Twitter Thread by Sahil Bloom

Sahil Bloom
@SahilBloom

Incentives are everything.

Here's a framework for establishing incentives (that actually create desired outcomes):

Great incentives create great outcomes.

But unfortunately, humans are astonishingly bad at establishing incentives—we consistently invite manipulation and unintended consequences.

This thread shares my framework for a establishing better incentives:

Incentives are anything that drives an individual to act in a specific manner.

Intrinsic incentives are internal—created by self-interest or desire.

Extrinsic incentives are external—created by outside factors (reward, punishment).

Today, we'll focus on extrinsic incentives.

In a very simple model, extrinsic incentives typically involve two key components: the measure and the target.

The measure is the metric that the individual or group will be judged upon.

The target is the level of the measure at which a reward (or punishment) will be initiated.

The problem?

This simple model of incentives—which will feel familiar to many—often leads to undesirable outcomes and unintended consequences.

Goodhart's Law explains why:

When a measure becomes a target, it ceases to be a good measure.

If a measure of performance becomes a stated goal, humans optimize for it, regardless of any consequences.

The measure is rendered useless.

The phenomenon has been observed time and again.

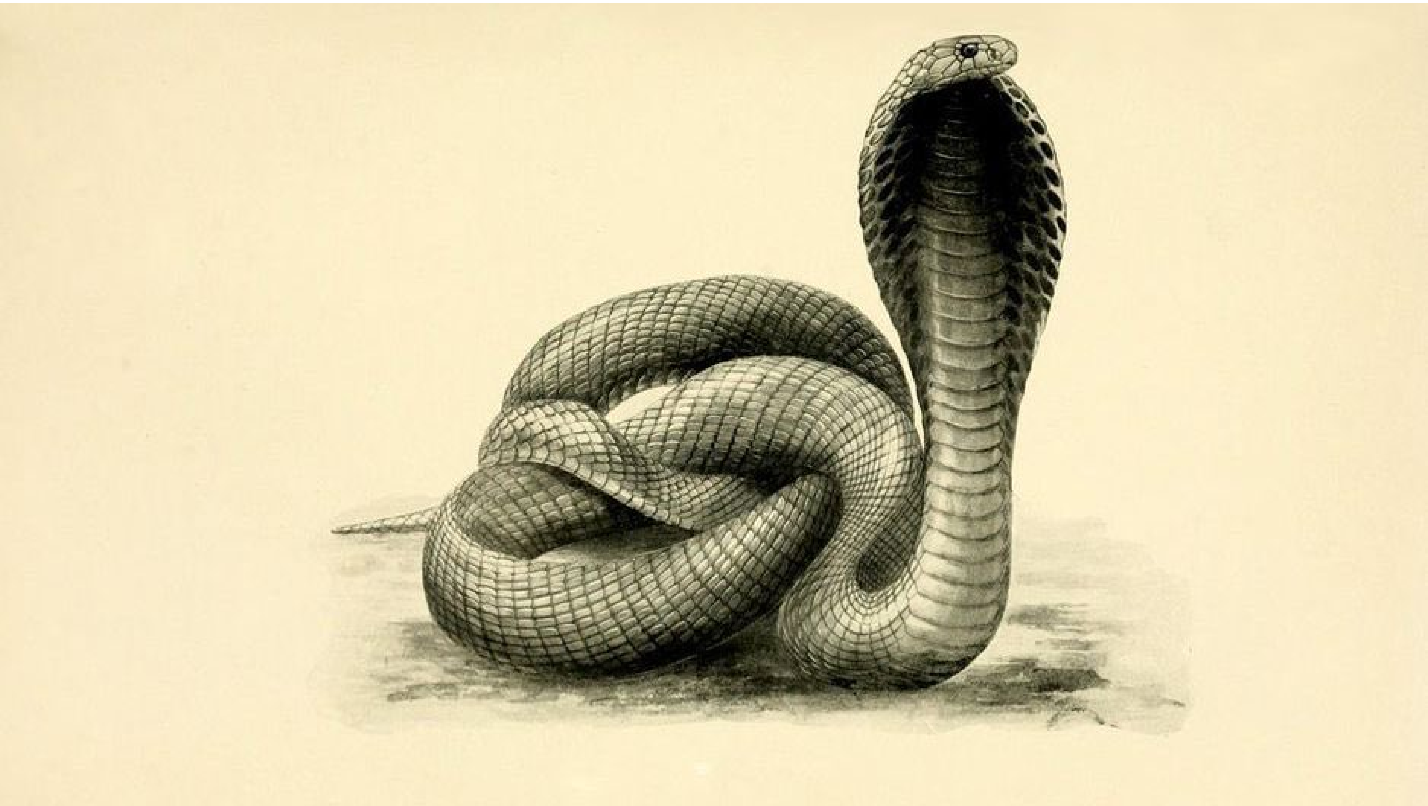
Let's look at a few examples and build a mental model for where incentives go awry:

