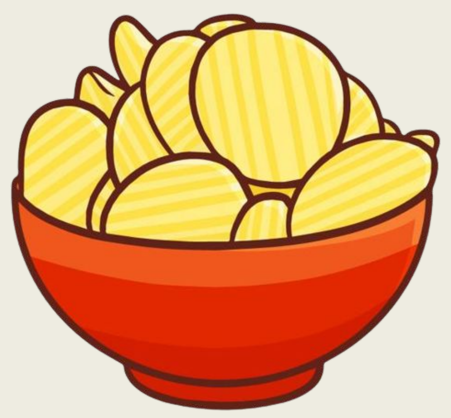
The Challenge of Understanding What Users Want: Inconsistent Preferences and Engagement Optimization

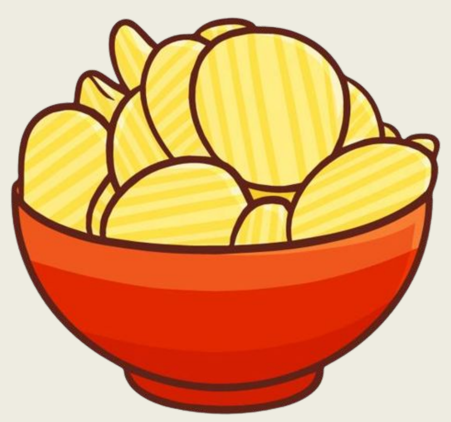
Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan

Inconsistent preferences in the offline world



You're at a party, and you finish the bowl of chips that your host put on the table

Your host notices and infers from your behavior that you want more chips



Your host refills the bowl

You finish the chips again

You leave the party feeling like you ate too many chips. What went wrong? In the moment, you wanted the chips. But after the fact you wish you had less. And the host made matters worse by constantly refilling the bowl.

Your host made a crucial **revealed preference** assumption: that they could infer your preferences from your behavior.

But your preferences were **inconsistent**, meaning your behavior in the moment did not reflect your long-term values.

Two selves

These inconsistencies manifest in a variety of settings

"Two-self" models: Different "selves" have different preferences in different times, contexts, etc.

Online groceries vs. in person [Milkman, Rogers, Bazerman 2010]
- "Want": ice cream
- "Should": vegetables

Mail-order DVDs vs. streaming [Milkman, Rogers, Bazerman 2009]
- "Want": action movies
- "Should": documentaries

Inconsistent preferences in the online world

Recommender systems tend to be built on these same revealed preference assumptions
- Maximize clicks, views, etc.

This can work well!
- For some content, behavioral signals can identify the best content. For example, the calculus tutorial with the most views might just be the best tutorial
- For other content, behavioral signals can be at odds with true preferences. Clickbait may get a lot of views but have low value

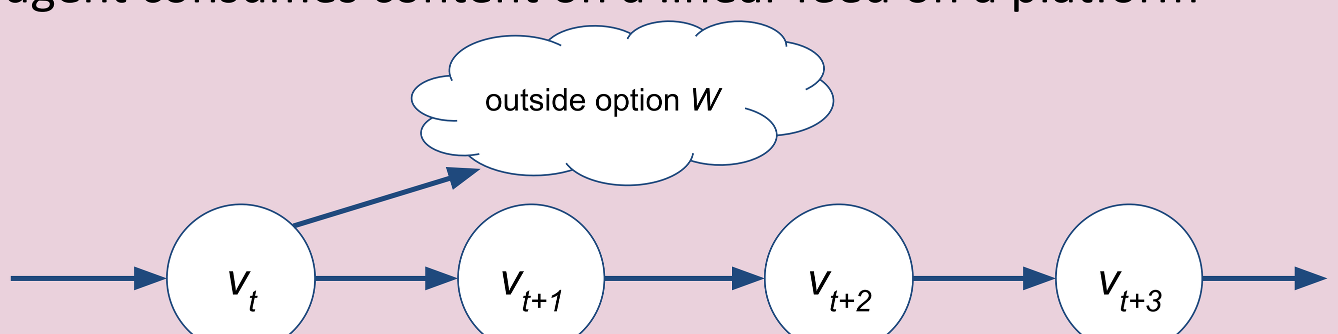
The same is true for food
- Refilling the salad bowl is less likely to run into preference inconsistencies
- But refilling the chips bowl may create problems

Social media contains both chips and salad

How could we tell the difference?
And how should we optimize differently for the two?

A formal model of consumption with inconsistent preferences

An agent consumes content on a linear feed on a platform



Content at time t has value v_t, and the agent has an outside option with value W.

The agent has two selves:
- System 1: quick, impulsive decisions
- System 2: long-term, reflective decisions

Each has different preferences and behavior. We think of System 2 as the agent's **true preferences**.

System 1 wants to remain on the platform as long as it's entertained. If it finds the current content sufficiently entertaining, it automatically moves to the next content.

System 2 remains on the platform as long as it gets positive expected utility for doing so.

Model (cont'd)

Let p be the probability that System 1 finds something entertaining. We can think of this as the content's **moreishness.**

Assume System 2 experiences diminishing returns. In particular, System 2 stops deriving value from the platform with probability $1-q$ at each timestep. Then, we can think of q as the **span** of the content, describing how long the agent truly wants to remain on the platform.

Finally let v be the expected value of content per unit time.

Results

There are two quantities of concern here: the agent's **utility** and **engagement** (how long it remains on the platform). In our model, we can write these as

Utility:
$$\mathbb{E}\left[S\right] = \max\left(\frac{\overline{v} - W}{1-q} - \frac{pW}{1-p},\ 0\right)$$

Engagement:
$$\mathbb{E}[T] = \left(\frac{1}{1-q} + \frac{1}{1-p} - 1\right) \cdot 1[\mathbb{E}[S] > 0]$$

(By assumption, the agent doesn't use the platform if it's expected utility for doing so is negative).

Consider an altruistic platform that wants to maximize System 2 utility, but only observes engagement (i.e., how long the agent remains on the platform). What's the effect of optimization?

Content manifolds

Platforms observe some features x about content. These features correspond to some latent parameters (p, q, v).
Optimizing for engagement $E[T]$ leads to some point ω_T in (p, q, v) space. What can we say about that point relative to ω_S, the point that maximizes utility $E[S]$?
1. We characterize content manifolds for which engagement optimization leads to suboptimal utility (Thm. 1).
2. We give conditions under which utility is nearly maximized (Thm. 2)

We also give strategies for identifying content manifolds, in an attempt to answer the question: how could we distinguish between salad and chips?