GDN's Next Horizons Essay Contest 2014*
THE FUTURE OF DEVELOPMENT ASSISTANCE
Supported by the Bill and Melinda Gates Foundation

Winning Entry
"MORE AUTONOMY FOR DONOR ORGANIZATIONS AND THEIR AGENTS (SOMETIMES): BRINGING ORGANIZATIONAL BEHAVIOR AND MANAGEMENT THEORY TO FOREIGN AID DELIVERY"

Abstract
Management practice is the low-hanging fruit of foreign aid delivery, the area where marginal investment is most likely to yield results. Leveraging over 100 in-person interviews and econometric analysis of the world's largest database of development interventions (over 14,000 interventions over 40 years), which I have assembled, I argue that for some (but not all) foreign aid tasks, the move towards output measurement and away from field agent autonomy is likely to be detrimental to performance. More generally I argue that inadequate attention is paid to applying the rich, evidence-based literature on organizational behavior and optimal performance to the practice of foreign aid delivery. There is good reason to believe that we can greatly improve the impact of aid simply by thinking more deeply about how organizational structure affects performance and how optimal structure is a function of recipient country context and the specific task being undertaken. Much attention is spent trying to alter elements of the development equation that, while very important, are not terribly tractable to external intervention (e.g. political will, corruption). It is high time that we concentrate on the levers of development that *are* primarily within the control of aid donors and are also significant determinants of outcome: management, incentives, and organizational behavior in aid agencies.

Author
Dan Honig**

Introduction

As a practitioner I have found myself working on behalf of the governments of developing countries, sitting across the table from counterparties who represented multilateral and bilateral donor agencies. I have often been struck by the fact that the donor agents who seem to be doing the best job – those most interested in contributing to national development trajectories, public goods, and citizens' welfare – often seem to spend much of their time in conflict with their own organizations over procedures and process. The organizations' missions are all laudable, and when I have met senior leadership of bilateral and multilateral agencies, I almost invariably have walked away impressed. What's going wrong, then? What stands between good leaders and their agents? Why does the forest of actual development impact so often get lost in the procedural trees?

This essay frames measurement as, under certain circumstances, circumscribing judgment by field level operators to move interventions forward based on contextual knowledge and experience. The judgment-versus-measurement debate is very much a live one in development, with some scholars noting the ongoing debate among practitioners (cf. Gulrajani 2011) and others arguing for a more iterative, agent-judgment-driven approach which plans less ex-ante and instead adapts to the soft, contextual information of recipient-country environments (Andrews, Pritchett, and Woolcock 2012; Barder 2009; Booth 2013; Easterly 2014; Ramalingam 2013). This essay argues that while neither extreme of the judgment vs. measurement debate is the optimal setting for aid organizations, the overall balance in international development organizations is off; aid organizations would see greater results if they and their field staff were given greater autonomy and subjected to fewer output-based reporting requirements.

What Theory Tells Us

One of the most striking features of aid delivery is the remarkable similarity in the structure of delivery of dissimilar tasks (judicial reform and road construction, say), and how little these structures change in response to dissimilar contexts (for example, whether the road under construction is in China or Chad, Turkey or Timor-Leste).[1] Long-established, foundational work in organizational behavior and management would suggest that whatever the optimal management practice might be, it is very unlikely to be the same across tasks and contexts.

One of the central lessons of these literatures is that context matters for optimal strategy, and that autonomy is a key lever for organizations, both public and private (Carpenter 2001; Huber and Shipan 2002; Lawrence and Lorsch 1967; Thompson 1967; Wilson 1989). Key in this theoretical literature is uncertainty, with greater environmental volatility (and thus less potential for task routinization) associated with a higher optimal level of agent discretion and

---

[1] While some organizations have special facilities or operational procedures for fragile states or particular tasks, this covers a small percentage of development projects as a whole.

autonomy (March and Simon 1958; Perrow 1967; Thompson 1967). In a more uncertain environment, flexibility and autonomy will be more highly prized.

