The 2012 Morris Hansen Lecture: Thank You Morris, et al., For Westat, et al.

Kenneth Prewitt[1]

This article, delivered as the 22nd Memorial Morris Hansen lecture, argues that the contract houses, typified by Westat, are uniquely situated in the cluster of institutions, practices, and principles that collectively constitute a bridge between scientific evidence on the one hand and public policy on the other. This cluster is defined in *The Use of Science as Evidence in Public Policy* as a policy enterprise that generates a form of *social knowledge* on which modern economies, policies, and societies depend (National Research Council 2012).

The policy enterprise in the U. S. largely took shape in the first half of the twentieth century, when sample surveys and inferential statistics matured into an information system that provided reliable and timely social knowledge relevant to the nation's policy choices. In ways described shortly, Westat and other social science organizations that respond to "request for proposals" (RFP) from the government for social data and social analysis came to occupy a unique niche.

The larger question addressed is whether the policy enterprise as we know it is prepared for the tsunami beginning to encroach on its territory. Is it going to be swamped by a data tsunami that takes information from very different sources than the familiar census/survey methods?

Key words: Policy enterprise; RFP; scientific integrity; scientific productivity; boundary organizations; big data.

1.   What's The *et al* In the Title About?

In my title, Morris is of course Morris Hansen, and though his contributions are properly celebrated, he was one of many who helped establish the dozen or so flagship contract houses in the U.S that produce a large share of the survey based social knowledge on which I focus. Other important contributors at Westat, for example, include Ed Bryant, Joe Hunt and Joe Waksberg. Well before Westat's founding in 1963, however, came the National Opinion Research Center (NORC) founded by Harry H. Field in 1941 and located at the University of Chicago. Field was advised by three social science giants of the period: Gordon Allport, Hadly Cantril, and Sam Stouffer. Other NORC luminaries are Paul Sheatsley, Peter Rossi, and Norman Bradburn. The Institute of Social Research (ISR) founded in 1946 at the University of Michigan brings to mind Rensis Likert, its storied first director, and also in its early days Charles Cannell, George Katona, Leslie Kish, and Dorwin Cartwright, who were soon followed by Donald Campbell, Philip Converse and many others. Clark Apt, and his Apt Associates, pioneered the for-profit base of contract

[1] Professor at Columbia University, 535 West 116th Street, New York, NY 10027, U.S.A. Email: Kp2058@columbia.edu

research. Among another dozen or so flagship institutions of the sort under discussion are Mathematica, RAND, and RTI.

These institutions are more than "contract houses," but I will use that term to draw attention to the RFP mechanism that shaped the growth of a new way to collect social information in the second half of the twentieth century.

The new way was much more than taking to scale the methodology of survey research, though it involved that. It was also and perhaps more consequentially a new way to structure the relation between science and public policy. It is this structure that needs re-engineering for the twenty-first century information environment.

2.   How Institutions Matter

Ian McNeely, writing with Lisa Wolverton (2008), observes that in the western tradition "organizing knowledge became as important as knowledge itself" (p. xix). In fact, "'the west' is better defined by its institutions for organizing knowledge than as a set of cultural values or a region of the world" (p. xiv). Libraries – stretching back to Alexandria and Timbuktu – were places where written knowledge could be stored in one place, made accessible, and added to, helping to establish the idea that knowledge is cumulative. Monasteries used multiple copies of the same text to standardize religious instruction that built a church with priests and parishioners scattered across a large region of the world. Museums, basically an Enlightenment creation, collected flora and fauna, which established a natural science based on taxonomy and comparison. Another major step was the nineteenth century invention of the research universities, initiated in the United States when Johns Hopkins, Clarke University, and the University of Chicago combined two traditions: Germany's great research institutes with England's great teaching colleges. These new institutions promoted scientific specialization, giving us the familiar disciplines housed in departments.

Subtract libraries, monasteries, museums, or research universities from western history and its substantive knowledge would of course look very different. I do not promote contract houses into the distinguished company of these institutions, but I do hold that they should be seen as new institutional forms that helped establish the conditions for social knowledge production from the 1950s to today. It has been a specific type of knowledge – largely quantitative and intended for use in public policy. The importance of this is apparent under the next heading.

