

市場調査における ベイジアン自白剤の活用

インサイト・ファクトリー
小野 滋

目次

1. ベイジアン自白剤

1.1 ベイジアン自白剤とは

1.2 インサイド・ベイジアン自白剤

1.3 既存研究

2. 市場調査からみたベイジアン自白剤

2.1 背景

2.2 意義

2.3 課題

3. 研究計画I. 消費者調査における活用 (Web公開版では割愛)

4. 研究計画II. 従業員調査における活用 (Web公開版では割愛)

1. ベイジアン自白剤

1.1 ベイジアン自白剤とは

- “Bayesian Truth Serum”

- 心理学者 D. Prelec によって提案された

- 質問に対する回答へのスコアリング手法

- 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」を予測させる設問

2択設問の例:

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？	_____ %	

1.1 ベイジアン自白剤とは

■ BTSスコア

対象者 i のカテゴリ 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。

下式で定義するBTSスコアの期待値は、真実回答において最大化される。
BTSスコアをインセンティブとすると、真実回答がベイジアン・ナッシュ均衡となる。

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

1.1 ベイジアン自白剤とは

■ 算出例

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

質が高い回答

質が低い回答

「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

1.1 ベイジアン自白剤とは

■ 算出例

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
25%	75%
「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
31%	65%

情報スコア	-0.22	+0.08	...

「はい」回答者に与える情報スコア	「いいえ」回答者に与える情報スコア
$\log(0.25/0.31)$ =-0.22	$\log(0.75/0.65)$ =+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	...
予測スコア	0.00	-0.01	...

(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	...
BTSスコア	-0.22	+0.07	...

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

1.1 ベイジアン自白剤とは

■ 特徴

外的な「正解」に依存しないスコアリング

「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スコアリング

少数意見であっても高いスコアを得る可能性がある

インセンティブ整合性

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

1.2 インサイド・ベイジアン自白剤

■ Prelec (2004) が実際に示したことはなにか？

■ セッティング

- 対象者 r の問1の回答を二値ベクトル x^r , 問2の回答をベクトル y^r で表す。
選択肢が(イエ, ハイ)の2つで、回答が問1「イエ」問2「イエは20%」だったら、 $x^r = (1, 0)$,
 $y^r = (0.2, 0.8)$
- 対象者 r の問1への本当の答えを二値ベクトル t^r で表す。
もし本当の答えがハイなら、 $t^r = (0, 1)$
- 選択肢 k の選択率についての対象者 r の本当の推測を $p(t_k | t^r)$ と表す。
- 母集団における t の分布をベクトル Ω で表す。
本当の答えがハイの人が全体の2割なら、 $\Omega = (0.8, 0.2)$

1.2 インサイド・ベイジアン自白剤

■ Prelec の前提

A. 対象者 $1, 2, \dots, n$ の本当の答え t^1, t^2, \dots, t^n は、 Ω の下で互いに独立である

B. 本当の答えが違っているとき、そのときに限り、本当の推測が異なる

もし $t^r \neq t^s$ なら、そのときに限り $p(t_k | t^r) \neq p(t_k | t^s)$

※ Prelec が挙げている、前提B.が破られる例:

(1)ふたりの t は異なるのに、ふたりの $p(t_k | t^r)$ がほぼ等しい場合

例: 公的情報が利用可能で、個人的意見が情動的でないとき。たとえば、母集団における女性の割合についての判断。本人の性別は効かない。

(2)ふたりの t は同じなのに、ふたりの $p(t_k | t^r)$ が異なる場合

例: 好みや性質のちがう人が混じっていて、違う理由で同じ答えを持ち、しかし母集団についての推測は異なる場合。

■ Prelec が示したこと

- **他の人は正直だと仮定したとき、正直に答えることがベイジアン・ナッシュ均衡となる。**
- **ほかの均衡解もあるうるが、情報スコアの期待値は正直に答えた時が最大。**