Put another way, setting clear and realistic objectives and measuring the right things – and incentivizing agents to deliver them – is clearly the best possible strategy; economics and contract theory make this abundantly clear. In tasks that are not tractable to output measurement, management by measurement may prove ineffective, but nonetheless crowd out the agent autonomy necessary for optimal organizational performance. In submitting this essay to the GDN Next Horizons Essay Contest, I am aware, of course, that a key partner – the Gates Foundation – is often associated with the push towards measurement in development. I would be the last to suggest that to the extent the Gates Foundation has pushed towards measurement it has not been positive. In fact, I would argue that the need to consider the nature of the task and environment in deciding what to measure is very much in keeping with the Gates vision. In his 2013 Annual Letter, Bill Gates highlights the development impact of measuring vaccine transmission and coverage rates rather than simply sending out health personnel to conduct vaccine drives (Gates 2013). He also, however, seems to implicitly endorse this work's conditional view that measurement's role depends on its ability to provide timely, appropriate, non-distortionary feedback in saying "You can achieve amazing progress if you set a clear goal and find a measure that will drive progress toward that goal" (ibid, p.1). A well-aligned measure and a clear objective are necessary conditions for measurement to be optimally beneficial. Measurement leads to greater production of whatever is measured. The question is when that is likely to be a good thing (e.g. aid linked to vaccination rates), and when not.

Innovative aid thinkers such as Nancy Birdsall and Bill Savedoff have pushed aid organizations to focus on stable long-term measures tied to clear objectives in their advocacy of Cash on Delivery aid (e.g. Birdsall & Savedoff 2010), a modality which promises greater autonomy to implementers by conditioning aid on the accomplishment of mutually agreed quantifiable goals. Laudable though this solution is, it may not always be implementable. It is not always possible to find good measures and clear objectives, to say nothing of political authorizing environments, project cycles, external shocks, and other constraints which limit the ability of aid agencies to contract on long term outcomes. As such we are often left with the output measurement regime under which aid generally labors at present. Cash on Delivery is an excellent tool, and may well be part of an optimal portfolio of aid instruments for a given country; it may not however be the whole of an optimal portfolio.

Some types of tasks are more tractable to measurement and external monitoring than others. If an organization is constructing a building, clear standards of output quality can be observed or contracted on. If an organization is supporting civil service reform, it becomes much harder to develop appropriate short-term output measures against which results can be measured. The notion that tasks are inherently different and pose different measurement challenges is well articulated in the management control systems literature on private sector contexts. It is

also a critical part of some of the most prominent theorizing in the public administration literature on bureaucratic functioning and contracting (Brown and Potoski 2003, 2005; Wilson 1989).

One critical difference between these two contexts is the degree to which tacit knowledge (Polanyi 1966) or soft information is critical to success. Stein defines soft information as

> [i]nformation that cannot be directly verified by anyone other than the agent who produces it. For example, a loan officer who has worked with a small-company president may come to believe that the president is honest and hardworking – in other words, the classic candidate for an unsecured "character loan." Unfortunately, these attributes cannot be unambiguously documented in a report that the loan officer can pass on to his superiors (2002, p. 1892).

In international development implementation, soft information includes (but is not limited to) assessing ministry personnel and their motivations; structuring or revising an intervention to maximize its likelihood of being line with the interests of important political actors and thus becoming fully implemented; or simply knowing whether an intervention under implementation is headed in the right direction. Many things that are hard to codify and communicate up a hierarchy may well be critical to a development intervention's success.[2]

Autonomy allows field staff to make judgments about program design, management, and revision that rely on soft information; that is, to navigate by judgment. Autonomy also leads to higher-quality staff (who migrate to roles that give them greater opportunity to make decisions) and superior organizational learning. More autonomous agencies can design interventions that are more appropriately calibrated and more likely to be "owned", and hence implemented and sustained, by domestic government actors. Such agencies are more able to adjust or revise interventions when needed, and can more adequately make more appropriate day-to-day implementation and supervision decisions.[3] More autonomous agencies are likely to be better able to work jointly with domestic government actors both in designing and adjusting interventions. As such, the push for more autonomy is in service of some of some of the core insights of the High Level Fora on Aid Effectiveness (i.e. Paris, Accra, Busan) which argued for giving developing countries greater agency and linked this agency to greater propensity for success and sustainability, particularly in difficult and complex environments.

---

[2] This line of argument shares much with a separate literature on observability and top-down control pioneered by James Scott's *Seeing Like a State* and the myriad "Seeing Like…" publications it has spawned. Soft information is, on this view, a first cousin of mētis, which Scott defines as "a wide array of practical skills and acquired intelligence in responding to a constantly changing natural and human environment" (Scott 1998, p. 313).

[3] Several mechanisms exist that incorporate soft information by autonomous agencies and agents, leading to better decisions and more successful development projects. These are explored in greater depth in qualitative case studies (Honig 2015).

4   *Next Horizons Essay Competition 2014-15 – Dan Honig*
"More autonomy for donor organizations and their agents (sometimes): bringing
organizational behavior and management theory to foreign aid delivery"

In sum, then, I am arguing that navigation by measurement will be most useful for relatively routine tasks and/or relatively predictable environments where (a) the desired outcomes are verifiable and thus contractible, and (b) it is easy to make frequent non-distortionary measurements which will also be stable, avoiding the problems of Goodhart's Law: when a measure becomes a target, it may cease being a good measure. Navigation by judgment, on the other hand, will be most useful when (a) tasks are difficult to routinize and/or environments are relatively unpredictable, and (b) it is hard to define appropriate targets ex-ante or find good measures.