3.   The Policy Enterprise

America, fresh from its victories in World War II, and especially appreciative of the role of science – radar, penicillin, and of course the atomic bomb – latched on to the idea that America's universities could build a knowledge base relevant to economic growth and national security. This spawned the National Science Foundation (NSF), Defense Advance Research Projects Agency (DARPA), government laboratories such as Lawrence-Livermore and the National Institutes of Health (NIH) laboratories. Big science had arrived. By the mid-1960s, social science was in the big science game. The Coleman Report (Coleman 1996) with its database of 600,000 students and 40,000 teachers in 6,000

schools addressed a "big question" – racial inequality in America's schools. By the standards of social science in 1964, this was big science.

The policy enterprise was underway. Its institutional base includes: policy think-tanks, now estimated at more than 5,000 worldwide, with 1,800 in the U.S. and about 450 of these in Washington; public policy schools and programs in higher education, providing career training for thousands of positions in the policy enterprise; for-profit consulting firms drawing on social science; advocacy groups and professional lobbyists organized around particular policy goals, repackaging social science to their purposes; government units for social and economic analysis in the executive branch, Office of Management and Budget (OMB) of course, but also in the intelligence agencies and many domestic departments such as education, health, and human services; other units attached to Congress, including the Congressional Budget Office and the Congressional Reference Service; and, of course, the statistical agencies and their arrangements to make data widely available for analysis. New research fields emerged, obviously around the statistical underpinnings of probability sample surveys but also fields such as science and technology studies and what came to be known as the knowledge utilization specialty.

I will not dwell on this history; in her discussion Margo Anderson goes into greater historical depth and advances an important structural hypothesis, with which I agree.

In short, the policy industry is a multi-billion dollar cluster of institutions bringing social knowledge to bear on policy design, implementation, and evaluation. It has given birth to evaluation research, social indicators, ranking schemes, performance metrics, evidence-based policy and practice, accountability measures, best practice, and, more generally, the quantification of policymaking.

Of course none of this is unique to this country. You can identify some version of a policy enterprise in at least 150 countries – China to Ghana, Brazil to Jordan, Britain to Australia.

I return below to the unique place of the contract houses in the policy enterprise, but here insert a historical footnote of considerable importance. The contract houses that began to take shape in the 1940s and 1950s, some rooted in research universities (NORC and ISR especially) and others drawing personnel from the census and statistical agencies (Westat being a prime example), were built on a foundation of scientific and academic principles. This was equally true for the for-profit as for the non-profit institutions.

For instance, there was the expectation that the contract houses would produce science of a quality that matched what was found in universities. In fact, when it made sense, they would go beyond current survey practices. NORC, from the day of its founding, insisted that interviewers be treated not as casual workers, but as specialists key to the quality of the survey effort (see Sheatsley 1981–82). Interviewers were trained before being sent into the field. More generally, the competition for contracts led to constant quality innovations– area probability sampling and randomized digital dialing being dramatic examples (Bryant 1997). Personnel in the contract houses would be active in professional academic societies; in the case of the America Association for Public Opinion Research (AAPOR), Henry Field and his NORC colleagues were among its founders.

The pioneers also insisted on openly produced and publicly accessible knowledge. As early as 1947 the ISR was turning down funders who would not agree to make results publicly available, including a study proposed by the Ford Motor Company. ISR did not

refuse corporate sponsors, but applied two criteria: Was it a problem of social importance and would the results be published (Frantilla 1998).

Much more could be said on the principles that guided the founding of the contract houses, whether nonprofit or for profit, but the principles continue today and thus are familiar to and probably taken for granted by readers of this Journal.

4. The Integrity and Productivity of Social Knowledge

I borrow the terms integrity and productivity from David Guston (2000). He points out that when the government purchases scientific knowledge, it needs a guarantee of the integrity and the productivity of the science. Integrity refers to the absence of fraud or other substandard practices.

Productivity is a more complicated concept. Certainly it involves cost-effective performance, but the term has a broader meaning. Productive knowledge is that which meets the exacting criteria of "usefulness" to its public sponsors.

Both integrity and productivity are reasonable demands. A government that wasted public funds on fraudulent or useless social knowledge puts its own legitimacy at risk. Fraud is fairly straightforward, but productivity takes us into tricky territory. This is clear if we compare social knowledge to engineering or biological knowledge. What constitutes fraud is comparable across these knowledge sources – the deliberate effort to get the user to accept as true something which is false.

No such comparable standard is available for productivity. What do the engineers or the natural sciences promise when they claim their knowledge is "useful"? Engineers promise bridges that won't fall down; physicists a missile that will hit its targets; biochemists vaccines that prevent diseases.

