

質問紙調査回答に表れるバイアス -係留ビネット法と項目反応モデルによる検討-

北條大樹

東京大学大学院 教育学研究科 D1 / 日本学術振興会



HP: <https://dastatis.github.io/index.html>

Twitter: @dastatis

- **名前:** 北條大樹 (ほうじょう だいき)
- **指導教員:** 岡田謙介
- **専門:** 心理統計学・ベイズ統計学
- **大きな興味:**
 - 統計学で、人間の行動を明らかにしたい。モデル化したい
 - 心理学で(統計学的に)良いモデルとはどういう意味なのか

- **第1部 調査回答に現れるバイアス**
 - 調査はなぜむずかしいのか？
 - 調査回答に現れるバイアス
- **第2部 Response Styleとそのモデル**
- **第3部 実践 係留ビネット調査(時間余れば)**
- **総合議論**
- **質問・コメントは適宜入れてもらえると嬉しいです**
- **各部ごとに休憩挟みます**

調査はなぜ難しいのか？

- 調査の特徴(研究者視点)

- 環境操作なし
- 観察なし
- 主観的な出来事(感情・動機・態度)を扱える
- 用意した問いへの回答を得る
- 相関関係を探る

「信用に足る研究法として用いて適切なデータを収集するには、適切な手続きを踏むことが必要」

(三浦, 2017. 第3章 調査, 心理調査の基礎, 有斐閣)

- なぜ調査は難しいのか?

- **様々なバイアス**がかかるから
- もう少し掘り下げる⇒次スライド



- 以下のように定義(あくまで1表現として)
 - θ : 主観的な出来事(感情・動機・態度)・構成概念・特性
 - ϵ : 誤差(後述)
 - $f()$: 関数 (関数: 内部が規則的な振る舞いをして結果を排出)
 - Y : 得られるもの(ex. 項目回答データ)
- 調査は以下のように表現できるのではないだろうか？

$$Y = f(\theta) + \epsilon$$

- $f()$ は、特性についての関数 (特性の表出を直接変える)
- ϵ は、調査そのものを歪めてしまう誤差(バイアス)
- これによって、「調査の難しさ」を考えてみると
 - 注) ここでの難しさ \neq 問題/項目の難しさ

- Social Desirability (社会的望ましさ)
 - 社会的に良いとされる方向に自己回答を歪めてしまう現象

- 数式に戻ってみると

$$Y = f(\theta) + \epsilon$$

- 社会的望ましさを決定しているのは、 $f()$ の部分
 - 社会的に良い (ex. 常識) という規則によって特性が表出
 - **社会的望ましさフィルターにかかり、回答が歪む**

- Satisficing (努力の最小限化もしくは満足化)
 - “目的を達成するために必要最小限を満たす手順を決定し、追求する行動”(三浦・小林, 2015, 社心研)
 - **“教示文や尺度項目を読まずに回答する行動”**
(三浦・小林, 2018, 行動計量学)
- 数式に戻ってみると

$$Y = f(\theta) + \epsilon$$

- Satisficeは、 ϵ の部分と考えられる
 - そもそも特性を表出しようとしていない(表出が雑)
 - **それ以上に ϵ が大きくなり、回答が歪む**

• 調査の構造

$$Y = f(\theta) + \epsilon$$

• 調査が難しい理由

- 特性が何らかの関数 $f()$ (フィルター)によって表出が歪む
- 特性とは全く別の要因(ϵ)によって歪む
- 関数・要因(誤差)が調査内で未知もしくは特定されていない
- そして、得られるのは Y (**項目回答**)だけ

だから、調査を“**信用に足る研究法**”にするのは大変

調査回答に現れるバイアス

偏り

出典: フリー百科事典『ウィキペディア (Wikipedia)』

偏り（かたより）、または**バイアス**という用語は、**統計学**で2つの異なる意味に用いられる。

1. 標本の偏りとは、母集団の要素が標本として平等に選ばれていないと考えられる場合をいう。
2. 推定量の偏りとは、推定すべき量を何らかの理由で高く、または低く推定しすぎている場合をいう。

