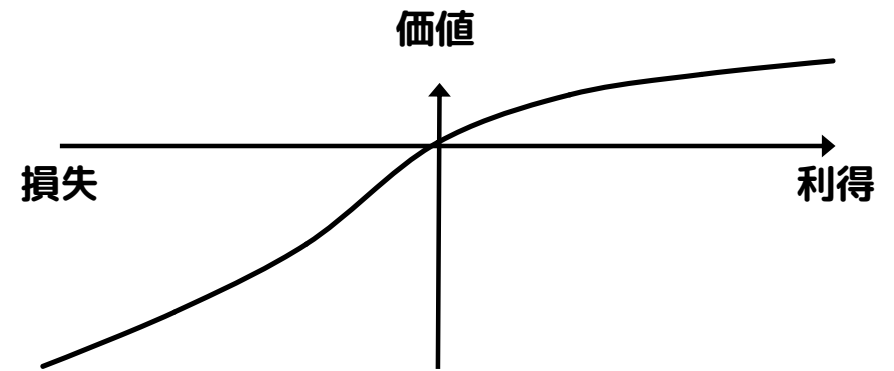


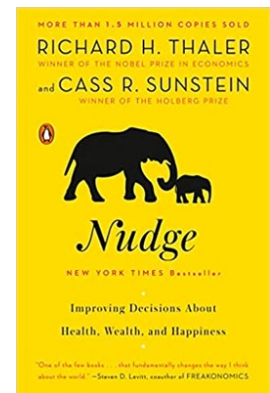
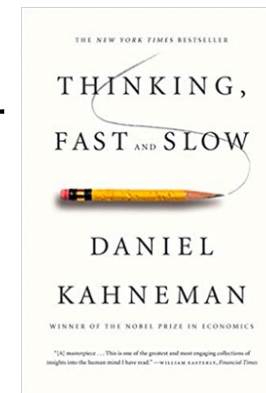
# 認知バイアスと デジタルプラットフォーム

大阪大学 大学院経済学研究科 大歳達也

# 認知バイアス



- 人の判断における系統的な偏り
- 近年注目を集めている
  - 2002年ノーベル経済学賞 ダニエル・カーネマン
    - システム1・システム2
      - 自動的で早い判断・意識的で遅い判断
    - プロスペクト理論
      - 損失回避・参照点・確率加重
  - 2017年ノーベル経済学賞 リチャード・セイラー
    - ナッジ（リバタリアン・パターンリズム）
      - 強制せず、より良い決定に導く



# システム1・システム2

- システム1

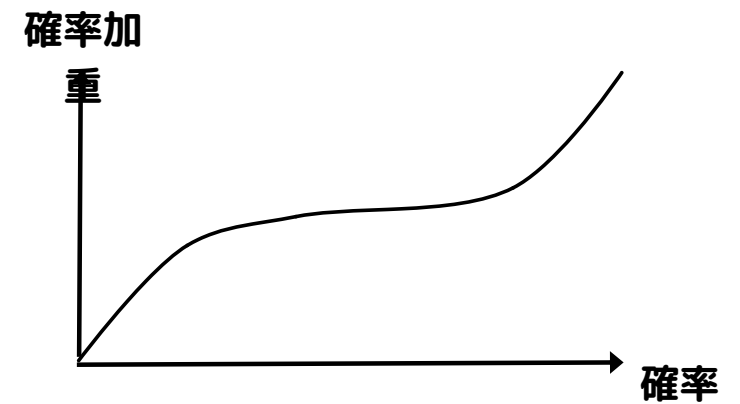
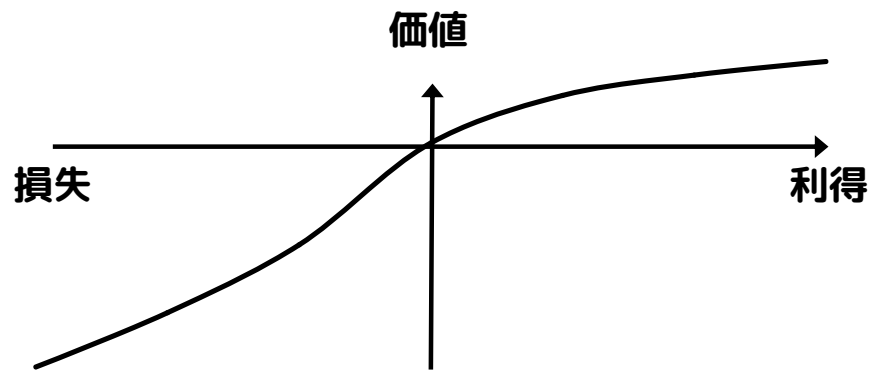
- 自動的
- 早い判断
- 直感的
- 連想的
- 認知バイアス

