

# 消費者調査における 真実申告誘発メカニズムの活用

株式会社インサイト・ファクトリー  
小野滋 ono@insightfactory.jp

# 背景

- いかにして質の高い回答を得るか？
  - 調査における本質的な問いのひとつ
  - 消費者調査においては、近年特に深刻な問題となっている
    - ネット調査パネルの普及
    - 調査のコモディティ化・低価格化



## アプローチ

- A. 個々の回答の質を評価する** (→ 質の低い回答を事後的に取り除く)
- B. 質の高い回答行動を促進する**

## A. 個々の回答の質を評価する

- 従来の方法：外的基準との比較
  - 外的基準
    - 「真剣かつ誠実に回答している調査参加者であれば、きっとこのように回答するだろう」と調査者が考える回答
  - 例
    - 社会的望ましさ尺度
    - 虚偽尺度項目
    - …
- 限界
  - 調査内容に適した、適切な外的基準を用意するのが難しい
  - 例) 社会的望ましさ尺度のスコアが高い人は…
    - 回答における社会的望ましさバイアスが高い人？
    - 本当に高潔かつ英雄的な人なのかも？

## B. 質の高い回答行動を促進する

- 従来の方法：調査票の改善
  - 教示の改善
  - 聴取項目のワーディングの改善
  - 質問紙レイアウトの改善
  - ...
- 暗黙の前提：「調査参加者は、もともと真剣かつ誠実に回答する傾向を持っている」
  - 調査票改善は、その行動傾向を阻害しないことを目指している
- 限界
  - なぜ、調査参加者は真剣かつ誠実に回答しようとしていると想定できるのか？

- 調査は**対人的行為**である

調査という行為において、回答者が目指している**対人的目標**とは？  
(Kuncel, Borneman & Kiger, 2012)

- Credible
- True to the Self
- Impressive



- 調査の設計 = <回答者の目標追求を支援する、対人相互作用のデザイン>

## ■ 調査回答行動へのゲーム理論的視座

- なぜゲーム理論？
  - 対人相互作用のデザインに対する、現代の代表的枠組み
- B. <質の高い回答の促進>とは？  
→ 回答行動に対する(有形・無形の)インセンティブを、真剣かつ誠実な回答行動と整合的な形にデザインするという問題  
(**真実報告誘発メカニズム** truth-telling mechanism)
- A. <回答の質の評価>とは？  
→ 個々の参加者が持っている、<与えられた情報とインセンティブ・メカニズムの下で合理的に行動しようとする傾向>を評価するという問題



## ■ ベイジアン自白剤

- 調査手法研究における、ゲーム理論の観点からの提案の代表例

# ベイジアン自白剤



# ベイジアン自白剤とは

---

Bayesian Truth Serum (Prelec, 2004)

- カテゴリカル質問への回答に対し、スコア(**BTSスコア**)を付与するルール
- 対象者がBTSスコアの最大化を目指すとき、真実申告がベイジアン・ナッシュ均衡解となる
  - = 他の対象者が真実を申告しているという仮定の下で、最適方略(BTSスコアの期待値を最大化する方略)は真実の申告である
- ほかの均衡解もありうるが、期待値は真実申告時に最大となる



- 意義
  - ・ BTSスコアは回答の真実性 (truthfulness) を表す
  - ・ スコアをインセンティブに連動させることで、真実性のある回答を引き出せる

- 必要な設問

Q1: 任意のカテゴリ質問 (K択質問)

Q2: 「Q1に他の人がどう答えるか」を予測させる質問

2択質問の例:

Q1. これまでに万引きをしたことはありますか?	Yes	No
--------------------------	-----	----

Q2. Q1にYesと答える人は、どのくらいいると思いますか?	_____ %
---------------------------------	---------

# ベイジアン自白剤とは

## ■ BTSスコアの例

Q1. 万引きしたことはありますか？	Q2. Q1にYesと答える人はどのくらいいると思いますか？		BTSスコア
はい	20%	⇒	+0.31
いいえ	10%	⇒	-0.18
はい	5%	⇒	+0.09
いいえ	30%	⇒	-0.09
...	...		...
はい	25%	⇒	+0.32