偏りという用語は悪い意味に聞こえるが、必ずしもそうではない。偏った標本は悪いものだが、偏った推定量のよしあしは状況による。

- **調査では、両方ともありえる**

- ただ、1 は基本的に考慮されることが多いので、今回は略

- **2.について**

- 何らかの理由 = ex.) 社会的望ましさ、努力の最小限化

- **バイアス ≠ 悪いことが原因**

- 無意識的に働くバイアスも存在

偏り

出典: フリー百科事典『ウィキペディア (Wikipedia)』

偏り（かたより）、または**バイアス**という用語は、**統計学**で2つの異なる意味に用いられる。

1. 標本の偏りとは、母集団の要素が標本として平等に選ばれていないと考えられる場合をいう。
2. 推定量の偏りとは、推定すべき量を何らかの理由で高く、または低く推定しすぎている場合をいう。

偏りという用語は悪い意味に聞こえるが、必ずしもそうではない。偏った標本は悪いものだが、偏った推定量のよしあしは状況による。

・調査では、両方ともありえる

- ただ、1 は基本的に考慮されることが多いので、今回は略

・2.について

- 何らかの理由 = ex.) 社会的望ましさ、努力の最小限化

・バイアス ≠ 悪いことが原因

- 無意識的に働くバイアスも存在

→ **どんな調査バイアスがあるだろうか？**

意図的(故意)
(調査協力的)

- Response Set  • Social Desirability
- Faking Good / Bad

意図的(故意)
(調査**非**協力的)

- Satisficing

非意図的(無意識)

- Method Bias (狭義)  • Reversed Item
- Response Style (狭義) • Item Wording
- Careless Responding


注) これ以上にもありますし、
別の呼び名、別のカテゴリ分け、
必ずしも(非)意図的とならない場合も
各バイアスが独立でない場合もある

個別に見ていく → 次スライド

• 特徴

- **項目内容に応じて**、自分の特定面だけ見せるバイアス
- 本心を隠す(防衛)・仮病・偽造に積極的・積極的に賛成
- 意図的でなく、非意図的な場合もあり
- Response Style(広義)と呼ばれることも(後述)

• これが、右の2つにつながっていく

- Response Set  • Social Desirability
- Faking

参考 : Rorer (1965). The Great Response-Style Myth.

*Response Setの訳は、増田・坂上 (2014)より

・特徴

- **社会的に良い(正解)**とされる方向・規範に自己回答を歪める
- 昔は、Social Acquiescence(社会的黙従?)と呼ばれていたことも
- 社会的なルールに則り、正しい選択肢へ
- 当然、**項目内容に大きく依存**する (ex. 就活テスト)
- これらのバイアス系で最も有名？
- 統制のために、**日本語版バランス型社会的望ましさ尺度** (谷, 2008)やLie scale等*がある

*これらが最も良い・悪いという意味ではない

・特徴

- **自分を良く(悪く)見せる**ように、回答を歪ませるバイアス
- 社会的望ましさとしてまとめられることも
- 項目内容に依存する
- **良く見せる例**
 - 職業適性テスト ⇒ 職に就きやすくなる
 - 性格診断(MMPI) (Graham et al., 1996などで検討)
- **悪く見せる例**
 - 心理検査 ⇒ 疾患がもらえると助成が受けられるから
 - 軍人適正検査 ⇒ 軍に入らなくて済むように
- 統制のために、MMPIではF scaleが同梱されている
- また、IRTによるアプローチも提案(Usami et al, 2016)

• 特徴

- “**教示文や尺度項目を読まずに回答する行動**” (三浦・小林, 2018)
- 上記のように適当に (DK選択肢に) 答えるバイアス
- 他に調査項目を理解し回答しようとするが、部分的な選択肢のみに注目して回答するSatisficeも存在する
- 前者をstrong satisfice, 後者をweak (or optimizing) satisfice (三浦・小林 2015; Krosnick, 1991)
- House Effects*との関係も指摘

意思決定に関する近年の研究で、人間の決定は「真空」状態でおこなわれるわけではないことがわかってきました。人が何かを決めるとき、その人の好みや知識、または、そのときどんな状況におかれているかなどのさまざまな特徴が、大きな影響を及ぼすのです。この調査では、こうした「人間の決め方」を研究するために、あなたの「意思決定者」としてのある特徴を知りたいと考えています。つまり、あなたがこの指示を時間をかけてよく読んでいるかどうかを知りたいのです。もし誰もこの問題文をお読みになっていなければ、問題文の内容を変えることが「人間の決め方」に与える影響を見たい、というわれわれの試みは意味を持たないからです。ここからがお願いします。**この指示をお読みになったことの証明として、以下のスポーツ競技に関する項目は全て無視してください。その代わり、「その他」にチ**