- システム2

- 意識的
- 遅い判断
- 論理的
- 分析的

# プロスペクト理論

- 不確実下での意思決定をモデル化した理論
  - 価値関数と確率加重関数による期待値により判断
- 価値関数
  - 利得と損失に対する価値が非対称→損失回避
- 確率加重関数
  - 低確率の過大評価



# ナッジ

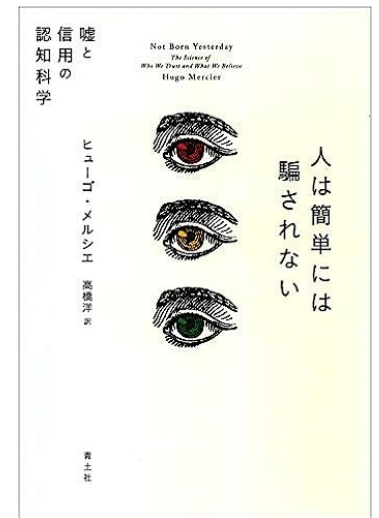
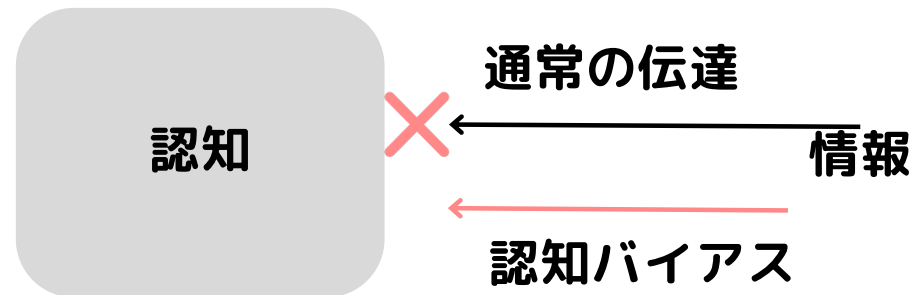
- 良い選択を取るように行動変容を「促す」こと
- 強制はせず、選択肢が用意されている（リバタリアン・パターナリズム）
- リバタリアン：個人が自由に選択できる
  - バイアスの影響が深刻に
- パターナリズム：「親」が良い選択を選ばせる
  - バイアスのない「親」によってバイアスの影響が排除されるが自由もない
- スラッジ：悪い選択を取るように行動変容を促す

# AIにも見られるバイアス

- データバイアス
  - ゴリラ問題
- Shape/Texture バイアス
  - CNN (Convolutional Neural Network): Texture バイアス
  - ViT(Vision Transformer): Shape バイアス
  - 人間の子供：Shapeバイアス
- Chain of Thought (LLM)
  - 「ステップバイステップで考えよう」 → 計算問題に正答
  - 人間：システム1/システム2
- 最初と最後に重要情報を与えると正答率が上がる (LLM)
  - 人間：Primary/Recency Effect