真実性が高い

真実性が低い

「はい」回答率  
25%

「はい」率予測の平均  
18%

※上の例のQ1は2択設問だが、任意のカテゴリ数のSA設問について同じ方法でBTSスコアを算出できる

## ■ BTSスコアの定義

対象者  $i$  ( $=1, \dots, n$ ) のカテゴリ  $k$  ( $=1, \dots, K$ ) に対する Q1 での選択有無を  $x_{ik}$  , Q2 での回答を  $y_{ik}$  とする。

$$\bar{x}_k = \frac{1}{n} \sum_i^n x_{ik}$$

$$\log \bar{y}_k = \frac{1}{n} \sum_i^n \log y_{ik}$$

$$BTS \text{ Score}_i = \underbrace{\sum_k^K x_{ik} \log \frac{\bar{x}_k}{\bar{y}_k}}_{\text{情報スコア}} + \alpha \underbrace{\sum_k^K \bar{x}_k \log \frac{y_{ik}}{\bar{x}_k}}_{\text{予測スコア}} \quad (0 < \alpha < 1)$$

# ベイジアン自白剤とは

## ■ 計算例

データ (架空例)

		対象者A	対象者B	...
回答	Q1. この製品を買ってみたいですか？	はい	いいえ	...
回答率 予測	Q2. Q1に「はい」と答える人は何パーセントいると思いますか？	25%	30%	...

集計結果	
実際の「はい」回答率	実際の「いいえ」回答率
<b>25%</b>	<b>75%</b>
「はい」回答率 予測の幾何平均	「いいえ」回答率 予測の幾何平均
<b>31%</b>	<b>65%</b>

情報スコア	-0.22	+0.08	...
-------	-------	-------	-----

「はい」回答者に与える情報スコア $\log(0.25/0.31)$ =-0.22	「いいえ」回答者に与える情報スコア $\log(0.75/0.65)$ =+0.08
---	--

「はい」についての対象者の回答率予測と、実際の回答率との比の対数	$\log(0.25/0.25)$ =0.00	$\log(0.30/0.25)$ =+0.18	...
「いいえ」についての対象者の回答率予測と、実際の回答率との比の対数	$\log(0.75/0.75)$ =0.00	$\log(0.7/0.75)$ =-0.07	...
上記の2つの値を実際の回答率で重みづけた和	$0.25 \times (0.00)$ $+ 0.75 \times (0.00)$ =0.00	$0.25 \times (+0.18)$ $+ 0.75 \times (-0.07)$ =-0.01	...
<b>予測スコア</b>	<b>0.00</b>	<b>-0.01</b>	...

(情報スコア) + a (予測スコア) (ここではa=1の場合を示す)	-0.22+0.00 =-0.22	-0.08+(-0.01) =+0.07	...
<b>BTSスコア</b>	<b>-0.22</b>	<b>+0.07</b>	...

情報スコアは、Q1においてその回答カテゴリが“surprisingly common”である程度を表す

予測スコアは、その対象者のQ2への回答の正確さを表す

A. <個々の回答の質を評価する>ための手法として…

- **外的な「正解」を用いない**  
「正解」が存在しない質問についてもスコアリングできる
- **回答の分布に依存しない**  
少数意見であっても高いスコアを得る可能性がある