<https://kgpsysci.au1.qualtrics.com/ControlPanel/Ajax.php?action=GetSurveyPrintPreview>

6/18

2017/10/29

Qualtrics Survey Software

エックを入れた上で、自由記述欄に 読んだ と入力し、次に進んでください。よろしくお願いします。

普段よく行う活動は以下のうちどれですか。当てはまるもの全てにチェックを入れてください。

- ☐ スキー
- ☐ 水泳
- ☐ サッカー
- ☐ テニス
- ☐ スノーボード
- ☐ バスケットボール
- ☐ ジョギング
- ☐ サイクリング
- ☐ 卓球
- ☐ その他

三浦・小林 (2018)で実際に使われたもの
オンライン付録より引用

上記は次より引用: <https://mfr.osf.io/render?url=https://osf.io/v56ye/?action=download%26mode=render>


意図的(故意)
(調査協力的)

- Response Set  • Social Desirability
- Faking Good / Bad

意図的(故意)
(調査非協力的)

- Satisficing


非意図的(無意識)

- Method Bias (狭義)  • Reversed Item
- Response Style (狭義) • Item Wording
- Careless Responding

注) これ以上にもありますし、
別の呼び名、別のカテゴリ分け、
必ずしも(非)意図的とならない場合も
各バイアスが独立でない場合もある

•特徴

- 回答者というよりは、**特定の状況や方法論によるバイアス**

- Method Bias (狭義) 
- Reversed Item
- Item Wording

•Reversed Item Effect

- 順向項目のみだと、**黙従バイアス**(Yea-saying)がかかる
⇒順向と逆向(逆転)項目をバランスよく入れたら良い
- **逆向項目入れると、因子構造悪くなるとの報告も**(二瓶 他, 2018)

•Item Wording

- 項目内容に依存
- **曖昧な表現 / 専門性が高い / 教育歴等も関係**
- 対処: 曖昧な表現を避ける(基本中の基本ですが)

•あとは...

- ある調査方法で収集したときのみ表れるバイアス

TABLE 1

Frequencies of College Student Samples, Research Designs, and Measurement Methods by Content Area in Psychological Research Studies

	Psychological Content Area					
	Applied	CH	Develop.	CBB	SP	Totals
Human Samples						
College	8 (26%)	3 (5%)	3 (27%)	12 (39%)	16 (50%)	42 (25%)
Noncollege	23 (74%)	61 (95%)	8 (73%)	19 (61%)	16 (50%)	127 (75%)
Totals	31 (18%)	64 (38%)	11 (7%)	31 (18%)	32 (19%)	169
Research Design						
Experiment	5 (15%)	14 (22%)	4 (36%)	41 (79%)	15 (47%)	79 (41%)
Cross-Sec.	7 (21%)	27 (42%)	4 (36%)	11 (21%)	14 (44%)	63 (33%)
Longitudinal	1 (3%)	10 (16%)	2 (18%)	0 (0%)	1 (3%)	14 (7%)
Qualitative	12 (36%)	3 (5%)	1 (9%)	0 (0%)	0 (0%)	16 (8%)
Survey	4 (12%)	5 (8%)	0 (0%)	0 (0%)	2 (6%)	11 (6%)
Archival	4 (12%)	5 (8%)	0 (0%)	0 (0%)	0 (0%)	9 (5%)
Totals	33 (17%)	74 (33%)	11 (6%)	52 (27%)	32 (17%)	192
Measurement						
SR Question.	14 (45%)	33 (52%)	3 (27%)	7 (13%)	28 (88%)	85 (45%)
Interview	15 (48%)	38 (59%)	4 (36%)	2 (4%)	3 (9%)	62 (33%)
Observ.	6 (19%)	4 (6%)	3 (27%)	5 (10%)	2 (6%)	20 (11%)
Behavioural	1 (3%)	0 (0%)	5 (45%)	30 (58%)	8 (25%)	44 (23%)
Physio.	1 (3%)	8 (13%)	1 (9%)	26 (50%)	0 (0%)	36 (19%)
No. of Studies	31 (16%)	64 (34%)	11 (6%)	52 (27%)	32 (17%)	190

Notes. Numbers in parentheses are percentages rounded to whole units. Percentages in the Totals rows and columns represent percentage of total sample; percentages in the interior represent percentage of the studies within a given content area. Column sums of frequencies (percentages) in the third panel on measurement exceed the number of studies (100%) as some studies used more than a single measurement method.