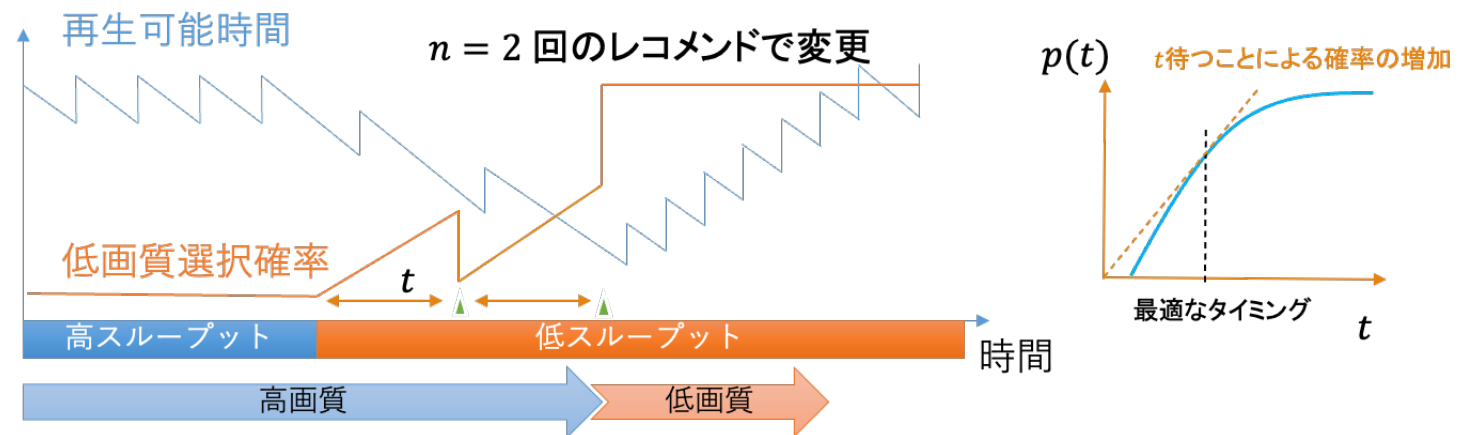
# チャンネルとしての認知バイアス

- 人の行動変容は意外と難しい
  - 習慣化された行動が選択されやすい
  - 自分の考えに合わない情報はフィルタリングされる
- 人の認知に働きかける手段（チャンネル）となり得る
- 人の行動変容への応用が期待されている
  - ナッジ
  - 仕掛学



# UI/UXへの応用は多い

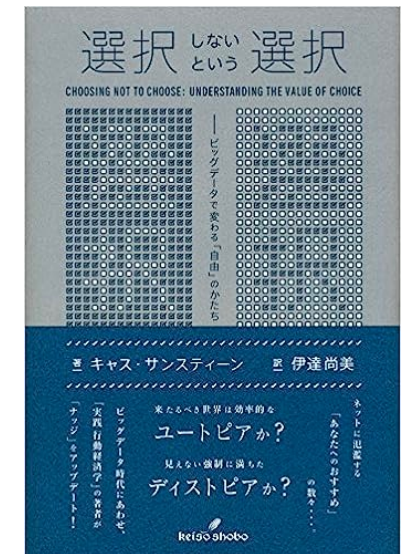
- 「ベストセラー」：社会的証拠
- 「空室はあと何室です」：時間制約
- 「プレミアム」：アンカリング
- レコメンドのタイミング：認知不協和





# 制度設計においても重要

- 強力なデフォルト効果
  - ex. ドナーカードのオプトアウト/オプトイン
- 中立的なデフォルトを用意することは困難
  - 何かを設計した時点で何かを選んでいる
    - オプトアウト・オプトアウト
    - 空欄の処理方法
- プラットフォームにもデフォルトは欠かせない
  - テンプレートを個別にカスタマイズ



# セキュリティにおけるナッジ

- ナッジ型のサイバーセキュリティ
  - 人間像の捉え直し
    - 必ずしも合理的ではない
      - セキュリティの対策をしないとリスクがあるが対策しない
    - 注意力は高くない
    - 気分・感情で行動することもある
  - 認知科学・心理学の知見を応用



# まとめ

- 人は常に合理的な選択を行うわけではなく、合理的な人を前提としたシステムには限界がある
- 人の生成物を学習したAIにも人の認知バイアスのような挙動が見られることもある
- 人、AIがアクターとなるデジタルプラットフォームでも認知バイアスを考慮した設計が今後重要となると考えられる