But what does the policy enterprise promise? The promise is not, or at least should not be, better data. It is not, or at least should not be, better policy. On first reading you may find this counterintuitive. But better social data does not, in the policy world, translate into more usable data. And better policy begs such questions as for whom, under what conditions, over what time frame. These are political more than scientific questions. For additional explanation, see *Using Science as Evidence in Public Policy* (National Research Council 2012).

This does not mean that we lack grounds on which to claim that social knowledge is productive. Social science can promise to describe social conditions, and whether they are changing, in what direction, and with what velocity. That is, we can describe an aging society and its many features that might require policy attention. Social science can also make estimates about what is likely to happen if a policy intervention occurs, and, post hoc, what did happen as a consequence of that intervention.

If the policy maker wants to know how rapidly the population is ageing and what that means for the social security system, we can provide useful knowledge. If he or she asks what is likely to happen if the age at which social security starts paying benefits is moved up or down, we can provide useful knowledge.

Incidentally, this limited but workable definition of productivity has nothing to do with basic versus applied science, terms that tell us nothing about whether or not the policy maker will find the science useful. And for the policy maker it is beside the point whether scientists call their knowledge discipline-based or interdisciplinary. These distinctions

might be of interest among scientists, but to the policy maker what matters is whether the knowledge can be put to productive use in the context of the available policy choices.

I return below to the issue of usefulness, but first I want to return to the period in which contract research was getting underway.

5. A Tacit Agreement Fades

In the postwar period, science policy was heavily influenced by Science the Endless Frontier (Bush 1945), especially its strong assertion on behalf of government investment in basic science. The National Science Foundation was the most visible result on the federal scene; peer review science was ascendant. Science policy in the U.S. was generally based on the assumption that science could perform most productively if free of government control, though of course not free of public obligations. Science, solely concerned with the truth, did not need to be tightly regulated or directed. Its internal policing mechanisms would guarantee scientific integrity and its culture of responsibility would guarantee productivity.

As Don Price (1965) makes clear, this tacit agreement was short lived. The generalized trust in science was gradually replaced by incentives, oversight, performance measures, and related institutional arrangements by which the government assures itself that publicly funded science meets the criteria of integrity and productivity. The current reflection of this is the concern in science policy circles with metrics that can assess "broader impacts" of the government's investment in science.

Principal-agent theory helps us see what the issues are. The principal – the government – lacks the expertise to produce knowledge it needs. It needs to delegate to an agent the task of producing expert knowledge, that is, scientific research. If the government trusts the integrity and productivity of its agent, nothing else is called for. The problem of science policy is solved.

If, however, the government worries that perhaps not all of the knowledge it is purchasing is free of fraud or rent seeking, and worries even more that scientists have an inclination to be more concerned with peer approval than in producing what society needs, the government will monitor its expert agents and create incentives to influence their behavior in desired directions.

RFP is an obvious mechanism. The basic RFP design specifies in detail what research is to be carried out and uses price as a prime criteria for awarding the contract. This is not the place to assess whether the particularities of the RFP optimize productive knowledge, though thoughtful participants believe it does not. At one point there was hope that the RFP might be designed differently – that is, the government agency would define its objectives, make clear how much it intended to spend, and then judge proposals in terms of their scientific merits within the budget constraints provided. This idea was never seriously considered.

My interest is a particular consequence of RFPs in structuring social knowledge intended for use in public policy. RFPs uniquely position the contract houses on the boundary between science and government. In fact, the contract house can be understood as a "boundary organization," using the term in the specific way it has been developed in science and technology studies, especially by Shelia Jasanoff (1990). She writes that

bringing the scientific community into "decisionmaking produces a stronger consensus than any achievable through the agency's in-house expertise alone" (p. 237).

As elaborated at greater length in Dan Gaylin's discussion, a boundary organization, in conducting research guided by an RFP, is responding to what the government has defined for itself as productive knowledge. This allows the government to claim that they are holding their agents accountable for the integrity and productivity of their work. Simultaneously, the contract house is advancing scientific goals. If the product is statistical data, the data will meet standard scientific criteria. If the product is analysis, the reports and the methods by which it is produced will enter the stream of public social knowledge. Stated most simply, the contract houses identified above are a successful example of institutions that respect the government's need for reliable and productive social knowledge and do so without compromise to scientific principles.

This is not a trivial observation. If the policy enterprise alluded to above is a four to five billion dollar annual effort, the contract houses are responsible for perhaps half of the public funds involved.