CH = Clinical/Health, Develop. = Developmental, CBB = Cognition, Brain, and Behaviour, SP = Social/Personality, Cross-Sec. = Cross-sectional, SR Question = Self-report questionnaire, Observ. = Observation, Physio. = Physiological measurement.

- ある調査方法で収集したときのみ表れるバイアス
 - 対処: 単一方法でなく、複数の方法を用いることで軽減
 - たとえば、
 - Likart法とSD法は変換しやすいので、両方やると方法による効果を軽減できるとの報告(Podsakoff et al., 2011内で複数紹介)

アイスクリームは朝食に適している

1. 全く同意できない
2. 同意できない
3. どちらともいえない
4. 同意できる
5. 非常に同意できる

リッカート法イメージ

ウィキペディアより引用

<https://ja.wikipedia.org/wiki/%E3%83%AA%E3%83%82%AB%E3%83%BC%E3%83%88%E5%B0%BA%E5%BA%A6>

きれい	— — — — — —	きたない
遅い	— — — — — —	速い※
かたい	— — — — — —	やわらかい
悪い	— — — — — —	良い※
動的な	— — — — — —	静的な
弱い	— — — — — —	強い※

※左右逆転項目

図-2 SD 法の項目例

SD法イメージ

谷口(2017) 図-2より引用

•特徴

- **不注意による回答バイアス**
- 意図的な場合は、弱いsatisficeとして考えることもできる?
- Carelessになってしまう(e.g., 注意資源をものすごい割く)場合は、**方法(項目)に問題ある可能性も**
- **欠損として表れる可能性**もあるので、規則的な欠損が表れたらCareless Responding になっている可能性を疑う必要
- 対処: **こちらから提供するもの(教示/項目)が万全か確かめる**

・特徴

- **項目内容に関係なく (独立に)、反応パターンによって回答するバイアス**
- たとえば、**国ごとで極端に答えやすい・にくい等**
- 意図的でなく、非意図的な場合もあり
- **Response Set(反応傾向)との違いは、項目内容に依るか否か**
- **Response Setが、“True (Social Acquiescence)”なものに反応するのに対し、Response Styleは、“Preference”なものに反応**
- **ただ、近年のResponse Style研究はResponse Setの意味を含んで使われているのがほとんど**
- 今回の講演後半のメイン


意図的(故意)
(調査協力的)

- Response Set  • Social Desirability
- Faking Good / Bad

意図的(故意)
(調査**非**協力的)

- Satisficing


非意図的(無意識)

- Method Bias (狭義)  • Reversed Item
- Response Style (狭義) • Item Wording
- Careless Responding

注) これ以上にもありますし、
別の呼び名、別のカテゴリ分け、
必ずしも(非)意図的とならない場合も
各バイアスが独立でない場合もある

- 調査を信用たる方法にするのは難しい
- その一つの原因が、**様々な反応バイアス**による影響


意図的(故意)
(調査協力的)

- Response Set 
 - Social Desirability
 - Faking Good / Bad

意図的(故意)
(調査非**協力的)**

- Satisficing

非意図的(無意識)

- Method Bias (狭義) 
 - Reversed Item
 - Item Wording
- Response Style (狭義)
- Careless Responding

- 回答者によって歪む、方法論によって歪む
- 様々な歪み方があって、各種知っておくだけでも違うかも

- **広義のResponse Styleは、**
 - 社会的望ましさを始めとする**項目依存による影響、**
 - 項目とは独立な**人口動態変数による影響、**
 - 上記のような**回答者の直接的な要因、**
 - 方法論として**間接的に影響する要因****の包括的なバイアスと捉えられる**
- **第2部では、これを様々な側面から掘り下げて行く**