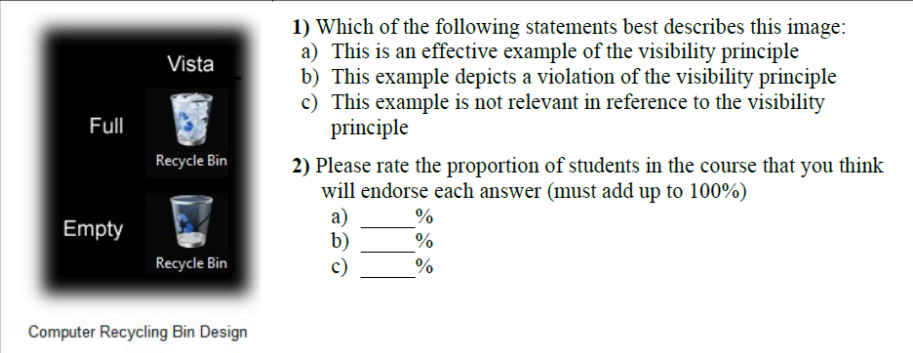
B. <質の高い回答行動を促進する>ための手法として…

- **真実報告誘発メカニズムを提供**  
スコアをインセンティブに連動させることで、望ましい行動 (真実申告) を引き出せる

## A. <個々の回答の質を評価する> タイプの実証研究

### ■ BTSによる教育評価 (Miller, Bailer, & Kirlik, 2014)

- デザイン教育では、デザインが守るべき「デザイン原理」を教えるが、受講者がそれを真に理解したかどうかを採点するのは困難
  - 原理を正しく理解していても、それを現実のデザインに当てはめると、簡単な「正解」は存在しないから
- デザイン例を示し、「デザイン原理を守っているか」「他の人はどう答えると思うか」を聴取。回答を教師が採点するかわりに、BTSスコアを算出。



1) Which of the following statements best describes this image:

- This is an effective example of the visibility principle
- This example depicts a violation of the visibility principle
- This example is not relevant in reference to the visibility principle

2) Please rate the proportion of students in the course that you think will endorse each answer (must add up to 100%)

- \_\_\_\_\_ %
- \_\_\_\_\_ %
- \_\_\_\_\_ %

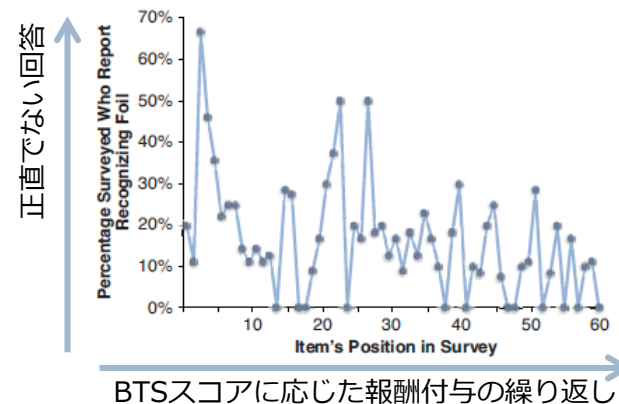
Computer Recycling Bin Design

- 教師の採点よりもBTSスコアのほうが、その後のデザイン制作実習における学生のパフォーマンスと相関が高かった

## B. <質の高い回答行動を促進する>タイプの実証研究

### ■ BTSによる正直さの促進 (Weaver & Prelec, 2013)

- 消費者にブランド名を示し「知っているブランドは？」と尋ね、「知っている」と答えたら10セント渡す
  - 不誠実な回答を促進するための手続き。調査参加者は実在しないブランド名に対してさえ「知っている」と答えやすくなる
- ブランド名を示し「知っているブランドは？」「他の人はなんと答えると思いますか？」と尋ね、答えるたびにBTSスコアをフィードバック
  - 知っていると答えたら10セント渡すが、さらにBTSスコアに応じた少額のコインを渡す
- **BTSスコアと連動した報酬を繰り返し与えると、参加者は次第に実在しないブランド名を「知らない」と答えるようになった**





# 実験

- マーケティング・リサーチにおけるベイジアン自白剤の有用性について検討する
- <個々の回答の質を評価する> タイプの用途に焦点を当てる
- 先行研究:
  - デザイン教育評価(Miller, Bailer, & Kirlik, 2014)
  - 医師の新薬処方意向 (Howie, Wang, & Tsai, 2011)
- 本実験の特徴:
  - マーケティング・リサーチにおいて一般的な文脈 (新商品評価)を用いて検証
  - 管見のかぎり、日本語での初の実証研究

## ■ あるメーカーの販売員に対する定例調査

- 販売員の製品知識は高い
- 勤務時間中に、業務の一環として回答
- 自社の上市済製品・上市前新製品について聴取
- 強い評価懸念バイアスが働いているものと考えられている
  - 本社部門に対する、自社製品についての回答であるため

## ■ 対象者

全国の販売員341名から回答を収集

実質的には匿名調査

ただし、調査回答の匿名性の程度について対象者がどう捉えているのかはわからない

- 調査票中の以下の4問について、直後にBTS設問(他の販売員の回答の予測)を聴取

[現在販売中の製品について]

Q1. お客様の反応はいかがですか

(良い～悪い, 5件法)

Q2. あなた自身の、現在の取り組みはいかがですか

(かなり積極的～積極的でない, 4件法)

[来季に発売される、現在紹介活動中の新製品について]

Q3, Q4 (同上)