1.2 インサイド・ベイジアン自白剤

■ 小野の素朴な疑問のコーナー

1. どんな質問であれ、他者の回答の推測に際しては、自分の信念以外にも、なんらかの手がかり情報があるのではないのでしょうか？

「うちのおばあちゃんならきっとイエと答えるだろう」とか…

このとき $t^r = t^s$ かつ $p(t_k | t^r) <> p(t_k | t^s)$ となるのでは？

2. 「回答」を情報不完備ゲームにおける行動として捉えたとき、「他の人は正直だ」という仮定は、どの程度まで現実的なのでしょうか？

3. 「あるスコアを最大化するために真実回答が均衡解だ」ということは、「そのスコアが高い回答は真実性が高い」ということを含意するのでしょうか？

4. そもそも、人の行動はどの程度合理的なものなのでしょうか？

聞くとところによれば、“平均値推測ゲーム”では、参加者の行動はナッシュ均衡解になかなか一致しないそうではないですか…？



1.3 既存研究

A. 調査結果の精度向上

BTSスコアでウェイトニングして集計・分析

B. 優れた回答者の特定

BTSスコアが高い回答者をピックアップ

C. 回答行動の変容

BTSスコアを報酬と連動させる

A. 調査結果の精度向上

■ BTSによる予測精度向上 (Howie, Wang, & Tsai, 2011)

医師の処方意向は、BTSスコアが高いときにあてになる

上市前の処方薬について医師に処方意向(「処方したいですか?」)を聴取しても、その回答は上市後の実際の処方とあまり関係しないことがわかっている。

↓

処方意向とともに「他の医師が処方する割合は?」と尋ね、BTSスコアを算出。

↓

BTSスコアでウェイトニングして分析すると、上市後の実際の処方を上市前処方意向で予測するモデルの説明率が向上した。

B. 優れた回答者の特定

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

BTSスコアが高い人は、質問をほんとうに理解している

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。



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。



The image shows two computer recycling bin designs. The top design is labeled 'Vista' and shows a 'Full' bin and an 'Empty' bin, both labeled 'Recycle Bin'. The bottom design is labeled 'Empty' and shows a single 'Recycle Bin'.

1) Which of the following statements best describes this image:
a) This is an effective example of the visibility principle
b) This example depicts a violation of the visibility principle
c) 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%)
a) _____ %
b) _____ %
c) _____ %

Computer Recycling Bin Design

C. 回答行動の変容

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

BTSスコアに応じて報酬を渡すと、回答者は正直になる

調査において、調査参加者は往々にして正直に答えていない。

↓

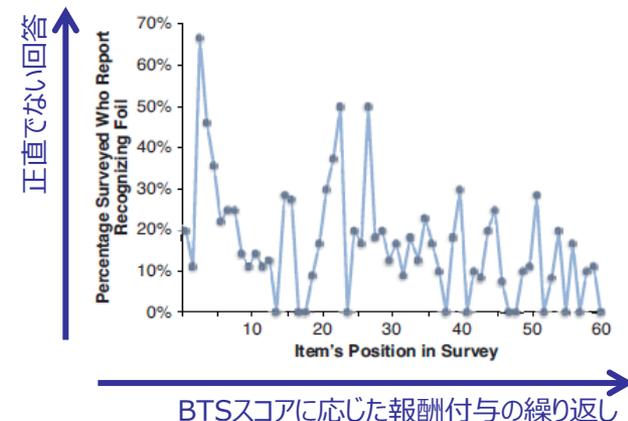
ブランド名を示し「知っているブランドは？」と尋ね、「知っている」と答えたら10セント渡す。このとき、調査参加者は実在しないブランド名に対してさえ「知っている」と答えやすくなる。(不誠実な回答を促進するための手続き)

