicable research" in the Reading Excellence Act before Jan 1st, 1998

"(6) RELIABLE, REPLICABLE RESEARCH.—The term 'reliable, replicable research' means objective, valid, scientific studies that—

"(A) include rigorously defined samples of subjects that are sufficiently large and representative to support the general conclusions drawn;

"(B) rely on measurements that meet established standards of reliability and validity;

"(C) test competing theories, where multiple theories exist;

"(D) are subjected to peer review before their results are published; and

"(E) discover effective strategies for improving reading skills.

Definition 2: „Scientifically based reading research" in the Reading Excellence Act after January 1st 1998

"(5) SCIENTIFICALLY BASED READING RESEARCH.—The term 'scientifically based reading research'—

"(A) means the application of rigorous, systematic, and objective procedures to obtain valid knowledge relevant to reading development, reading instruction, and reading difficulties; and

"(B) shall include research that—

"(i) employs systematic, empirical methods that draw on observation or experiment;

"(ii) involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;

HR 2614 EAS

6

"(iii) relies on measurements or observational methods that provide valid data across evaluators and observers and across multiple measurements and observations; and

"(iv) has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review.

Definition 3: “Scientifically based research standards” in the “original Castle Bill” (H.R. 4875, 2000, p. 3), definition taken from (Towne 2003).

(5) SCIENTIFICALLY BASED QUANTITATIVE RESEARCH STANDARDS.—The term “scientifically based quantitative research standards”—

(A) means the application of rigorous, systemic, and objective procedures to obtain valid knowledge relevant to education activities and programs; and

(B) includes research that—

(i) employs systematic, empirical methods that draw on observation or experiment;

(ii) involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;

(iii) relies on measurements or observational methods that provide valid data across evaluators and observers and across multiple measurements and observations and across studies by the same or different investigators;

(iv) is evaluated using experimental designs in which individuals, entities, programs, or activities are assigned to different conditions with appropriate controls to evaluate the effects of the condition of interest through random assignment experiments, or other designs to the extent such designs contain within-condition or across-condition controls; and

(v) ensures experimental studies are presented in sufficient detail and clarity to allow for replication, or at a minimum offer the opportunity to build systematically on its findings.

(6) SCIENTIFICALLY BASED QUALITATIVE RESEARCH STANDARDS.—The term “scientifically based qualitative research standards”—

(A) means the systematic collection and analysis of data often associated with traditions of inquiry historically based in the humanities, such as narrative analysis; and

(B) includes research that—

(i) uses some combination of participant observation, in-depth interviewing and document collection;

(ii) is intended to explore issues and hypotheses whose underlying dynamics and factors are not sufficiently well refined, understood, or amenable to experimental control to permit adequate study through quantitative research;

(iii) may include case studies, ethnographies, life histories, multi-site case studies, and participatory action research; and

(iv) uses approaches to assess the experimental knowledge acquired to assure that the findings are scientifically valid and replicable.

Definition 4: „Scientifically based research in the "No Child Left Behind Act" of 2000

The term "scientifically based research":

(A) Means research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs; and

(B) Includes research that:

(i) Employs systematic, empirical methods that draw on observation or experiment;

(ii) Involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;

(iii) Relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators;

(iv) Is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition;

(v) Ensures that experimental studies are presented in sufficient detail and clarity to allow for replication, or, at a minimum, offer the opportunity to build systematically on their findings; and

(vi) Has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review.

Definition 5: “Scientifically based research standards” in the “Educational Sciences Reform Act” of 2002

SCIENTIFICALLY BASED RESEARCH STANDARDS—

(A) The term “scientifically based research standards” means research standards that—

(i) apply rigorous, systematic, and objective methodology to obtain reliable and valid knowledge relevant to education activities and programs; and

(ii) present findings and make claims that are appropriate to and supported by the methods that have been employed.