## 評価懸念バイアスが強いと思われる項目で、ポジティブ回答のBTSスコアが低い

顧客反応の評価 (Q1, Q3)においては

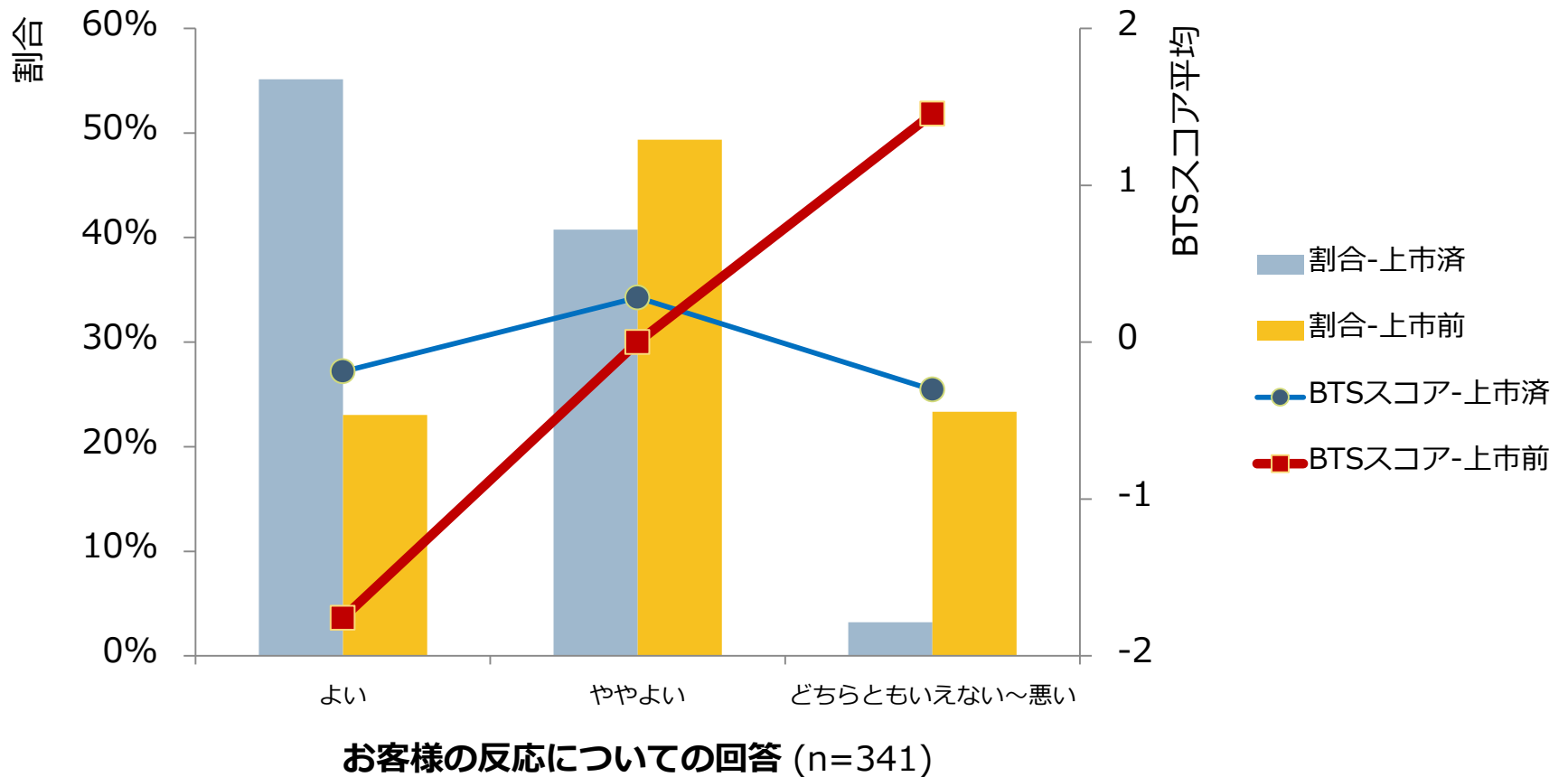
- 上市前製品については、ポジティブな評価のBTSスコアが低い
- 上市済製品については、評価とBTSスコアは関連しない
  - 顧客反応は関係者にとって既知の事実であるため