↓

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。

↓

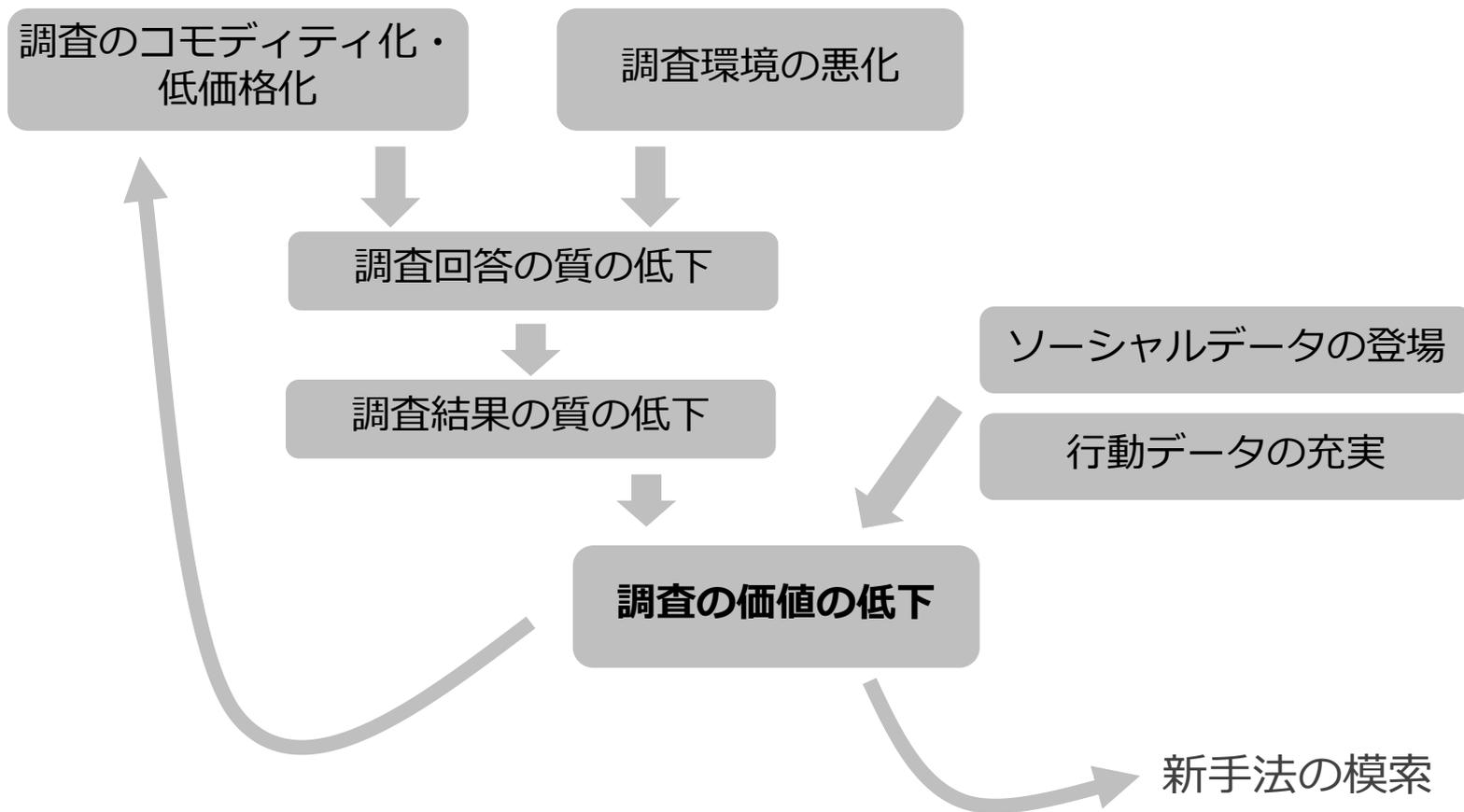
これを繰り返すと、参加者は次第に実在しないブランド名を「知らない」と答えるようになる。