One of the main reasons that organizations measure and engage in hierarchical control is for legitimacy, to 'show results' to a skeptical political authorizing environment or public (DiMaggio and Powell 1983; Meyer and Rowan 1977). This dynamic has been echoed in the aid literature (Easterly 2002; Eyben 2013), and connected to political authorizing environments and the fight for continued funding and resources (Barnett 2009; Bush 2011; McMahon 2001). Criticism, including that from political authorizers, constrains what IDOs do and what they imagine themselves capable of doing; as Tendler (1975) puts it regarding USAID:

> It has been generally recognized that criticism of the foreign aid program weakened [USAID] and kept it from doing what it wanted to do. Less understood is the fact that the process of living with criticism profoundly affected what the agency *wanted* to do and what it was capable of doing (p. 40).

Constraints emanating from political authorizing environments change the incentive structure of aid organizations and their agents, shifting them toward navigation by measurement and away from navigation by judgment. It is no surprise, then, that aid organizations often focus on changing what can be easily measured (policy, structures) in recipient countries, at the expense of areas where measurement is more difficult (Andrews 2011; Eyben 2013; Pritchett and Woolcock 2004).

There is perhaps no more eloquent description of these dynamics than that of Andrew Natsios, the former head of USAID. Natsios (2010) describes what he calls,

> Obsessive Measurement Disorder (OMD), an intellectual dysfunction rooted in the notion that counting everything in government programs (or private industry and increasingly some foundations) will produce better policy choices and improve management… [Relatedly] demands of the oversight committees of Congress for ever more information, more control systems, and more reports have diverted professional USAID (and now MCC) staff from program work to data collection and reporting requirements (p. 8).

Natsios relates the inappropriateness of measurement directly to the difficulty of measuring outcomes. His prescription is that "USAID should decentralize aid programming and decision making to the lowest possible organizational level, where officers have the greatest knowledge of what is happening on the ground" (p. 72). Natsios also notes that staff are often frustrated by the lack of autonomy and by the "risk aversion" that results from this environment (pp. 57-58).

Variation in the Structure of Aid Organizations

In thinking about the role that organizational structure might play in explaining the impact of foreign aid delivery, we can take advantage of differences in the way that aid delivery agencies are structured. By looking across aid organizations and seeing if variation in structure is associated with variation in performance, we can start to gather suggestive evidence as to whether organizational structure and management does indeed have an impact on performance. Table 1 below compares two aid organizations, USAID and DFID.

Table 1: Comparison of USAID and DFID's Political Authorizing Environment

	Political status of aid agency head	*Budget security*	*Response to 2008 financial crisis*	*Workplace satisfaction surveys*	*Rank (out of 33) on autonomy measure used in econometric work below*
DFID	Full ministerial rank, limited coordination with Foreign Affairs	Three-year budget allocations; few earmarks	Only ministry spared from across-the-board cuts; budget has continued to increase	Top 2%	3
USAID	Head of USAID (Administrator) reports to State Department	Yearly, often delayed; USAID budget heavily earmarked	Cutting aid-funding promises *literally* the first thing mentioned by Obama ticket (as candidate)	Bottom third	29

Sources: 2012 US Federal Employee Viewpoint Survey Global Satisfaction Index (USAID 25th of 36); 2013 UK Civil Service People Survey Employee Engagement Index (DFID tied for 2nd of 98); Biden-Palin Debate, October 2 2008; author.

Do we see, then, differences in performance between these organizations? The short answer is a definitive *yes*. A few illustrative examples:

In 2006, Liberia was just emerging from two decades of conflict. A strong Minister of Health was looking for international help in improving Liberia's woeful health statistics, which were among the world's worst.[4] Faced with a ministry that had not produced a financial statement in over a decade, and having no idea where funds allocated to the ministry were flowing, the Minister approached the US Agency for International Development (USAID) about establishing an office of financial management. USAID declined. The Minister then approached the UK's Department for International Development (DFID), which was excited by the idea and quickly began to implement it.[5] At a point when it was still too early to measure the new office's performance and generate quantitative data, DFID staff on the ground realized that their mission was not succeeding. They used their judgment that the wrong personnel had been assigned and arranged to have them replaced. Today, the Liberian health ministry's office of financial management is thriving, praised for its professionalism and effectiveness.