自分の取り組みについての評価 (Q2, Q4) においては

- 上市前・上市済を問わず、ポジティブな評価のBTSスコアが低い
  - 上市済製品であっても、個々の回答者の取り組みの実態について調査者が知ることはきわめて困難なため

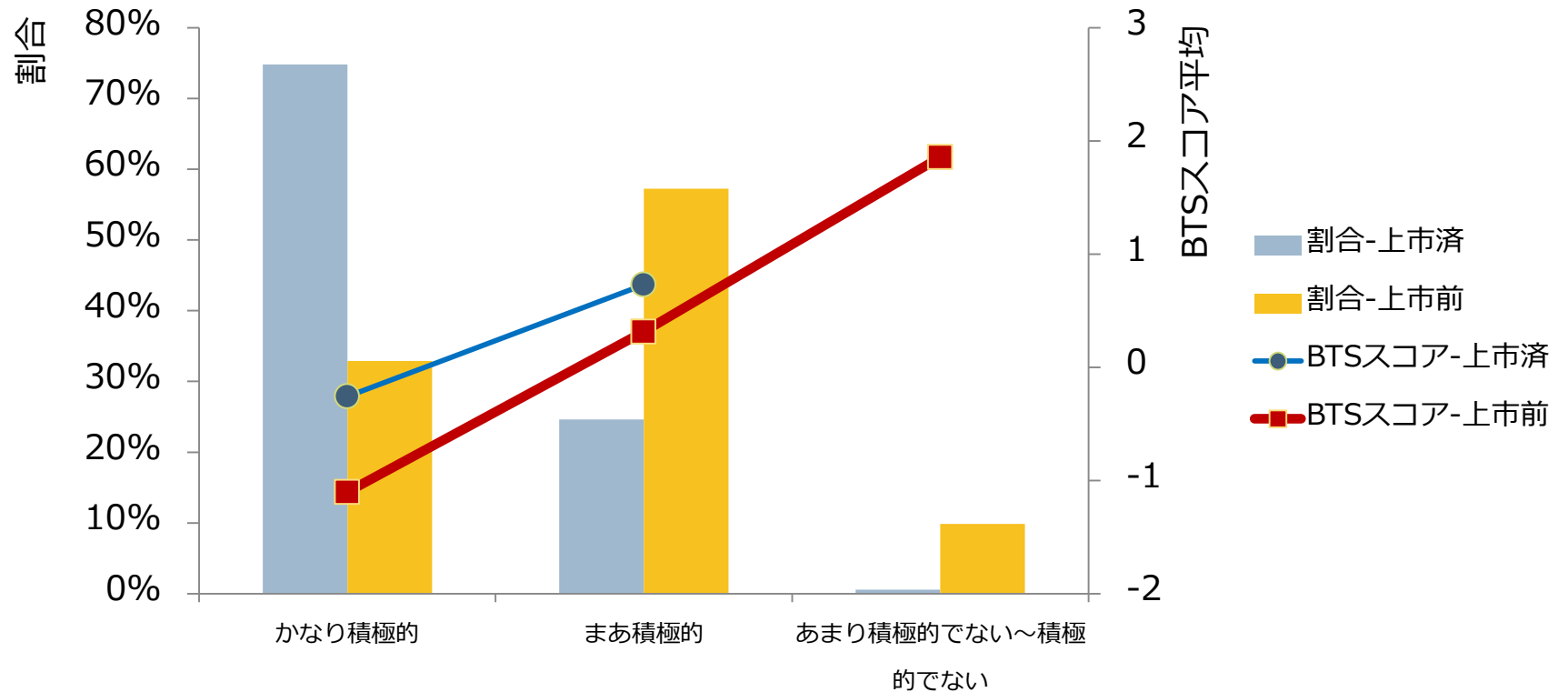
## 顧客反応の評価

上市前製品についての評価で、ポジティブな評価のBTSスコアが低い



## 自分の取り組みについての評価

上市前製品・上市済製品のいずれにおいても、ポジティブな評価のBTSスコアが低い



自分の取り組みについての回答 (n=304)

(n<10の箇所はBTSスコア平均の図示を省略)