2. 市場調査からみたベイジアン自白剤

2.1 背景

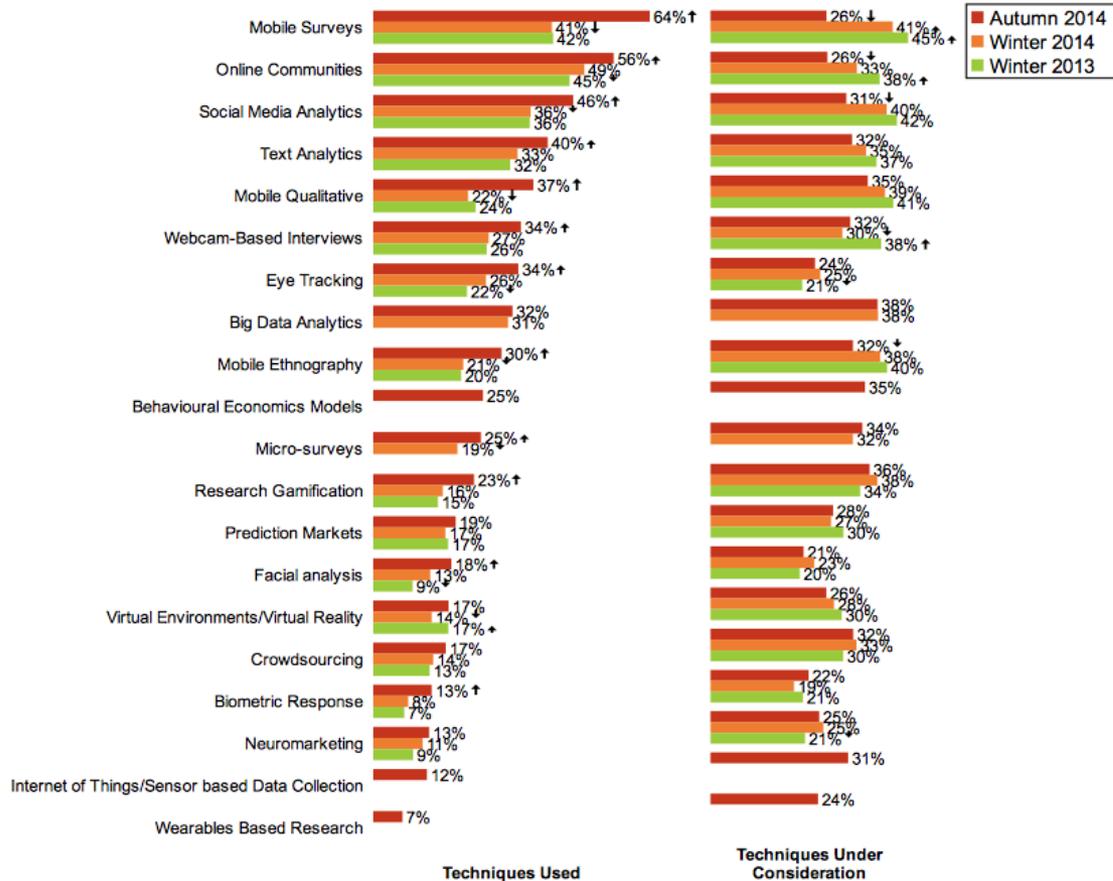
■ 消費者調査の悲劇



2.1 背景

■ 新手法の模索

Techniques Used/Considered: Trend



2013 to 2014; Base: 435 to 494 Autumn 2014; 0 to 2229 Winter 2014; 0 to 644 Winter 2013