(B) The term includes, appropriate to the research being conducted—

(i) employing systematic, empirical methods that draw on observation or experiment;

(ii) involving data analyses that are adequate to support the general findings;

(iii) relying on measurements or observational methods that provide reliable data;

(iv) making claims of causal relationships only in random assignment experiments or other designs (to the extent such designs substantially eliminate plausible competing explanations for the obtained results);

(v) ensuring that studies and methods are presented in sufficient detail and clarity to allow for replication or, at a minimum, to offer the opportunity to build systematically on the findings of the research;

(vi) obtaining acceptance by a peer-reviewed journal or approval by a panel of independent experts through a comparably rigorous, objective, and scientific review; and

(vii) using research designs and methods appropriate to the research question posed.

Figure 1: Ngram of “evidence-based *”

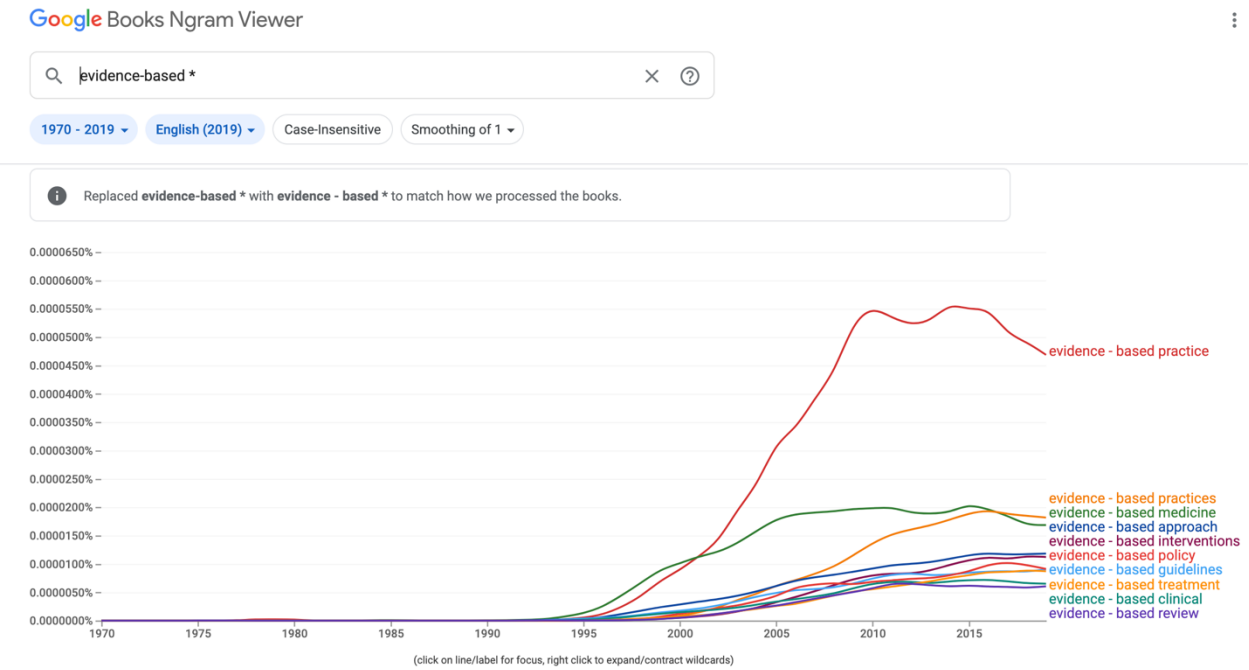


Figure 2: Ngram of „evidence-based medicine, evidence-based policy, scientifically based research, scientifically based reading”