- Bodner, T. E. (2006). Designs, participants, and measurement methods in psychological research. *Canadian Psychology/Psychologie canadienne*, 47(4), 263.
- Graham, J. R., Watts, D., & Timbrook, R. E. (1991). Detecting Fake-Good and Fake-Bad MMPI-2 Profiles. *Journal of Personality Assessment*, 57(2), 264–277. https://doi.org/10.1207/s15327752jpa5702_6
- Krosnick, J. O. N. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, 5, 213–236.
- 増田真也, & 坂上貴之. (2014). 調査の回答における中間選択 —— 原因, 影響とその対策 ——. *Japanese Psychological Review*, 57(4), 472–494.
- 三浦麻子, & 小林哲郎. (2015). オンライン調査モニタのSatisficeに関する実験的研究. *社会心理学研究*, 31(1), 1–12. https://doi.org/10.14966/jssp.31.1_1
- 三浦麻子, & 小林哲郎. (2018). オンライン調査における努力の最小限化が回答行動に及ぼす影響. *行動計量学*, 45, 1–11.
- 三浦 麻子. (2017). 第3章 調査, In サトウタツヤ・鈴木直人, *心理調査の基礎*, 有斐閣
- 二瓶 正登, 荒井 穂菜美, 前田 香, 青木 俊太郎, 土屋垣内 晶, 岩野 卓, 富岡 奈津代, 岡村 尚昌, 三原 健吾, 城月 健太郎, 堀内 聡, 坂野 雄二. (2018) Fear of Negative Evaluation Scale日本語短縮版の因子構造, 信頼性および妥当性の再検討, *不安症研究*, 10 巻, 1 号, p. 54-63, https://www.jstage.jst.go.jp/article/jsad/10/1/10_54/article/-char/ja
- Podsakoff, P. M., Mackenzie, S. B., & Podsakoff, N. P. (2011). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. <https://doi.org/10.1146/annurev-psych-120710-100452>
- Rorer, L. G. (1965). The great response-style myth. *Psychological Bulletin*, 63(3), 129–156. <https://doi.org/10.1037/h0021888>
- 谷口 高士, 心理評価実験における尺度構成の方法, *日本音響学会誌*, 2017, 73 巻, 12 号, p. 774-782, https://www.jstage.jst.go.jp/article/jasj/73/12/73_774/article/-char/ja
- 谷伊織. (2008). バランス型社会的望ましさ反応尺度日本語版 (BIDR-J) の作成と信頼性・妥当性の検討. *パーソナリティ研究*, 17(1), 18–28.
- Usami, S., Sakamoto, A., Naito, J., & Abe, Y. (2016). International Journal of Testing Developing Pairwise Preference-Based Personality Test and Experimental Investigation of Its Resistance to Faking Effect by Item Response Model. <https://doi.org/10.1080/15305058.2016.1145123>

第2部 Response Styleとそのモデル

- RS研究の歴史と動向
- RS(Response Style)の種類
- RSと様々な変数との関係性
- RSへの対処アプローチ
 - 方法論的アプローチ
 - 統計学的アプローチ
- RSについての3つの疑問
 - **Psychometric level** – RSを測定するには?
 - **Substantive level** – 何でRSが説明できる?何に影響される?
 - **Applied level** – RS除去したい(興味があるのは特性だけ)
- RSに対処するモデル
 - **Item Response Model**
 - **Item Response Tree Model**

RESPONSE SETS AND TEST VALIDITY

LEE J. CRONBACH

University of Chicago

A PSYCHOLOGICAL or educational test is constructed by choosing items of the desired content, and refining them by empirical techniques. The assumption is generally made, and validated as well as possible, that what the test measures is determined by the content of the items. Yet the final score of the person on any test is a composite of effects resulting from the content of the item and effects resulting from the form of item used. A test supposedly measuring one variable may also be measuring another trait which would not influence the score if another type of item were used. This paper attempts to demonstrate these influences in a variety of tests, to examine the effect of these extraneous factors, and to suggest means of controlling them.