# 考察



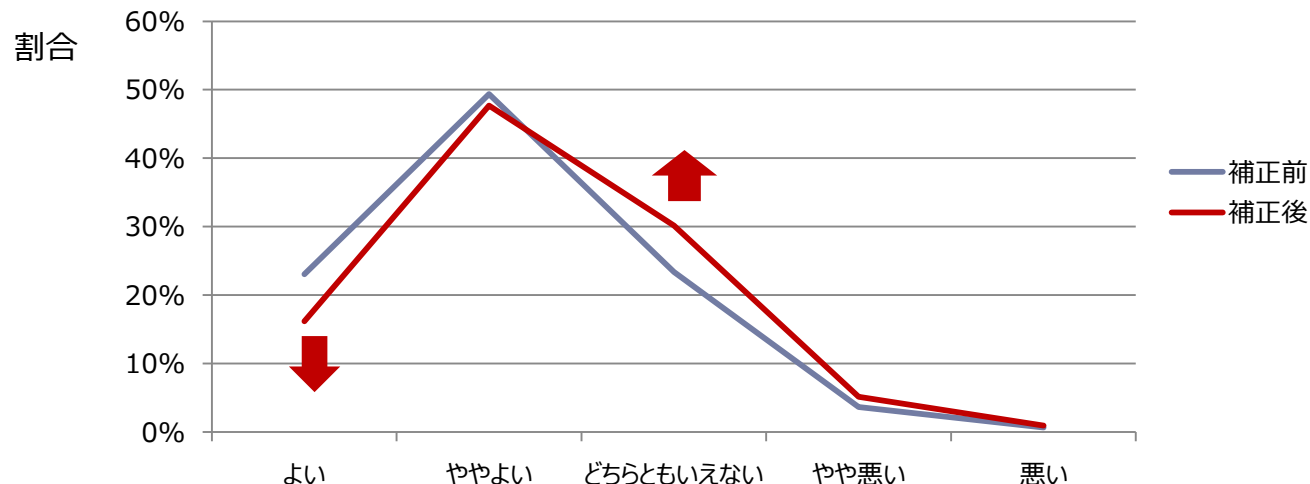
# 本実験から得られた知見

## ■ 仮説を支持する結果が得られた

評価懸念バイアスが強いと思われる項目で、ポジティブ回答のBTSスコアが低い

## ■ 調査回答の質の評価手法としての、**ベイジアン自白剤の有用性**を示唆

想定される用途 … BTSスコアによる重みづけ集計によって回答バイアスを補正



### Q3回答分布の補正例

回答を $\exp(\text{BTSスコア} \times 0.2)$  で重みづけて集計

### A. <回答の質を評価する>手法として

- 外的基準(“正解”)を用いた検証  
例, 行動意向のBTSスコアと、実際の行動との関連性の検証  
→ 現在取り組みを進めている
- BTSスコアを用いた、適切な回答分布補正手法の開発
- BTSスコアと個人特性との関連性の検討  
本実験におけるBTSスコアは、4設問の間で  $r=+.21\sim+.54$

### B. <質の高い回答行動を促進する>手法として

- インセンティブとBTSスコアの連動についての実証研究  
消費者調査における例は少ない → 現在取り組みを進めている

### ベイジアン自白剤を支える**心理的機序**

- 認知科学・行動計量学の観点からの基礎研究が望まれる

# ご清聴ありがとうございました

## 引用文献

Howie, P.J., Wang, Y., Tsai, J. (2011) Predicting new product adoption using Bayesian truth serum. *Journal of Medical Marketing*, 11, 6-16.

Kuncel, N.R., Borneman, M., & Kiger, T. (2012) Innovative item response process and Bayesian faking detection methods: More questions than answers. in Ziegler, M., Maccann, C., & Roberts, R.D. (eds.) "*New perspectives on faking in personality assessment*", Oxford University Press.

Miller, S.R., Brailey, B.P., Kirlik, A. (2014) Exploring the utility of Bayesian truth serum for assessing design knowledge. *Human-Computer Interaction*. 29(5-6), 487-515.

Prelec, D. (2004) A Bayesian Truth Serum for Subjective Data. *Science*, 306(5695), 462-433.

Weaver, R. & Prelec, D. (2013) Creating truth-telling incentives with the Bayesian Truth Serum. *Journal of Marketing Research*, 50(3), 289-302.

小野滋(2014b) 正直で真剣な回答を引き出す仕組み, マーケティング・リサーチャー, 124.