In the same country, in the same ministry, both DFID and USAID wished to support the same reform-minded Minister by putting the ministry in greater control of external funding. DFID set in motion the development of a pooled fund – a collective funding mechanism with contributions from multiple donors and a governing board composed of donor and health ministry representatives. While at least some of the critical USAID decision makers would have liked to contribute to the fund, Congressional restrictions prevented USAID from comingling its funds in this way; USAID ultimately set up a parallel system with much higher transaction costs and predetermined performance targets which, due to Liberia's inherent unpredictability, require frequent and costly revision.

In South Africa in the mid-2000s, both USAID and DFID wished to strengthen municipal governments. DFID's primary mode of engagement was to embed flexible advisers in municipal governments and let them guide support over the long term. USAID considered a similar strategy but initially rejected it, in part because it would be difficult to develop consistent measures for these activities. USAID instead initially worked primarily via the delivery of trainings, an approach for which the outputs (such as the number of participants and trainings) could be more easily measured.

This is not to suggest that measurement is always the inferior strategy. Also in South Africa in the mid-2000's, the US government's PEPFAR response via USAID and CDC, with its focus on targets and delivery, was – while not without shortcomings, particularly regarding exit

---

[4] These accounts come from in-person interviews with individuals who were part of these interventions, and receive fuller treatment in Honig 2015.

[5] Later, following a conversation with the US Ambassador and his intervention, USAID did indeed offer to provide support to establish the unit, though on a much slower timeline than that of DFID.

7   *Next Horizons Essay Competition 2014-15 – Dan Honig*
"More autonomy for donor organizations and their agents (sometimes): bringing
organizational behavior and management theory to foreign aid delivery"

strategy and sustainability – quite effective in responding to the AIDS epidemic, much more so than DFID's focus on high-level government strategy.

Quantitative Data and Results

However compelling these accounts may seem – and however useful they may be in understanding the mechanisms of action at play – they cannot aspire to universality, to general claims of relevance to a wide variety of aid agencies. After all, while DFID and USAID may face different political authorizing environments, these are not the only differences between them. Additionally, the interventions described above may have been outliers, idiosyncratic for one reason or another.

Thankfully, we may draw upon a much more general data source in this regard; I examine differential returns to autonomy in a dataset that I compiled of over 14,000 unique interventions in 178 counties carried out by nine donor agencies over the past 40-plus years.[6] Figure 1 below gives an overview of the distribution of projects by country.
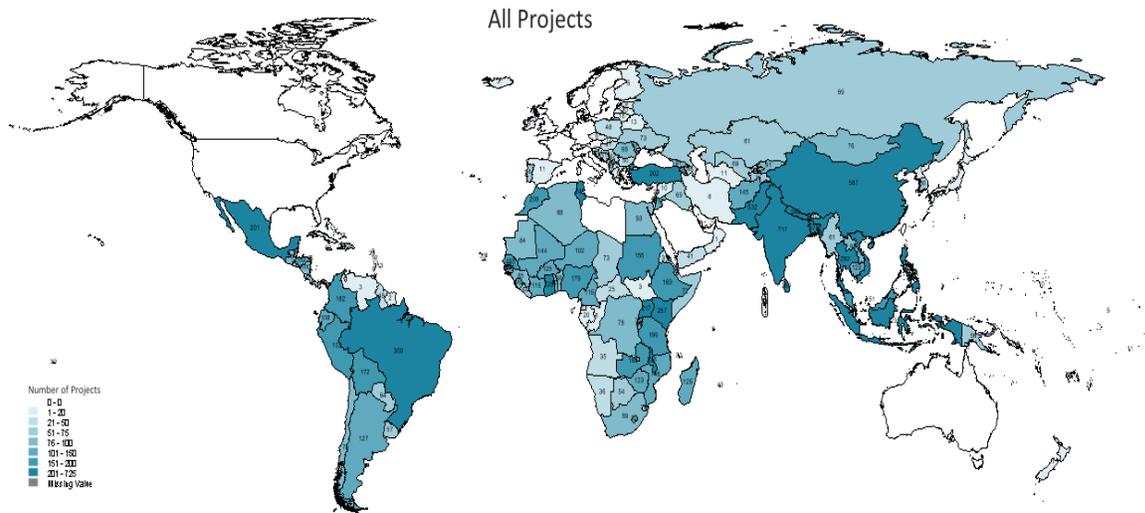

Figure   1: Overview of Projects in Dataset

The key dependent variable in the analysis is overall intervention success, a holistic rating undertaken by independent evaluators (either external evaluation contractors or independent evaluation units) or by project staff in project completion reports. For most IDOs, project