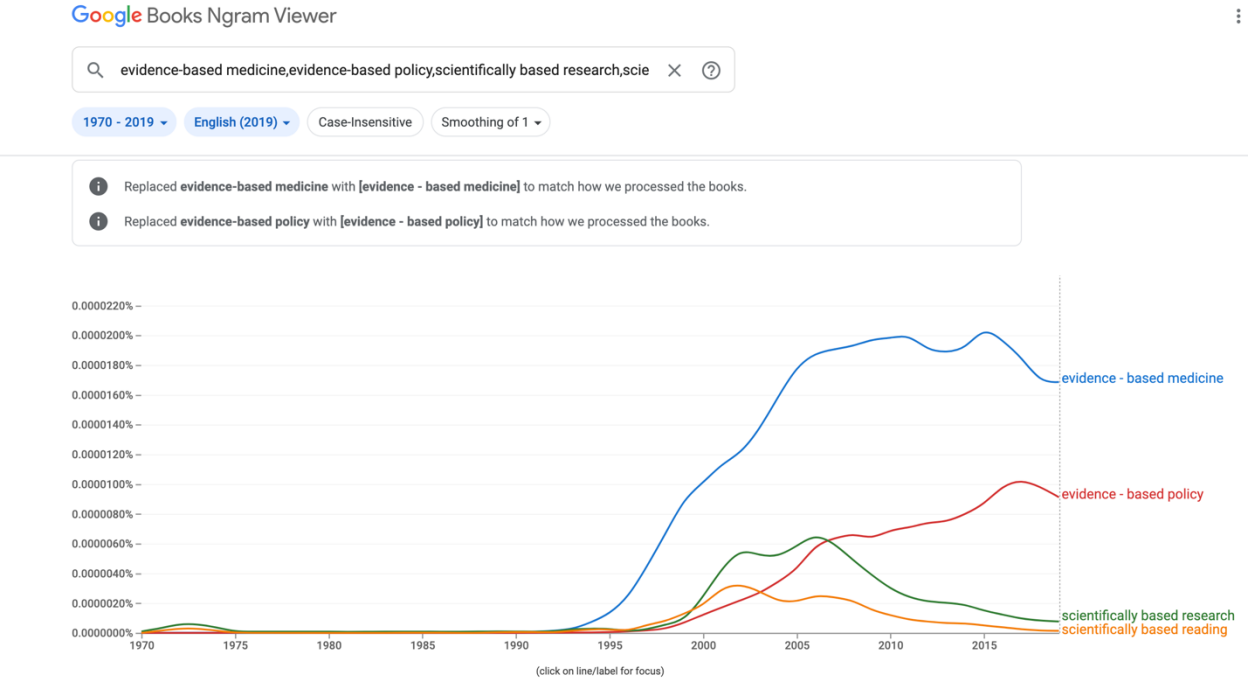
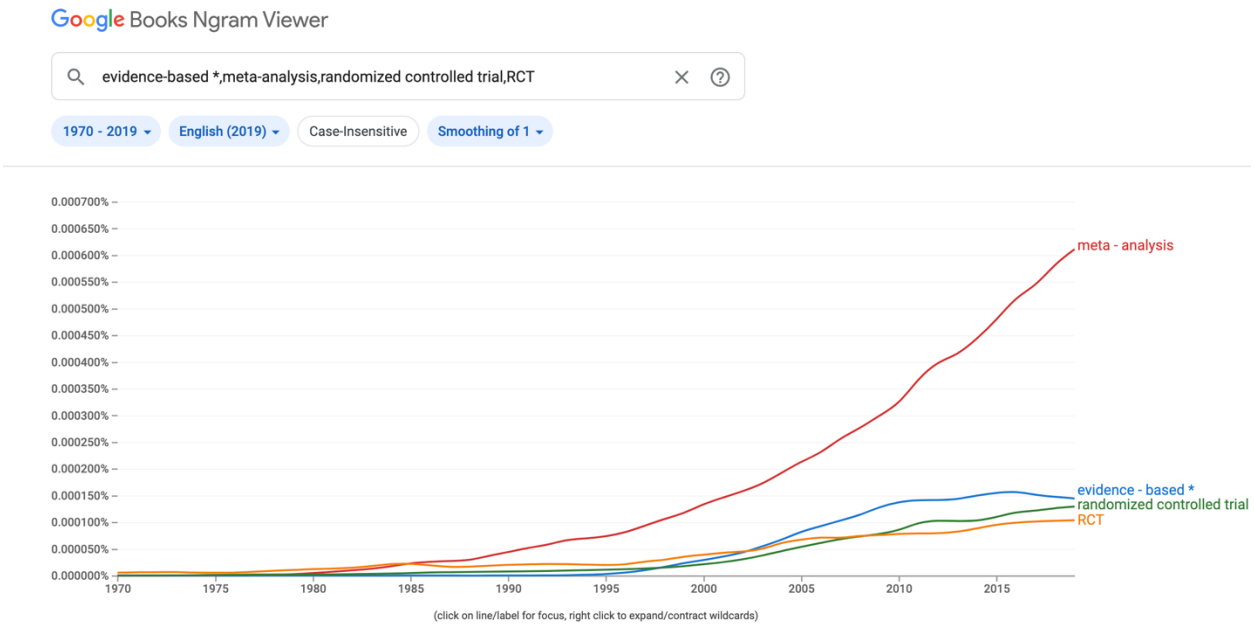