On integrity – I know of no instance of scientific fraud associated with any of our flagship contract houses. Certainly there have been cost over-runs and an inability to meet deadlines or other targets such as high response rates. It does not trivialize these failures to note that the occurrence is low and the magnitude is nothing at the scale routinely reported about the government's purchase of weapon systems.

The issue of productivity is, of course, more difficult to assess – but the growth in dollar volume, in size of studies, in methodological innovation, in timely delivery, and most other metrics we might mention is indirect evidence that the contract houses have proven their worth against the exacting criteria that principle-agent theory puts on the table. They have been productive.

At a moment of skepticism, even cynicism, about the contribution of science to public policy, it is reassuring to have this "success" story available. It is even more important to have this asset in place as we turn to a challenge unimaginable to the founders of NORC, ISR, RTI, Rand, Mathematica, and, of course, Westat.

6. The Digital Data Tsunami

The tsunami is the large and growing supply of social information from digital sources – credit card transactions, surveillance cameras, internet search patterns, social media, with many more technologies yet to come. Before addressing the challenge this poses to the policy enterprise, I want to emphasize what is probably the greatest achievement of the practices and principles we associate with that enterprise.

It has produced a high quality, *shared* information base for the nation's polity, economy, and society. This has made for healthy democratic debate about policy choices. This is most clearly seen in the arrival of the Great Society policies of the 1960s and 70s, and the critique and partial dismantling of them in the decades since. The critical point is that the information order used to design and implement the Great Society policies was used to evaluate and then challenge them as negative unintended consequences were documented – welfare dependency; the hidden taxes in government regulations; the poor record of urban school reform leading to demands for choice; the mixed record of affirmative action

as it benefited upper income African Americans and immigrant groups, especially Asians and West Indians, for whom it was not originally intended.

But it is not the policies that interest us here. It is the fact that America created social knowledge facilitating robust debate about public policies. The statistical underpinnings of the debates had characteristics so familiar that we assume their permanence – high quality, peer reviewed, publicly disseminated, theoretically guided, data representative of the entire population, all provided as a public good.

Whether the digital data tsunami now on its way will disrupt the system carefully assembled over the last century is unclear. I have no crystal ball. But at least in its early days we know much of the digital data from the private sector to be proprietary, of unknown quality, guided less by social theory than commercial benefit, largely unconcerned with privacy/confidentiality, unrepresentative in troubling ways and only incidentally provided as a public good.

These weaknesses notwithstanding, it is difficult for the government to let digital data go unused. It is too much, too cheap, too easy. What is already underway in the national security sector will surely migrate to the domestic policy sector on which today's policy enterprise is largely focused. On March 29, 2012, the federal government announced the "Big Data Research and Development Initiative," with the Office of Science and Technology Policy challenging "industry, research universities, and non-profits to join the Administration to make the most of the opportunities created by Big Data" (Office of Science and Technology Policy 2012). A report issued a few months later, titled the "Big Data Gap", observes that the "promise of big data is locked away in unused or inaccessible data" and that most federal "agencies are still years away from using it" (Big Data Gap 2012). We are not surprised to learn that government wants to use the tsunami of digital data, nor to learn that it is ill-prepared to do so. One concern is that it is the IT departments and not the program units that "own" the data.

The questions are obvious. What institutional platform will collect, manage, house, and analyze the tsunami of digital data? Where will "social knowledge" live in the new information order? Will today's statistical agencies remain involved? What parts of government will draft the RFPs? Who will bid? Today's academically rooted contract houses or Google, Apple, McKinsey, and commercial firms yet to be invented? For readers of this Journal, deeply committed to knowledge for the public good, the question is whether you will be in the game or marginalized – important to the second half of the last century but historical relics by the second half of this century.

Five issues are worrisome: ethics, quality, representativeness, theory, and productivity. This is not an exhaustive list of worries, but enough to alert us to the challenge.

Ethics. The policy enterprise worked out some difficult ethical challenges – informed consent, privacy, confidentiality, and access. How can millions of surveillance cameras and billions of electronic censor devices offer informed consent? Some of you have heard me complain about the failure to distinguish between privacy and confidentiality in the census/survey world (Prewitt 2011). Privacy is the public saying, "you don't have a right to know that about me" – don't ask. Confidentiality is the public saying, "don't share my information in any way that it could be used against me" – don't tell. I doubt that privacy protection has a future in the digital data world, but confidentiality protections are more plausible – though only if taken seriously. We have figured out how to protect

confidentiality and still provide robust access for legitimate policy analysis – migrating the arrangements from census/survey data to digital data should be possible, but will require alert attention.