- RS(論文内はResponse Set)の定義

Tendencies such as those just described are characterized in this paper as “response sets.” A response set is defined as any tendency causing a person consistently to give different responses to test items than he would when the same content is presented in a different form. This is a theoretical rather than a practical definition, since it is never possible wholly to separate the content of an item from its form. Yet many acquiescent

Response Setとは、

一貫して、テスト項目に対する回答と異なる反応をする傾向

- これは、実務上の定義というより理論的な定義だ
- なぜなら項目内容を形式から分離することはほぼ不可能だから

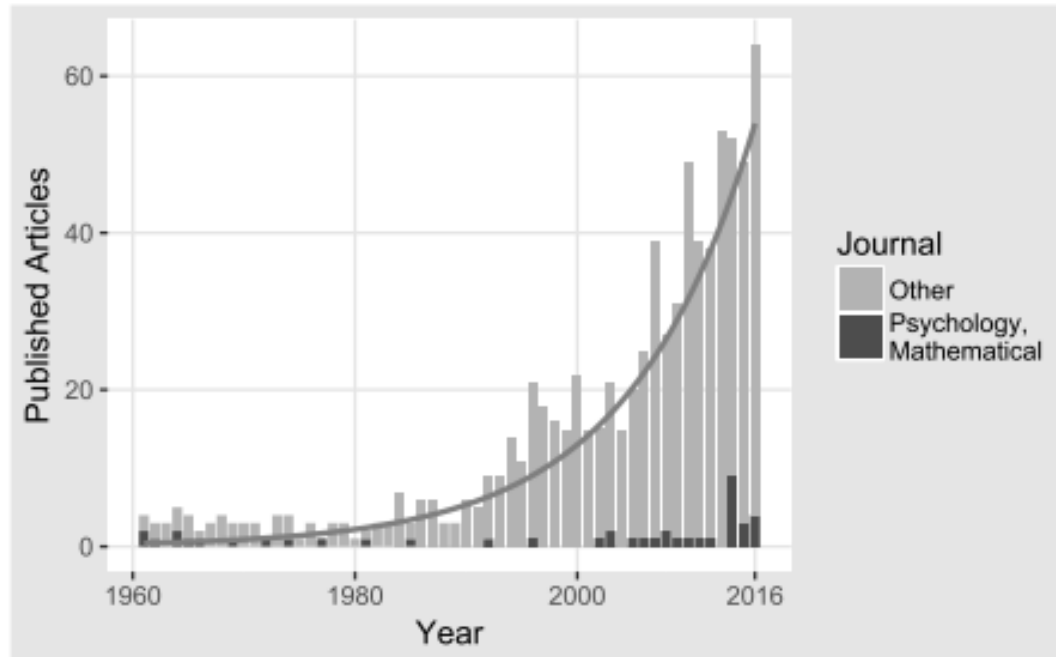


FIGURE 1.1: Bar chart of published journal articles on response styles per year and exponential growth curve.

- Web of Scienceにおいて「response style」で検索
- **心理学以外の分野でも急増している**

Plieninger, H. (2017). Figure 1.1より引用

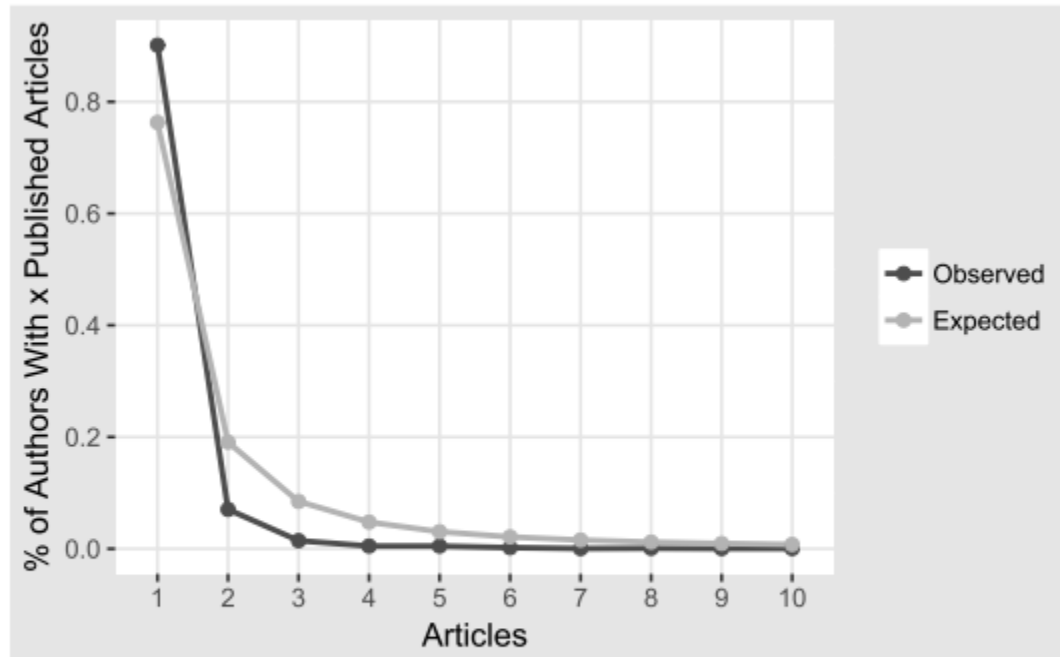
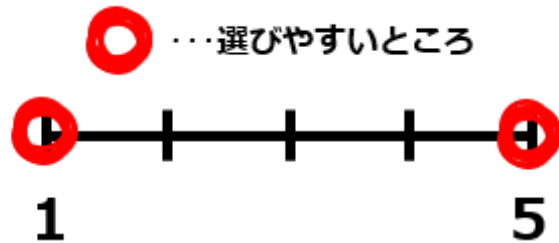