Figure 3: Ngram of “evidence-based *, meta-analysis, randomized controlled trial, RCT”


Google Books Ngram Viewer
evidence-based *,meta-analysis,randomized controlled trial,RCT
1970 - 2019
English (2019)
Case-Insensitive
Smoothing of 1
0.000700%
0.000650%
0.000600%
0.000550%
0.000500%
0.000450%
0.000400%
0.000350%
0.000300%
0.000250%
0.000200%
0.000150%
0.000100%
0.000050%
0.000000%
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
meta - analysis
evidence - based *
randomized controlled trial
RCT
(click on line/label for focus, right click to expand/contract wildcards)

References

Adams, Marilyn Jager. "Beginning to read: Thinking and learning about print." 1994.

Airy, G. B. *On the algebraical and numerical theory of errors of observations and the combination of observations*. London: Macmillan. 1861

Anderson, Richard C. "Becoming a nation of readers: The report of the Commission on Reading." 1985.

Citizens' Commission on Civil Rights, "Analysis of President George W. Bush's Education Plan." Washington, DC: Citizens' Commission on Civil Rights, 2001. https://web.archive.org/web/20081113071946/http://www.cccr.org/BushPlan.pdf

U.S. Congress, "Why Children Can't Read", HRG-1997-HEW-0082, 1997. https://congressional-proquest-com.proxy.uchicago.edu/congressional/docview/t29.d30.hrg-1997-hew-0082?accountid=14657

AERA Report From the Hill, "Gorillas in Our Midst: Emerging Themes on How to Improve Education Research," 1999. http://www.aera.net/gov/rpn/n-06.htm

Baron, Jon. "A Brief History of Evidence-Based Policy." The ANNALS of the American Academy of Political and Social Science 678, no. 1 (July 2018): 40–50. https://doi.org/10.1177/0002716218763128.

Beard, R. "The National Literacy Strategy: review of research and other related evidence." London, Department for Education and Employmen, 1999.

Beard, R. "The Relevance of Educational Research" Oxford Review of Education, 2000, pp. 421-436. https://www.jstor.org/stable/1050768

Bond, Guy L., and Robert Dykstra. "The cooperative research program in first-grade reading instruction." *Reading research quarterly* (1967): 5-142.

Boruch, R., Petrosino, A., & Chalmers, I. "The Campbell Collaboration: A proposal for systematic, multinational, and continuous reviews of evidence." In P. Davies, A. Petrosino, & I. Chalmers (Eds.), The effects of social and educational interventions: Developing an 32 Evaluation & the Health Professions, 1999.