---

8   *Next Horizons Essay Competition 2014-15 – Dan Honig*
"More autonomy for donor organizations and their agents (sometimes): bringing
organizational behavior and management theory to foreign aid delivery"

success is an ordinal variable ranging from 1 to 6, with 6 being "Highly Satisfactory" and 1 being "Highly Unsatisfactory."[7]

It would be ideal to have time-varying data on organizational autonomy for every organization, including variation at the country (or even intervention) level. The data available only varies at the organization level and is time-invariant.[8] This work, therefore, cannot test for the effect of autonomy on success directly, as different organizations have different measurement standards; a rating of 4 given by aid organization A may or may not mean an intervention is more successful than one that received a rating of 3 from aid organization B. This work can, however, examine the *differential* performance of organizations with varying levels of autonomy in interaction with other explanatory variables, thus leveraging the idea that a rating of 4 given by organization A means a project succeeded more than a project assigned a 3 by organization A.

Figure 2 below (drawn from Honig 2015) demonstrates the main findings. Using the State Fragility Index as a measure of environmental unpredictability and an autonomy measure drawn from the Paris Declaration monitoring surveys, an organization with a greater level of autonomy is predicted to have a much more consistent performance across countries of varying fragility than an organization with a lower level of autonomy.
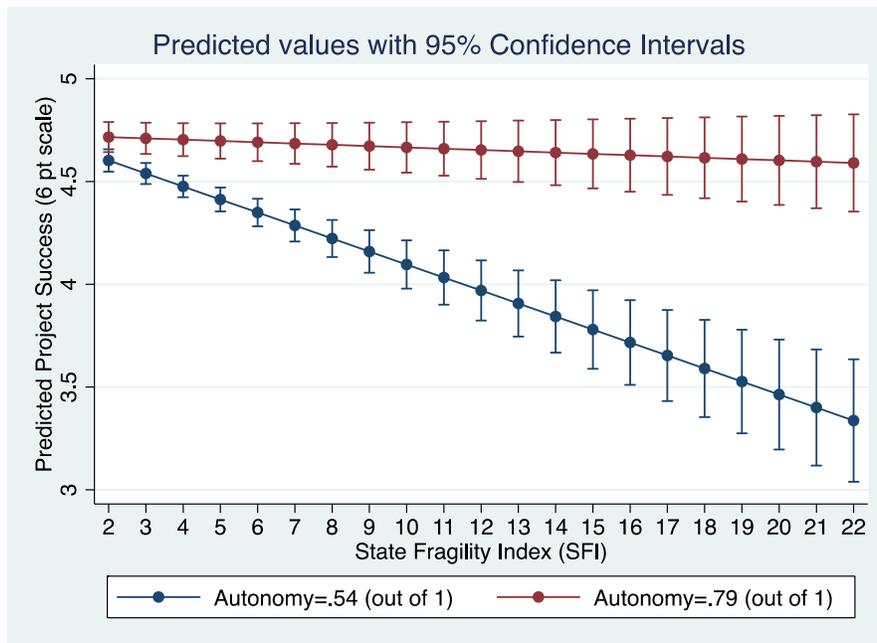

Figure 2: Returns to Autonomy in Countries of Differential Environmental Unpredictability

---

[7] These are the World Bank's designations. No agency has significantly different names/standards in this regard, which would, in any case, be removed by agency fixed effects.

[8] This study's focus on measurement at the organizational level is not intended to suggest that there is not recipient and recipient-year variation in autonomy, only that this is the level at which measurement is most clean and broad. Controls below ensure that my results are not biased by these other levels of variation in autonomy.

These results have been subject to a wide range of tests of robustness. They are, for example, robust to use of a variety of fixed effects (including time and recipient country fixed effects), which should allay any concerns that the results are driven by heterogeneous agency intervention performance over time or by heterogeneous entry of agencies into and out of recipient countries over time. These results are also robust to the use of an alternative direct survey measure of autonomy and to controlling for evaluation type, thus assuaging concerns that differential reliance on project staff for evaluations might drive differential bias between aid organizations.[9] This autonomy measure is not merely a proxy for good donor practices in general; placebo tests demonstrate that the autonomy scale has a unique relationship with intervention outcomes, one that parallels the theory above.

We also see differences by OECD-DAC Creditor Reporting Service (CRS) purpose codes, which help us differentiate between different kinds of tasks. Also drawn from Honig 2015, the tables below focus, on the one hand, on purpose codes which relate to infrastructure construction or observable service delivery (for which we might not expect as strong a relationship between autonomy and outcome) and, on the other hand, on purpose codes which focus on related policy or administration tasks but which are more difficult to observe. Focusing on related but difficult-to-observe domains helps to ensure that the results are not driven by something like the fact that it is much easier to deliver electricity than to deliver education.

