

Conversation with Angela Hawken, September 16, 2013

Participants

- Angela Hawken – Associate Professor of Public Policy, Pepperdine University
- Holden Karnofsky – Co-Founder, Co-Executive Director, GiveWell
- Howie Lempel – Research Analyst, GiveWell

Summary

GiveWell spoke to Angela Hawken as part of its investigation of criminal-justice reform. The conversation covered Hawken's proposed organization, BetaGov, which will enable practitioners to conduct randomized controlled trials (RCTs) of public-sector programs, initially focusing on criminal justice. Hawken also commented on Hawaii's Opportunity Probation with Enforcement (HOPE) program and described her work conducting RCTs of HOPE in Arizona and Washington.

Note: This set of notes was compiled by GiveWell and gives an overview of the major points made by Angela Hawken.

Problem: Lack of rigorous testing in public sector

In the private sector, thousands of randomized controlled trials (RCTs) are conducted every day, but in the public sector rigorous testing is rare. For example, within criminal justice, about five RCTs are conducted every year. Ideally, legislatures would require government programs to be rigorously tested and would consider programs to be "evidence based" only after multiple RCTs found positive results, but this is not currently the case. What shows promise in one jurisdiction might not work in another. We should be careful about drawing strong conclusions about "what works" based on single studies.

Many issues with the current system lead to this lack of testing. Public-sector practitioners do not feel equipped or encouraged to test innovative ideas and approaches because the tools for such studies are not readily available and there is a fear of exposing programs as ineffective. Nearly all RCTs of public-sector programs are conducted through academic institutions or think tanks and are time consuming and typically very expensive. When practitioners submit proposals for RCTs to be conducted by an outside institution, it may take two years for proposals to be funded, at which point the innovating practitioner may no longer even work at the agency.

BetaGov

BetaGov aims to transform knowledge generation in the public sector and increase the evidence available for public-sector programs by enabling practitioners to conduct RCTs of their own programs. Much of the data required for assessing key outcomes of RCTs is already being collected by government agencies for administrative purposes, so RCTs can be implemented immediately and inexpensively.

BetaGov will serve three main functions:

- **Repository for ideas.** BetaGov will serve as a repository for ideas to be tested; this will attract ideas that would otherwise not have risen to the surface.
- **Toolkits for conducting RCTs.** BetaGov will provide toolkits that government and nonprofit practitioners can use to run RCTs within their own agencies, rather than relying exclusively on universities or think tanks. This will reduce the time and financial cost for practitioners to test ideas.
- **Learning across studies.** Practitioners will be required to submit their data and report results to BetaGov, which will enable it to make comparisons and share sector-wide knowledge.

Role and potential impact

- **Increasing the evidence available.** By collecting ideas, enabling and encouraging practitioners to conduct RCTs, and sharing the results, BetaGov will dramatically increase the evidence available for public-sector programs. Within criminal justice, it aims to bring the number of RCTs started every year from about five to several hundred, when BetaGov is operating at scale. RCTs can test new or existing approaches, and could focus on small parts of a process or entire programs. BetaGov staff will not lead the research, but may monitor aspects of some RCTs, such as checking that the process described matches what was implemented. They will also be available to provide ongoing advice and support during the period of the trial. Any one of these RCTs might not be perfect (field experiments rarely are), but producing a large number of “good enough” RCTs will still be a massive contribution to the field. All RCTs run using BetaGov resources will be required to share their results.
- **Sharing failures as well as successes.** It is well established that practitioners are more inclined to publish findings of programs that work than those that do not. This discourages rigor and transparency. Learning from failure is as important as learning from success. BetaGov will encourage practitioners to broadcast “failures” (ideas found to have negligible or negative results), because this information is of great value to the field—if we know that something *doesn't* work,

we should stop doing it. It will award prizes for well-conducted RCTs, regardless of whether they found positive or negative results.

- **Helping promising ideas to get recognized.** There is not currently a good system for matching programs with evaluators. Conducting practitioner-led RCTs is a great way to identify promising ideas. After this stage, practitioners can continue to test and implement the programs on their own or an outside institution may conduct further evaluations.

Areas of focus

BetaGov's first area of focus will be criminal justice in developed countries because Hawken's team has experience and strong networks in this field. Another benefit of working on criminal justice is that the agencies are fairly hierarchical, so if a higher-level official is interested in running a trial, it will happen. The launch of BetaGov comes at an auspicious time. There appears to be bipartisan interest in finding cost-effective solutions in criminal justice that would reduce incarceration without compromising public safety. Decisionmakers are therefore willing to consider ideas they might not have before.