FIGURE 1.2: The number of authors publishing x response style articles deviates from what would be expected under Lotka's law.

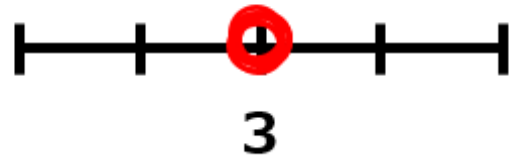
- その一方で、RS研究の**著者一人に対する論文数がかなり少ない**
- 1本だしておしまいな人が多い

Plieninger, H. (2017). Figure 1.2より引用



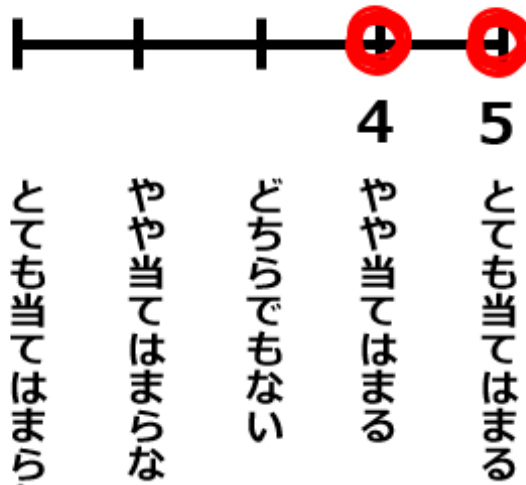
極端反応スタイル (Extreme RS)

⇒ 1, 5(極端なカテゴリ)を回答しやすい



中間反応スタイル (Midpoint RS)

⇒ 3 (中間のカテゴリ)を回答しやすい



黙従反応スタイル (Acquiescence RS)

⇒ 4, 5 (ポジティブなカテゴリ)を回答しやすい

RSの2つの源 (Weijters, 2006; Vaerenbergh et al., 2012) 36

- **Stimulus level**
 - 項目もしくは、他の課題に関連する要因
- **Respondent level**
 - 回答者やその特性要因

- **Stimulus level**
 - 項目もしくは、他の課題に関連する要因
- **Respondent level**
 - 回答者やその特性要因
- さきほどの様々な反応バイアスの話へつながる

- **広義のResponse Styleは、**
 - 社会的望ましさを始めとする**項目依存による影響、**
 - 項目とは独立な**人口動態変数による影響、**
 - 上記のような**回答者の直接的な要因、**
 - 方法論として**間接的に影響する要因**

の包括的なバイアスと捉えられる

• 尺度フォーマット

- 件法が大きくなるとERSが低下、MRSが増加
- フルラベルの調査ではERS低下、むしろ極端なラベルのみだとERS増加

• データ収集方法

- 電話は、紙やWebよりもARSが増加
- 紙よりもWebのほうがERS増加（逆の研究もある）

• 認知的負荷

- 認知的負荷高いと、ARS・NARSが増加

• 調査言語

- 母国語⇒ERS増加, 第二言語⇒MRS増加

• 話題関連度

- 関連度高⇒ERS増加

•教育

- ほぼ例外なく、教育とRSは逆の関係性
- 教育歴が浅い⇒RS出やすい

•年齢

- 関係ありそうだが、様々な報告
- ARS・ERSは、年齢と線形(非線形)の関係が報告
- 非線形⇒若い時ERS強いけど、一度落ち着いて、年取ってまたERS高くなる(Schneider, 2018)

•性別

- いくつかの研究では、女性の方がARS高いという報告
- 性差はないという報告も

•他にも

- 収入や人種と関連との報告も...