Some empty rows or columns have been removed

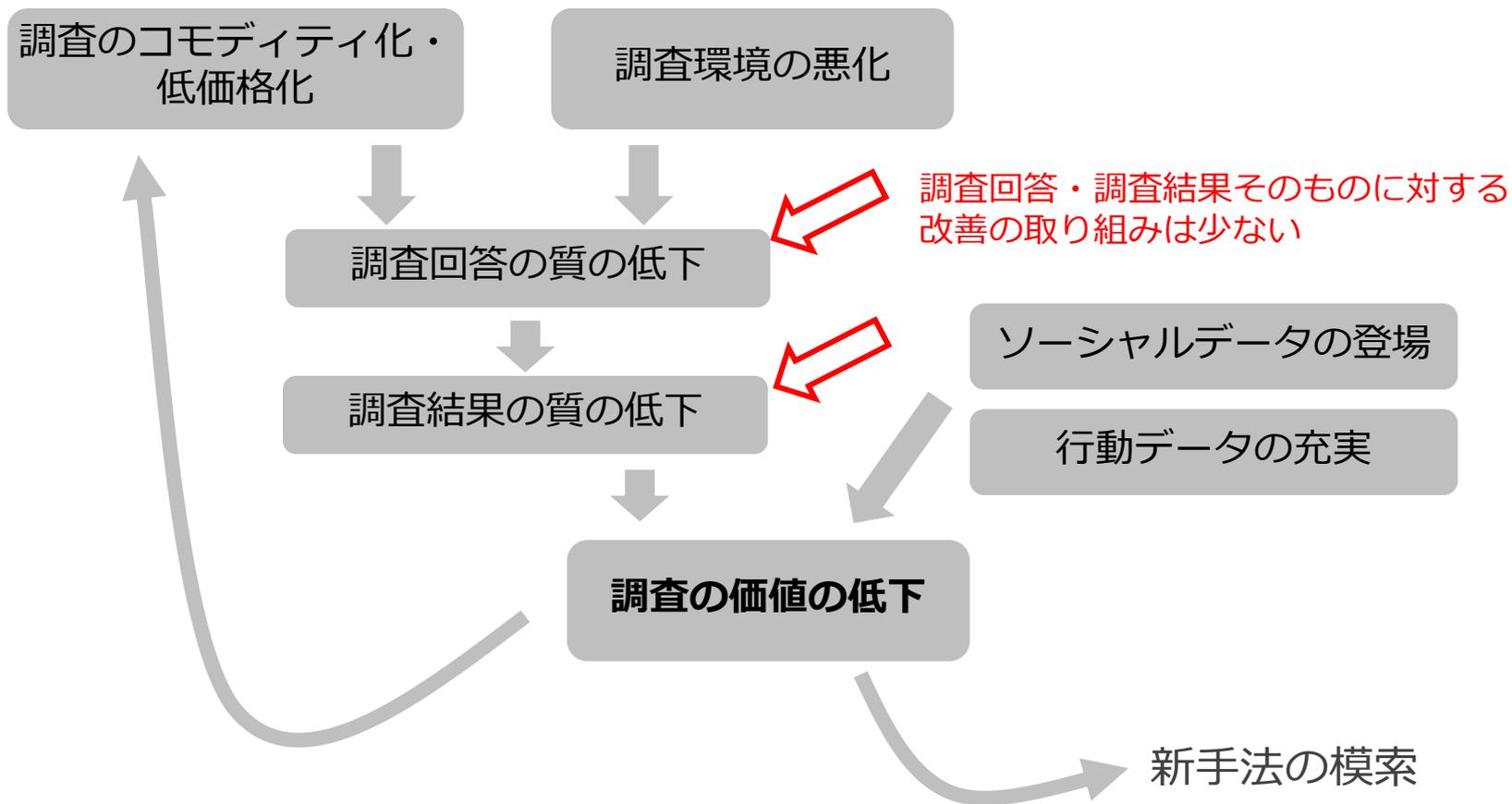
What use of these techniques and approaches do you see ahead in your future? For each statement, click an option.