Brookings Press Forum, 1999, http://stanford.edu/~hakuta/www/docs/brookings.html

Brown, A. James G., Lauren R. B., Mehan, H., and Lampert, M.. "Recommendations regarding research priorities: An advisory report to the National Educational Research Policy and Priorities Board." New York: National Academy of Education, 1999.

Cartwright, Nancy, and Jeremy Hardie. Evidence-based policy: A practical guide to doing it better. Oxford University Press, 2012.

Chall, Jeanne S. *The great debate*. McGraw-Hill, 1967.

Chalmers, Hedges & Cooper. A brief history of research synthesis. *Evaluation and the Health Professions,* 2002.

Cooper, H., & Hedges, L.V. *The handbook of research synthesis.* New York: Russell Sage Foundation, 1994.

Eisenhart, Margaret, and Lisa Towne. "Contestation and change in national policy on "scientifically based" education research." Educational researcher 32, no. 7 (2003): 31-38.

Erickson, F., & Gutierrez, K. "Culture, rigor, and science in educational research." Educational Researcher, 31(2002), 21–24.

Glass, G.V. "Primary, secondary, and meta-analysis of research". Educational researcher(1976): 5(10), 3–8.

Graham, Patricia, Remarks at the National Academy of Sciences' release of Scientific Research in Education. National Research Council, Washington, D.C., January 7, 2002

Guyatt, G. Guidelines for reading literature reviews CMA J, 138 (1988), pp. 697-703

"Harris Master of Public Policy Brochure", (2020) https://harris.uchicago.edu/files/master_of_public_policy_brochure_0.pdf

Heckman, James J. "Evaluation Revisited." *Randomized Control Trials in the Field of Development: A Critical Perspective*. Oxford, 2020.

Hunt, M. *How science takes stock: The story of meta-analysis.* New York: Russell Sage, 1997.

Hirschman, Daniel, and Elizabeth Popp Berman. "Do economists make policies? On the political effects of economics." Socio-Economic Review 12, no. 4 (2014): 779-811.

Kim, J. "Research and the reading wars." In: Hess FM When Research Matters: How Scholarship Influences Education Policy. Cambridge, MA: Harvard Education Press ; 2008. pp. 89-111.

Knorr-Cetina, K. Epistemic cultures: How the sciences make knowledge. Harvard University Press, 1999.

Krause, Monika. Model Cases: On Canonical Research Objects and Sites. University of Chicago Press, 2021.

Latour, Bruno. Science in action: How to follow scientists and engineers through society. Harvard university press, 1987.

Last, J. M. *A dictionary of epidemiology.* Oxford: Oxford University Press. 2001

Manna and Petril. "Double Standard?" In: Hess FM When Research Matters: How Scholarship Influences Education Policy. Cambridge, MA: Harvard Education Press ; 2008.

Merton, Robert K. *The sociology of science: Theoretical and empirical investigations*. University of Chicago press, 1973.

Mosley, Jennifer E., Nicole P. Marwell, and Marci Ybarra. "How the "what works" movement is failing human service organizations, and what social work can do to fix it." Human Service Organizations: Management, Leadership & Governance 43, no. 4 (2019): 326-335.

Mosteller, F., & Boruch, R.F. (eds.). *Evidence matters: randomized trials in education research.* Brookings Institution Press. 2002.

National Reading Panel (US), National Institute of Child Health, Human Development (US), National Reading Excellence Initiative, National Institute for Literacy (US), United States. Public Health Service, and United States Department of Health. *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health, 2000.

National Research Council. "Preventing reading difficulties in young children." 1998.

National Research Council. *Improving student learning: A strategic plan for education research and its utilization*. National Academies Press, 1999.

NIE employees charge Reagan appointees with destroying their agency from within. (1982, November 15). Department of Education Weekly.

NIE turns from the study of social change to emphasize the “basics.” (1982, August 24). Education Daily, p.5.

Paterson, Frances R. A. "The Politics of Phonics," Journal of Curriculum and Supervision 15, no. 3 (Spring 2000): 179-211.