The Cobra Effect

There were too many cobras in India, so the British colonists started offering bounties for cobra heads.

Locals began breeding cobras, killing them, and turning in the heads to earn bounties.

Many cobras were released, increasing the population of cobras.



The British viewed cobra heads as a simple way to measure cobra elimination, so it gave the population an incentive to deliver cobra heads.

The result? Locals bred cobras to kill them and earn bounties.

An incentive designed to reduce the cobra population actually increased it!

Amazon's "Hire-to-Fire" Issue

Amazon believed employee turnover was healthy, so it gave its managers a target rate for annual turnover.

The result? Managers hired employees they planned to fire in order to meet their targets.

Clearly not what Jeff Bezos had in mind...

Wells Fargo Account Openings

Wells Fargo viewed new account openings as an easy way to track business growth, so it gave its employees target account opening goals.

The result? Employees opened millions of fake accounts to hit their targets.

Wells Fargo was fined billions.

With these examples as a backdrop, we can begin to formulate a rough mental model for where incentives go awry.

Poorly designed incentives typically share one or more of the following three characteristics:

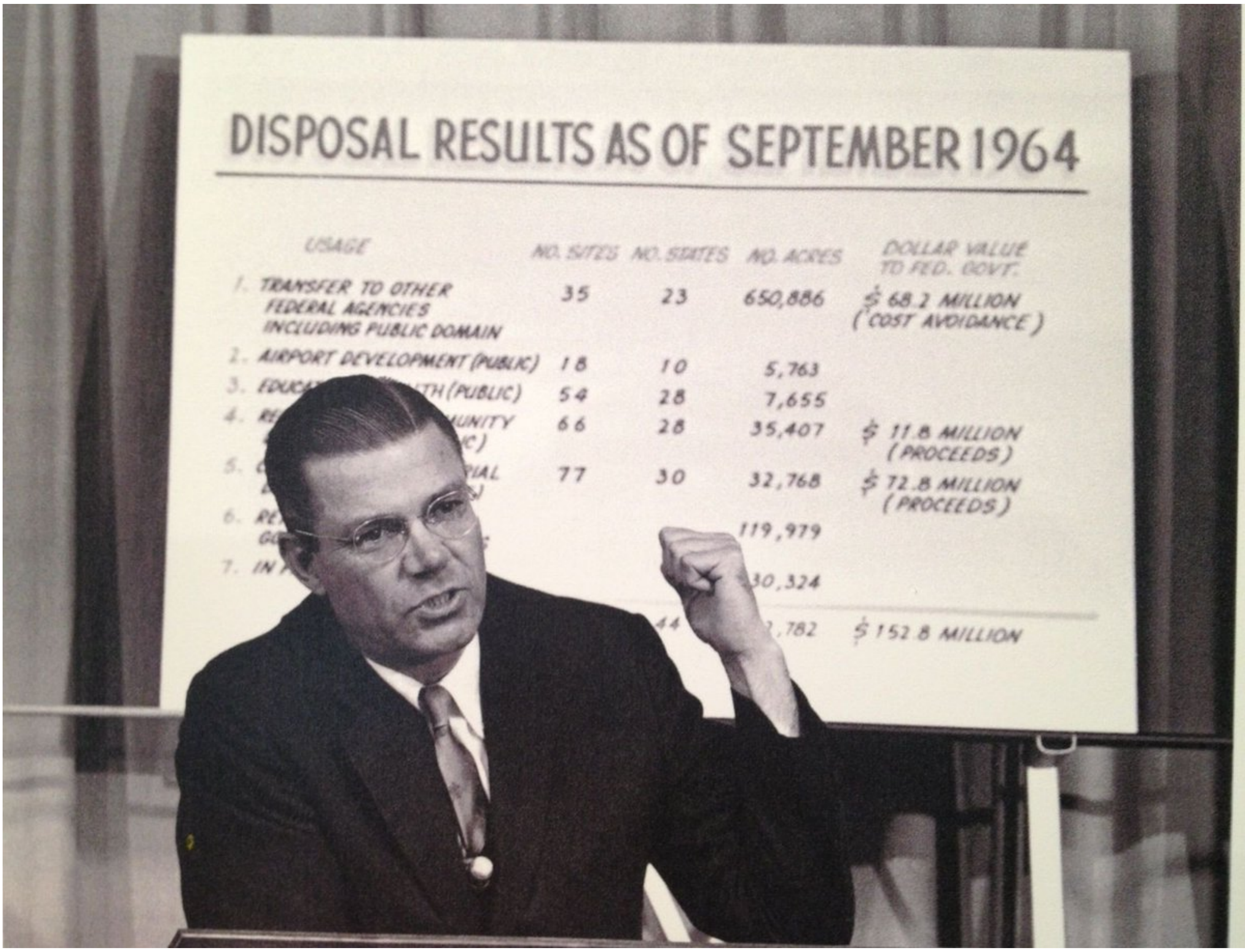
(1) McNamara Fallacy
(2) Narrow Focus
(3) Vanity > Quality