There is no relationship between autonomy in interaction with state fragility and intervention success in the first set of task domains, where the focus is on constructing something or delivering a service that is tangible and relatively easy to monitor, but the relationship is relatively strong in related administrative sectors. These results are consistent with my contention that the kind of task mediates the relationship between intervention success and environmental unpredictability.

---

[9] This is but a small number of the robustness tests in Honig 2015, which also include: Using ordered logit models on six point project outcome scales (rather than OLS); using z-scores as outcomes (rather than the six-point scale where employed); compressing success and failure to a binary outcome and employing logit models; employing quantile regression (were results to be driven by only part of the distribution); restricting the state fragility index to common support; that is, only the range of SFI realized in all donors' data (2-22, rather than 0-25 in the main analysis); dropping the latter two waves of the Paris Declaration survey in generating the autonomy measure (to allay concerns that donors responded to measurement by changing their practices) ; dropping any individual organization from the sample; double-clustering standard errors at the organization-recipient level (rather than clustering on organization alone); dropping either subscale of state fragility (legitimacy or effectiveness); and using any of the four domains of state fragility (security, political, economic, or social), among others.

Table 2: Relationship between Autonomy and State Fragility by Sector (Outcomes Easily Contracted; Sector by CRS Code)

DV: Project Success (6-pt scale)	(1) Road Infrastructure & Transport	(2) Buliding Power Transmission Lines	(3) Agricultural Irrigation & Water	(4) Basic Drinking Water Supply & Sanitation
State Fragility Index (SFI)	-0.262	0.586*	-0.516	-0.298
	(0.352)	(0.128)	(0.343)	(0.152)
Autonomy*SFI	0.356	-0.958*	0.735	0.386
	(0.561)	(0.201)	(0.536)	(0.233)
Constant	5.010***	5.120***	4.588***	4.621***
	(0.161)	(0.0796)	(0.152)	(0.0486)
IDO Fixed Effects	Y	Y	Y	Y
R^2-Within	0.030	0.031	0.024	0.054
R^2-Between	0.018	0.263	0.153	0.000
Observations	469	167	165	271

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Relationship between Autonomy and State Fragility by Sector (Outcomes Difficult to Contract; Sector by CRS Code)

DV: Project Success (6-pt scale)	(1) Transportation Management	(2) Agricultural Policy & Administration	(3) Social/Welfare Services (Administration, Capacity Building)	(4) All Administration/ Policy Management
State Fragility Index (SFI)	-1.030***	-0.670***	-0.371***	-0.151***
	(0.0271)	(0.123)	(0.0178)	(0.0125)
Autonomy*SFI	1.716***	0.928**	0.561***	0.192***
	(0.0407)	(0.182)	(0.0305)	(0.0195)
Constant	2.978***	4.587***	4.508***	4.554***
	(0.0266)	(0.246)	(0.0288)	(0.0210)
IDO Fixed Effects	Y	Y	Y	Y
R^2-Within	0.234	0.077	0.025	0.019
R^2-Between	0.058	0.437	0.031	0.296
Observations	39	55	160	1530

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This provides further evidence that a consideration of the effects of measurement is critical in determining where measurement is likely to have a negative effect on intervention success – that is, in harder-to-observe task domains – and where its effects are likely to be more ambiguous and potentially beneficial. Soft information seems to matter to development success, with more autonomous agencies thus better able to manage more unpredictable

contexts and task domains less tractable to navigation by measurement. This suggests that autonomy can have positive effects inasmuch as it provides support for the acquisition and use of soft information.

Implications

While the argument presented here relates specifically to organizational autonomy and measurement, I do not mean to suggest that this is the only dimension upon which managerial practices and organizational structure affect intervention performance, nor that these practices are the only source of variation in intervention performance or development impact. I mean for these data to simply demonstrate that managerial practices *do* matter – that they are one of the things that contributes to foreign aid intervention performance. These results suggest that this effect is substantively significant; the data underlying Figure 2 above suggest that, in comparing recipient-country environments one standard deviation above and below the mean, a relatively high-autonomy development organization would see a difference of about .05 points in performance on a six-point scale, while a relatively low-autonomy development organization would see more than 10 times the difference.

Donors like DFID and the World Bank are just the first among a complicated web of organizations, of course; contractors, implementers, NGOs, and CSOs are all organizations that have varying degrees of autonomy, play a critical role in development success, and face their own complex web of incentives that include reasons to 'show results' and engage in legitimacy-seeking behavior. We can do more to understand these relationships and their net results: where the complex web is optimally oriented towards development impact, and where it can be improved.