Data quality. We know what quality means for census/survey data – sampling error, respondent burden, cognitive bias, imputation, response rates, external validity, attrition in panel studies, and on and on.

What are the equivalent quality issues for digital data? There are few professional meetings with hundreds of papers debating the error structure of digital data. There is no generally accepted understanding of what constitutes errors when it is machines collecting data from other machines and passing the data along to algorithms for analysis and on to clouds for storage and dissemination.

Representativeness. The probability sampling method basic to the census/survey-based information order provides deep theory about what groups are represented in any given study. Representativeness has not been a major issue in the analysis of digital data. On the one hand, the entire Facebook population or users of a specific browser are, in principle, available. On the other hand, persons who are not online and not candidates to be enrolled – too old, too young, too poor, and so on – are of marginal interest to businesses with a product to sell. Again, the unrepresentativeness of digital data is a problem that can be addressed, but not likely by the current providers of digital data.

Theory. I noted Jim Coleman's pioneering study of racial equality. It was less its size than its finding that we recall today – he reported that family characteristics had as much bearing on educational outcomes as school characteristics – cost per pupil, classroom size, and so on. We now examine the out-of-school versus in-school influences on school performance with batteries of questions, longitudinal data sets, and powerful statistical tools. But some readers will recall that Coleman added but two simple questions: do parents go to Parent-Teacher Association (PTA) meetings; is there an encyclopedia in the house. Coleman brought sociological theory to bear.

Survey data may be case poor but they are variable rich. It is our control over the variables that provides theory-derived social knowledge. Digital data are case rich and variable poor, and insofar as the variables are constructed by theory it is likely to be theory drawn from marketing concerns with consumer behavior. The theory deficit in digital data is a problem that can be fixed, but it will take an active three-way partnership that replicates what the contract houses helped build: government defining its policy needs, theory-designed measurement strategy, and data collection expertise.

This twentieth-century partnership of government, academic social science and contract house created a productive social knowledge base because each of the three players brought relevant technical expertise to the table and a shared understanding of what was required. It is a model for what needs to be created as new data, in very large quantities, becomes available.

Productivity. I opined above that the failure to persuasively explain productivity – useful knowledge that was used in public policy – weakened public confidence in the social sciences. The contract houses have been an important corrective, especially in government circles that appreciate the value of high quality data. But this time around – given the hovering presence of commercial players with deep pockets and lobbying skills – even that might not be enough.

7.   References

Big Data Gap. (2012). Based on a survey of 151 Federal government CIOs and IT Managers in March, 2012. Available at: http://www.meritalk.com/bigdatagap (accessed October 15, 2012).

Bryant, E.C. (1997). Westat: Still Counting. Maryland: Rockville.

Bush, V. (1945). Science The Endless Frontier. A Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development. Washington, DC: United States Government Printing Office.

Coleman, J.S. (1966). Equality of Educational Opportunity. Washington DC: US Department of Health, Education and Welfare, Office of Education.

Frantilla, A. (1998). Social Science in the Public Interest: A Fiftieth Year History of the Institute for Social Research. Ann Arbor, Michigan: The University of Michigan, Bentley Historical Library, Bulletin No. 45. September 1998.

Guston, D.H. (2000). Between Politics and Science: Assuring the Integrity and Productivity of Research. Cambridge, United Kingdom: Cambridge University Press.

Jasanoff, S. (1990). The Fifth Branch: Science Advisers as Policymakers. Cambridge, MA: Harvard University Press, 237.

McNeely, I.F. and Wolverton, L. (2008). Reinventing Knowledge: From Alexandria to the Internet. New York: W.W. Norton & Company. Norton paperback edition, 2009.

National Research Council (2012). Using Science as Evidence in Public Policy. Committee on the Use of Social Science Knowledge in Public Policy. Division of Behavioral and Social Sciences and Education, K. Prewitt, T.A. Schwandt, and M.L. Straf (eds). Washington, DC: The National Academies Press.

Office of Science and Technology Policy (2012). Big Data is a Big Deal. Available at: http://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal (accessed October 15, 2012).

Prewitt, K. (2011). Why It Matters to Distinguish Between Privacy & Confidentiality. Journal of Privacy and Confidentiality, 3(2), Article 3. Available at: http://repository.cmu.edu/jpc/vol3/iss2/3/.

Price, D.K. (1965). The Scientific Estate. Cambridge, MA: Harvard University Press.

Sheatsley, P. (1981–82). NORC: The First Forty Years. Chicago: University of Chicago. NORC Report 1081–82.