GRIT Q1-Q2 2014 Report "Adoption of New Research Methods"

<http://www.greenbookblog.org/2014/09/29/the-top-20-emerging-methods-in-market-research-a-grit-sneak-peek/>

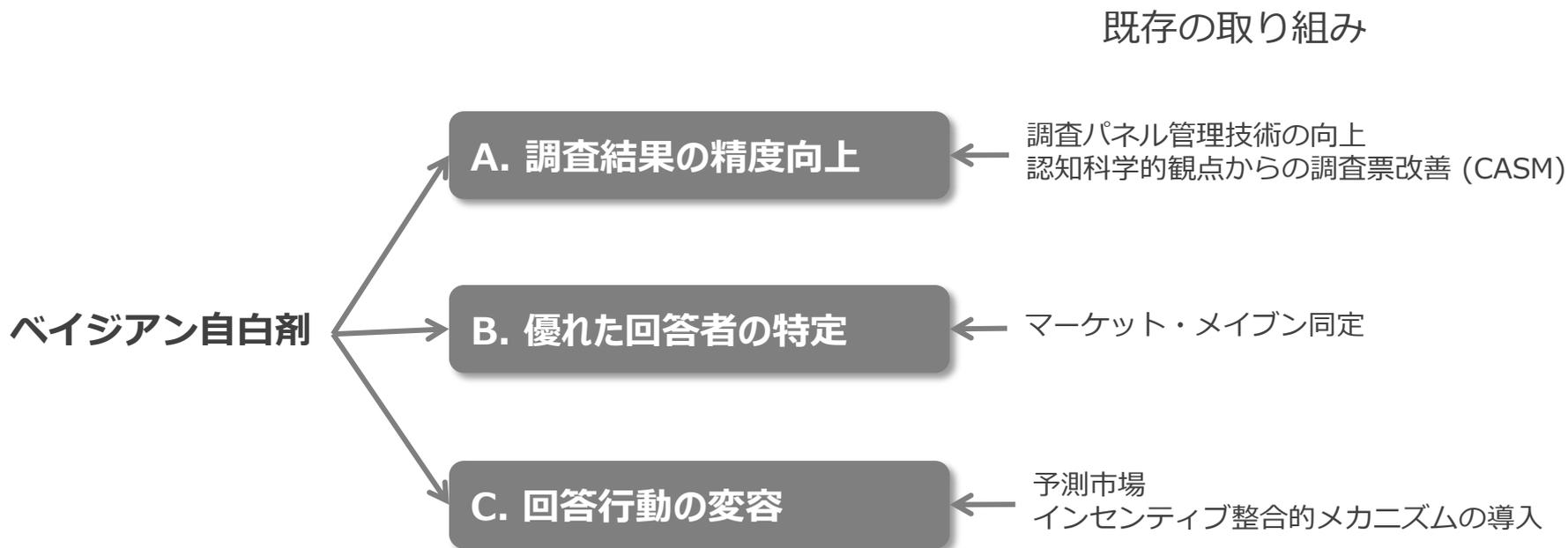
2.1 背景

■ 消費者調査の不思議



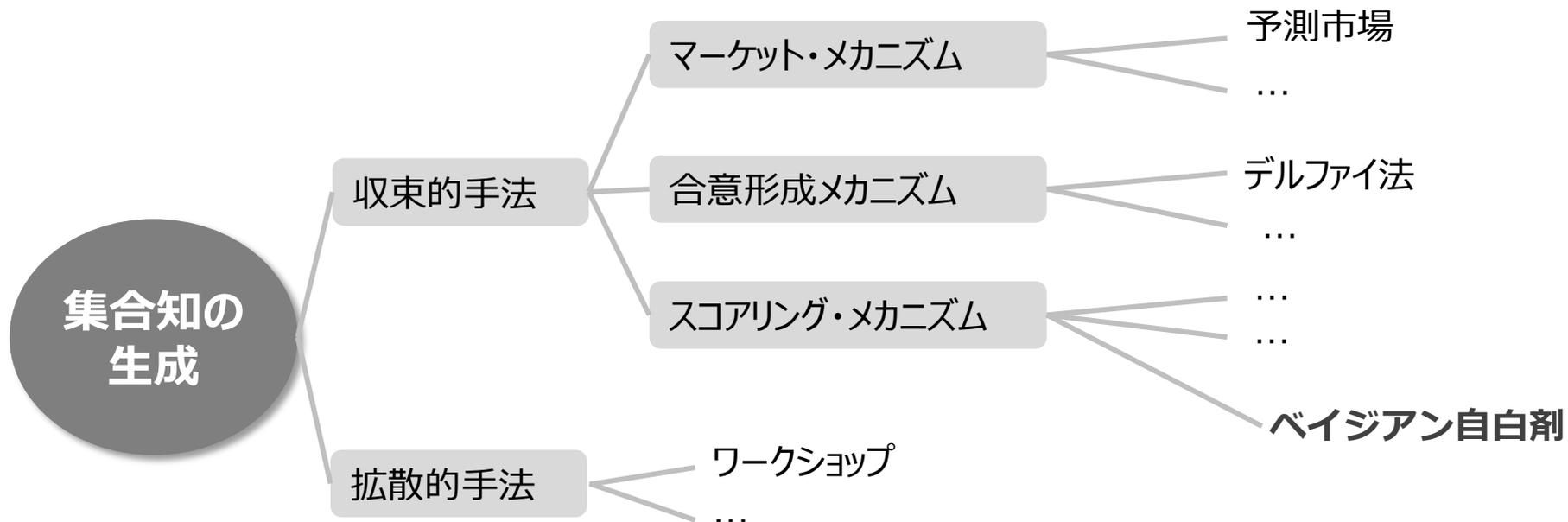
2.2 意義

■ 期待される貢献



2.2 意義

■ 集合知活用における位置づけ？



※集合知: 多くの人の知識・思考から、その単純な総和を超えた、価値ある情報を導き出すこと

2.3 課題

■ 基準関連的妥当性の検証

BTSスコアが高い回答は、実際に真実性が高いのか？

BTSスコアをインセンティブとすると、真実回答が引き起こされるか？

→ 外的基準 (なんらかの「正解」) を用いた検証は、いまだ十分ではない

■ 認知的基盤 (作用機序) の解明

1. 「BTSスコアが高い回答は真実性が高い」として、それはなぜ？

- Q1, Q2の質問文を正確に理解しているから
- Q1, Q2の回答にあたって十分な注意・推論を行っているから
- Q1, Q2の回答に必要な知識を十分に持っているから
- BTSスコアは真実性を表し、その人は誠実だから

2. 「BTSスコアによって真実回答が引き起こされる」として、それはなぜ？

- 教示におそれをなしたから (“bogus-pipeline effect”)
- フィードバックに伴い、十分な注意・推論を行うようになったから
- 正直な回答が個人合理的であると判断したから

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.

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.