Furthermore, this large-N analysis can capture only the tip of the management iceberg. Denizer, Kaufmann, & Kraay (2013) find that Task Team Leaders are critical to explaining World Bank project success; the bottom line is that people matter in development delivery. Organizational structure and incentives are surely connected to employee entry and exit: who decides to join an aid agency, and who stays in that agency.

We also need to do more to think about the idiosyncratic features of individual sectors, or individual sector-country combinations. The Gambia and Uzbekistan have the same rating on the 2012 State Fragility Index, but this does not mean that the DFID Water & Sanitation team visiting Banjul ought be structured, measured, monitored, and incentivized identically to the DFID Water & Sanitation team headed to Tashkent.

Development has seen a move towards measurement in the past few decades, much of it linked to New Public Management and a drive towards what are sometimes framed as private sector solutions and efficiencies. (Gulrajani 2011) Differences in the nature of tasks, the availability of summary performance statistics (e.g. profit), and the range of possible

12   *Next Horizons Essay Competition 2014-15 – Dan Honig*
"More autonomy for donor organizations and their agents (sometimes): bringing
organizational behavior and management theory to foreign aid delivery"

organizational setups has led a number of famous economists and political scientists to conclude that private sector solutions sometimes appropriate in the public sector, and sometimes they are not (Dewatripont, Jewitt, and Tirole 1999; Dixit 1997, 2002; Wilson 1989). Where private sector techniques make sense, the world of international development would do well to refer to the literatures on private sector management and organizational behavior and note that these literatures demonstrates that 'private sector solutions' are not limited to, or primarily focused on, measuring performance as a means to achieve results.

Recommendations

Aid agencies would do well to exploit organizational design, the "low-hanging fruit" of international development - the factor in development outcomes arguably most changeable by Western governments and polities. By the estimate of one interviewee with long experience at the United Nations Development Programme (UNDP), approximately 30% of all staff time is spent on measurement design and reporting (Interviews). For fiscal year 2013, this works out to approximately $350 million;[10] if a move towards more navigation by judgment and less navigation by output measurement were to reduce this figure by even 25%, the administrative savings – not to mention the efficiency gains from greater impact of UNDP's nearly nine billion dollars in annual development spending – would be quite significant.

Where output measurement and tight control by distant principals work well, management by measurement should be used to better deliver vaccines or more efficiently build electricity transmission infrastructure. But many of the environments in which organizations most seek profit or impact could be described as unfamiliar, unpredictable, or both; the effect of output measurement and tight control in these contexts may not be positive.

In the contexts where aid has the potential to make the most difference – in the most fragile states – measurement is the least useful; rather, navigation by judgment is the optimal strategy. My findings suggest that not only are we not doing all we can to improve aid delivery, the move towards measurement and control across all aid sectors in recent years may actually be making things worse in some sectors. While measurement may lead to the construction of many successful dams, it may also leave recipient countries without the capacity building necessary to manage and maintain those dams or to put the electricity to use. If our drive for results leads us to control aid too tightly, we may end up accomplishing precisely the opposite of what we intend.

We need to begin this project today, directing our funding and attention toward management challenges and incentive problems. When this focus is coupled with aid organizations willing to experiment with altered practices we will be able to confirm that these altered practices work once the proverbial rubber hits the road and ought be used more broadly. Optimal design will

---

[10] This is drawn from UNDP's estimates of administrative and policy coordination cost (United Nations Development Programme 2013, p. 6).

not ensure that foreign aid is universally successful, but it will ensure that those features wholly under the control of donor countries are calibrated so as to give aid the best chance to realize maximum impact.

The evidence suggests the pendulum has swung too far in recent decades towards measurement driven management. The solution I propose is not to return to a time where results were not considered; it is, rather, to keep results in mind as we design our organizations to achieve maximum impact. I advocate that aid organizations take a more nuanced approach to measurement and control regimes, namely:

- Considering the role of soft information and thus agent autonomy in a given task and environment;
- Specifying clear outcome measures where this is practicable;
- Reducing measurement where accurate, reliable, and holistic measures cannot be fully specified ex-ante in order to crowd in autonomy; and
- Carefully considering the way management, promotion, staffing, and incentive-based pay policies interact in determining whether organizations' structures are appropriate to achieve maximum impact.

Next Horizons Essay Competition 2014-15 – Dan Honig
"More autonomy for donor organizations and their agents (sometimes): bringing
organizational behavior and management theory to foreign aid delivery"

References

Andrews, Matt. 2011. "Which Organizational Attributes Are Amenable to External Reform? An Empirical Study of African Public Financial Management." *International Public Management Journal* 14(2): 131–56.

