Bob Behn's Public Management Report

An occasional (and maybe insightful) examination of the issues, dilemmas, challenges, and opportunities in leadership, governance, management, and performance in public agencies.

On why public executives need to remember that

The Data Don't Speak for Themselves

The glass sits on the table, partially filled with a liquid. The level of the liquid is midway between the top and bottom.

Looking at the glass, the liquid, and its level, the optimist declares: "The glass is half full." The pessimist does the same but reaches a different conclusion: "The glass is half empty."

The optimist and the pessimist are using precisely the same data point. Yet they reach very different conclusions. Clearly, this easily observed, easily measured, very unambiguous data point does not speak with a single, unequivocal message.

Despite the universal appeal of the seductive cliché, the data *never* speak for themselves.

When the data speak, they do so only through some framework, some theory, some causal model, some logical construct, some perception of the world and how it works. After all, any set of data is just a collection of abstract numbers. The data acquire meaning only when they are connected to some version of reality.

For decades, astute observers have railed against this data-speak-for-themselves nonsense. In 1932, Carl Becker, then president of the American Historical Association, observed that "to suppose that the facts, once established in all their fullness, will 'speak for themselves' is an illusion."

"Be aware of the intellectual traditions and choices out of which the 'data' emerge," cautioned Gary Marx, professor emeritus at M.I.T. "The facts do not speak for themselves. Look for the ventriloquists in the wings."

When James Heckman accepted the Nobel Prize in economics, he was unequivocal: "The data do not speak for themselves."

Yet, the cliché lives.

The data are just data—collections of apparently random ones and zeros. To interpret these data, we need a framework. We need a lens through which to observe the data and extract from the otherwise incomprehensible gibberish some information: a coher-

ent pattern, a revealing story, a valuable lesson, an implication for action.

Sometimes, to analyze the data, we are careful and explicit in selecting a framework. Often, however, we choose a framework by default, employing a very cherished but strictly implicit theory about how the world works (or should work). And, given the large number of different frameworks that can be used to analyze any collection of data, the number of possible interpretations, implications, and conclusions that can be drawn from these data is equally large.

> People who claim that "the data speak for themselves," are saying nothing more than: "If you employ *my* favorite framework for looking at the data, you can only interpret the data as I do and thus can only arrive at my policy prescriptions."

The choice of an analytical framework may be based on previous analyses—on the successes and failures of prior efforts to find patterns in similar data. Still, this choice is subjective. Moreover, the conclusions drawn from the data may be more dependent on the analytical framework chosen than on the data themselves.

This means that if different people —sincere, honest people—observe the data from different frameworks, they can (and often will) reach quite different (even contradictory) conclusions from identical data.

People who claim that "the data speak for themselves," are saying nothing more than: "If you employ *my* favorite framework for looking at the data, you will see precisely what I see." Moreover, these people are saying, "If you employ only *my* favorite framework, you can only interpret the data as I do and thus can only arrive at my policy prescriptions."

Those who assert that the data speak for themselves are not willing

to accept that there might exist other frameworks for looking at the data. They are not willing to accept that there might be alternative interpretations and thus different possible policy implications.

Peter Blau, in his classic book, *The Dynamics of Bureaucracy,* writes: "Data do not speak for themselves but only answer questions the investigator puts to them." In choosing a framework with which to examine some data, the analyst is choosing to ask some questions—and *not* to ask a large number of other questions.

The data do not speak independently of the analyst. The data do not speak independently of the framework. The data do not speak independently of the questions. The data only speak when an analyst, using a specific framework, asks a specific question. The data cannot answer any question until someone establishes (again, either implicitly or explicitly) a framework for creating an answer.

The data never talk back to the analyst: The data don't say: "Hey, you dumb analyst, you are using the wrong framework." The data don't say: "Please, analyst, if you would ask a different question, I could give you a much more revealing answer."

If the data spoke for themselves, we humans would always interpret every set of data in precisely the same way. Occasionally, however, this is not the case. In fact, it is very often not the case. Different humans can frequently reach polar opposite conclusions from precisely the same data. That is not necessarily because they are evil or stupid. It is simply because they are analyzing the data using different frameworks. ◼

Robert D. Behn is a lecturer at Harvard University's John F. Kennedy School of Government where he chairs the executive-education program "Driving Government Performance: Leadership Strategies that Produce Results." His latest publication is: *What All Mayors Would Like to Know About Baltimore's CitiStat Performance Strategy.*