Paige, R. .Opening remarks. Rigorous Evidence: The Key to Progress in Education. Coalition for Evidence-Based Policy. 2002. Retrieved from http:// www.excelgov.org/usermedia/images/uploads/PDFs/Final_Online_Version_-_November_18_Forum_Transcript.doc

Pearson, K. Report on certain enteric fever inoculation statistics. British Medical Journal (1904), 3, 1243-1246.

Peterson et al., “Can We Make Education Policy on the Basis of Evidence? What Constitutes High Quality Education Research and How Can it Be Incorporated Into Policymaking?” 1999. stanford.edu/~hakuta/www/docs/brookings.html

Petrosino, P., Chalmers, I. The effects of social and educational interventions: Developing an infrastructure for international collaboration to prepare, maintain and promote the accessibility of systematic reviews of relevant research. London, England: School of Public Policy, University College London.1999.

Popper, Karl. The logic of scientific discovery. Routledge, 2005.

Pflaum, Susanna W., Herbert J. Walberg, Myra L. Karegianes, and Sue P. Rasher. "Reading instruction: A quantitative analysis." Educational Researcher 9, no. 7 (1980): 12-18.

Pyle, E., “Te$t Market: Sandy Kress Has Ridden High-Stakes Accountability to Fame and Fortune. What about the Kids?” Texas Observer, May 13, 2005.

Office of Inspector General, U.S. Department of Education. “Final Inspection Report: The Reading First Program’s Grant ApplicationProcess.” September. Washington, DC, 2006: ED-OIG/I13-F0017.https://www.everycrsreport.com/files/20080206_RL33246_8787ef7fc2c8246a8e76cef7f76cd9d7a12053e5.html#fn18

U.S. Congress, “Reading Excellence Act (105 H.R. 2614 RFS Referred to Senate Committee).” Bill Text. (November 09, 1997)

U.S. Congress, “Reading Excellence Act (105 H.R. 2614 EAS Engrossed Amendment Senate).” Bill Text. (January 01, 1998)

U.S Congress, Why Children Can't Read, congressional hearing, 1997.

Shanahan, Timothy. "Research-based reading instruction: Myths about the National Reading Panel report." The Reading Teacher 56, no. 7 (2003): 646-655.

Slavin, Robert E. "Evidence-based education policies: Transforming educational practice and research." *Educational researcher* 31, no. 7 (2002): 15-21.

Slavin, R. E. "A reader's guide to scientifically based research." Educational Leadership (2003), 60 (5).

Smith, Adrian FM. "Mad cows and ecstasy: chance and choice in an evidence-based society." *Journal-Royal Statistical Society Series A* 159 (1996): 367-384.

Star, Susan Leigh, and James R. Griesemer. "Institutional ecology,translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." Social studies of science 19, no. 3 (1989): 387-420.

Zodhiates, Philip Phaedon. "Bureaucrats and Politicians: The National Institute of Education and Educational Research Under Reagan." Order No. 8907626, Harvard University, 1988. http://proxy.uchicago.edu/login?url=https://www.proquest.com/dissertations-theses/bureaucrats-politicians-national-institute/docview/303564477/se-2?accountid=14657.

The PCAST *Report to the President on the Use of Technology to Strengthen K-12 Education in the United States* can be read on-line at http://www.whitehouse.gov/WH/EOP/OSTP and copies can be requested through e-mail"information@ostp.eop.gov.

Yatvin, J. (2002). Babes in the woods: The wanderings of the National Reading Panel. Phi Delta Kappan, 83, 364–369.

Rhodes, Jesse H. *An education in politics*. Cornell University Press, 2012.

Teske, Paul. "President Bush and the US Department of Education: The Texas Mafia, Scientific Education Policy, and No State Left Behind." In *President George W. Bush's Influence over Bureaucracy and Policy*, pp. 105-121. Palgrave Macmillan, New York, 2009.