The McNamara Fallacy

Named after Robert McNamara—the US Secretary of Defense from 1961-1968.

His over-reliance on quantitative metrics led the US astray during the Vietnam War.

The McNamara Fallacy is the flawed assumption that what can't be measured isn't important.



It leads to a focus on measuring what is easy to measure vs. what is actually important.

Cobra heads, employee turnover, and new account openings were all easy to measure quantitatively, but missed the bigger picture.

All three programs were victims of the McNamara Fallacy.

Narrow Focus: Thinking too narrowly about the desired outcomes of the program.

Vanity > Quality: The reliance on vanity metrics (cobra heads, new account openings) that will impress superiors or the public.

So with these in mind, let's create a better framework for incentives:

The incentive framework involves 5 key pillars:

(1) Objectives
(2) Metrics & Anti-Metrics
(3) Stakes & Effects
(4) Skin in the Game
(5) Clarity & Fluidity

It provides a structure through which to create, evaluate, and adjust incentives.

Let's walk through the pillars…

Objectives

Deep consideration of the ultimate objectives of the incentives are a critical first pillar.

What does success look like? Not the surface level objectives—go deeper.

Without upfront deep thought on objectives, intelligent incentive design is impossible.

Metrics

Establish metrics that you will measure to track success.

Avoid the McNamara Fallacy—never choose metrics on the basis of what is easily measurable over what is meaningful.

Identify a wish list of metrics with no regard for feasibility. Work backwards from there.

Anti-Metrics

Even more important than the core metrics, establish "anti-metrics" that you measure to track unintended consequences.

Anti-metrics force you to consider whether your incentives are fixing one problem here, but creating another problem over there.

(h/t @joulee)

Stakes & Effects

Always consider the stakes:

High = costly failure, hard to reverse
Low = cheap failure, easy to reverse

If high-stakes, conduct a rigorous, second-order effects analysis (tip: read my thread on this).

Iterate on metrics accordingly.

Skin in the Game

To avoid principal-agent problems, the incentive designer should have skin in the game.

Never allow an incentive to be implemented where the creator participates in pleasure of the upside, but not the pain in the downside.

Skin in the game improves outcomes.

Clarity & Fluidity

An incentive is only as effective as:

(1) the clarity of its dissemination.
(2) the ability and willingness to adjust it based on new information.

Takeaway: Create even understanding playing fields for all constituents and avoid plan continuation bias.

So to recap, my framework for incentives:

(1) Objectives
(2) Metrics & Anti-Metrics
(3) Stakes & Effects
(4) Skin in the Game
(5) Clarity & Fluidity

I will be expanding on this topic in a newsletter this week. Subscribe so you don't miss it: https://t.co/qMB8i60ney

Follow me @SahilBloom for threads on frameworks and more.

If you are a job seeker looking to make the next leap forward in your career, check out my job board, where I curate roles at high-growth companies in finance and tech.

New roles every week! https://t.co/33hwlEqYaW