Early on, BetaGov will also do some exploratory work in less-developed countries, though this will not be a primary focus. BetaGov eventually expects to work both in the United States and abroad and to broaden its focus to include any number of public-sector issues.

Next steps

In order to launch, BetaGov needs to develop a website and create the toolkits for conducting RCTs. BetaGov plans to form a researcher-advisory group to ensure that RCTs meet research standards and that the tools are user-friendly for practitioners. Similarly, a practitioner-advisory group will ensure that the products developed are relevant for the end user.

Funding

Setting up BetaGov as a nonprofit funded by foundations or private donors will allow it to be nimble and get off the ground quickly.

Hawken is in contact with some philanthropists whose interests align well with BetaGov, though no funding has been secured. On the two occasions when Hawken has spoken publicly about the idea for BetaGov, it has generated great excitement and she feels confident that enough funding could be secured from private donors to get it off the ground. More funding would enable BetaGov to support more RCTs.

[In December 2013, after this conversation took place, Good Ventures awarded a grant of \$200,000 to Pepperdine University to support BetaGov.]

Promising ideas for criminal justice reform

There are many promising ideas for criminal-justice reform that should be tested further.

- **Reducing waiting time.** People are held in custody, often for weeks or months, while awaiting a hearing or trial. This represents a fiscal cost to the state and a time cost to the person in custody. Furthermore, punishments given closer to the time of the offense are found to be stronger deterrents.
- **Reducing time in custody.** There are data indicating that the first few days in jail are the most uncomfortable for inmates, so there may be “diminishing returns” to time in custody. Some judges are already giving shorter sanctions in response to technical violations because they deter about as effectively as longer ones. There would be high value (in fiscal and humanitarian terms) to reducing time in custody when public safety is not at issue.
- **Alternative punishments.** For many people on community supervision (people who cycle through the criminal-justice system), standard sanctions are not always the most effective or efficient deterrent to violate. There are promising alternative punishments that need to be explored further, such as having to write a reflective essay or serve on a work crew, that better reflect human motivations.

Hawaii’s Opportunity Probation with Enforcement (HOPE)

Hawken was involved with the first randomized controlled trial of Hawaii’s Opportunity Probation with Enforcement (HOPE) program. HOPE has received attention for being a “victory” and an “evidence-based” program, but there is a danger in labeling it that way. The trial certainly indicated that HOPE was promising, but it did not answer the question of *which* elements of the program were key for its success. There is still a lot to be learned from further testing. While HOPE’s premise makes sense and is based on knowledge of human behavior, HOPE is challenging to implement and there is still a great deal to learn about how to improve implementation. Simply handing states a handbook on how to roll out the HOPE program is not going to work. These programs need to be tailored to work well within the adopting jurisdiction’s practices, so long as the key features of the program (swift, certain, and fair) are maintained.

One of the criticisms of HOPE is that the effects may not be sustained in the long-term. The extent to which effects are sustained will likely depend on the length of a jurisdiction’s probation and parole periods, which vary roughly from 6 months to 5 years.

HOPE is designed to reinforce that bad decisions lead to bad consequences, good decisions lead to good consequences, and how one reacts to bad consequences is important because consequences compound. Hawken expects that there will be some regression after the period of the intervention, but that substantial effects will be sustained.

HOPE trials in other states

A rapidly launched trial in a western state

A jurisdiction in a western state wanted to run a trial of the HOPE program (with random drug testing once per month instead of six times per month), but didn't have additional funds for the research. Rather than apply for funding from the National Institute of Justice and delay the trial, Hawken's team helped the jurisdiction implement the trial immediately using their own resources and data it was already collecting. The trial was completed in short order and did not replicate HOPE's success. The trial did not cost the agency any additional funding.

Hawken's team is helping this same state with another RCT as well.

Washington

Hawken's team is currently doing a substantial amount of work with Washington State. Washington is innovative and eager to test its programs, and Hawken's team is working closely alongside practitioners there to launch RCTs of variations of HOPE. For example, varying the frequency of drug testing and the length of sanctions. The goal of the Washington RCTs is to find the least-intrusive testing schedule and minimum sanctions that still achieve the desired behavior change.

Washington rolled out its version of HOPE across the entire state in a few months; more than ten thousand offenders were enrolled in the program. Hawken noted that this is the largest transformation in community corrections that she's seen in her career. She initially predicted the statewide rollout would fail (she was of the opinion that the large-scale effort would overwhelm community corrections and law enforcement). Having observed the experience in Washington closely, she now believes that the expansion has gone fairly well. She attributes this to good leadership from DOC management and an inspired implementation team who helped coordinate the participation of the thousands of people who are involved in implementing the program.

All GiveWell conversations are available at <http://www.givewell.org/conversations>