Andrews, Matt, Lant Pritchett, and Michael Woolcock. 2012. *Escaping Capability Traps Through Problem Driven Iterative Adaptation (PDIA)*.

Barder, Owen. 2009. *Beyond Planning: Markets and Networks for Better Aid*.

Barnett, Michael. 2009. "Evolution Without Progress? Humanitarianism in a World of Hurt." *International Organization* 63(04): 621.

Birdsall, Nancy and William D. Savedoff. 2010. Center for Global Development.

Booth, David. 2013. Governance and Politics *Facilitating Development: An Arm's Length Approach to Aid*.

Brown, Trevor L., and Matthew Potoski. 2003. "Managing Contract Performance: A Transaction Costs Approach." *Journal of Policy Analysis and Management* 22(2): 275–97.

———. 2005. "Transaction Costs and Contracting: The Practitioner Perspective." *Public Performance & Management Review* 28(3): 326–51.

Bush, Sarah. 2011. *The Taming of Democracy Assistance*.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton University Press.

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2013. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." *Journal of Development Economics* 105: 288–302.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." *Review of Economic Studies* 66(1): 199–217.

———. 2000. "Multitask Agency Problems: Focus and Task Clustering." *European Economic Review* 44(4-6): 869–77.

DiMaggio, PJ, and W Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American sociological review* 48(2): 147–60.

Dixit, A. 1997. "Power of Incentives in Private versus Public Organizations." *The American Economic Review* 87(2).

———. 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review." *Journal of human resources* 37(4): 696–727.

Easterly, William. 2002. "The Cartel of Good Intentions: The Problem of Bureaucracy in Foreign Aid." *The Journal of Policy Reform* 5(4): 223–50.

———. 2014. *The Tyranny of Experts: How the Fight Against Global Poverty Suppressed Individual Rights (Google eBook)*. Perseus Books Group.

Eyben, Rosalind. 2013. … -a-framing-paper-for-development-practitioners (… *Uncovering the Politics of "Evidence" and "Results". A Framing Paper for Development Practitioners*.

Gates, Bill. 2013. *2013 Gates Foundation Annual Letter*.

Gulrajani, Nilima. 2011. "Transcending the Great Foreign Aid Debate: Managerialism, Radicalism and the Search for Aid Effectiveness." *Third World Quarterly* 32(2): 199–216.

Holmstrom, B., and P. Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7: 24.

Honig, D. 2015. Navigating by Judgment: Organizational Structure, Autonomy, and Country Context in Delivering Foreign Aid. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Huber, John D., and Charles R. Shipan. 2002. *Deliberate Discretion: The Institutional Foundations of Bureaucratic Autonomy*. Cambridge University Press.

Lawrence, Paul R., and Jay William Lorsch. 1967. *Organization and Environment: Managing Differentiation and Integration*. Harvard Business School Press.

March, James G., and Herbert Alexander Simon. 1958. *Organizations*. Wiley.

McMahon, E. R. 2001. "Assessing USAID's Assistance for Democratic Development: Is It Quantity versus Quality?" *Evaluation* 7(4): 453–67.

Meyer, JW, and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American journal of sociology* 83(2): 340–63.

Natsios, A. 2010. *The Clash of the Counter-Bureaucracy and Development*. Washington, DC.

Ostrom, Elinor, Clark Gibson, and Sujai Shivakumar. 2002. *Aid, Incentives , and Sustainability: An Institutional Analysis of Development Cooperation*.

Perrow, Charles. 1967. "A FRAMEWORK FOR THE COMPARATIVE Analysis of Organizations." *American sociological review* 32(2): 194–208.

Polanyi, Michael. 1966. *The Tacit Dimension*. Doubleday.

Pritchett, Lant, and Michael Woolcock. 2004. "Solutions When the Solution Is the Problem: Arraying the Disarray in Development." *World Development* 32(2): 191–212.

Ramalingam, Ben. 2013. *Aid on the Edge of Chaos: Rethinking International Cooperation in a Complex World*. Oxford University Press.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

Stein, JC. 2002. "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms." *The Journal of Finance* LVII(5).

Tendler, Judith. 1975. *Inside Foreign Aid*. Johns Hopkins University Press.

Thompson, James D. 1967. *Organizations in Action: Social Science Bases of Administrative Theory*. Transaction Publishers.

United Nations Development Programme. 2013. 39319 *UNDP Integrated Budget Estimates for 2014-2017 (DP/2013/41)*.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do And Why They Do It*. Basic Books.