・パーソナリティ (心理関連変数)

- RSはStability(縦断)/Consistency(横断)があるか？
- 縦断的に安定しているとの報告が多い(e.g., Okada, Hojo, Takahashi, revised)
- 横断的に安定とみるかは、RSの定義による？
- ERSは、BigFiveの外向性・誠実性と正の関係
- MRSとevasiveness (回避) に正の関係
- ARSは、衝動性と外向性と正の関係

・文化・国レベルの特徴

- 民主主義や政治的自由度とARS・ERSは関係無いとの報告
- (日本人の) 政治的態度(イデオロギー)のMRSは高い
(清水・稲増, in press)

- 日本人に政治的態度を10件法で尋ねてみると…
- MRSやばい

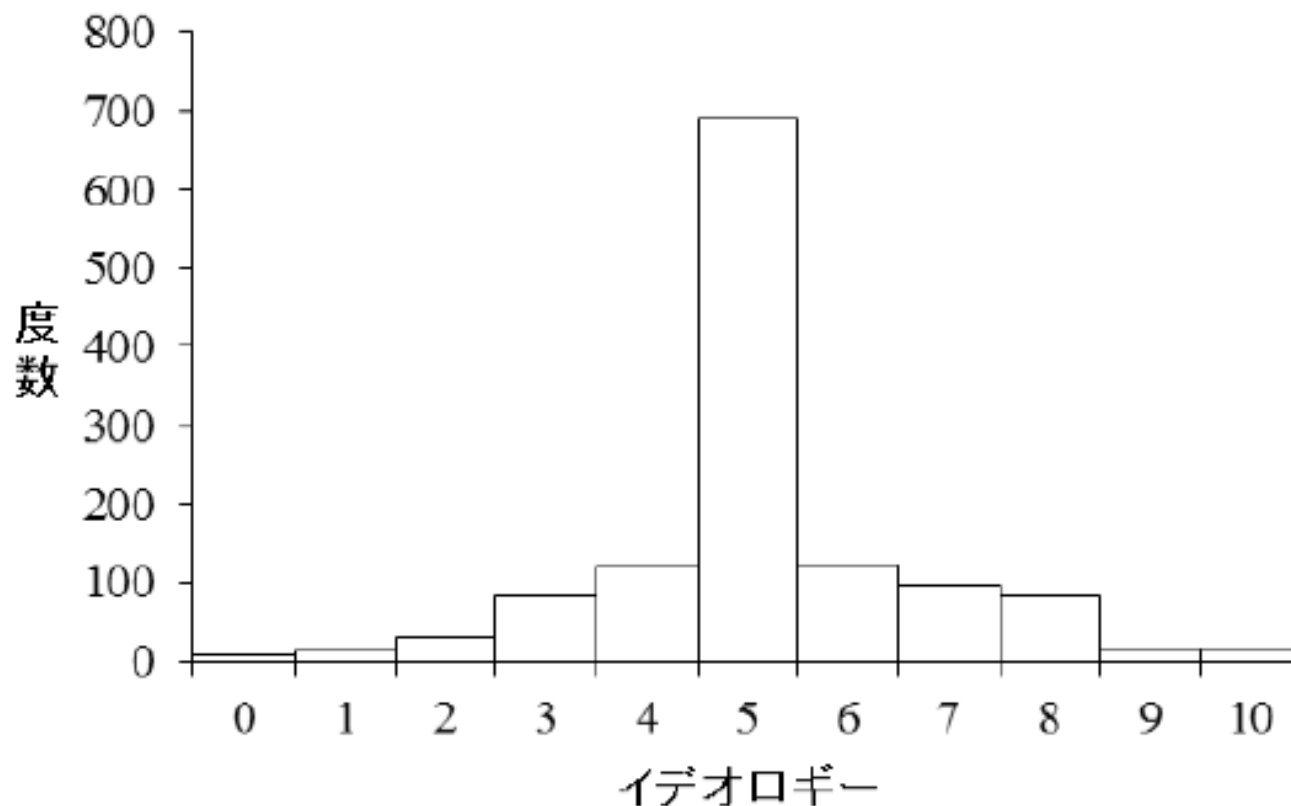
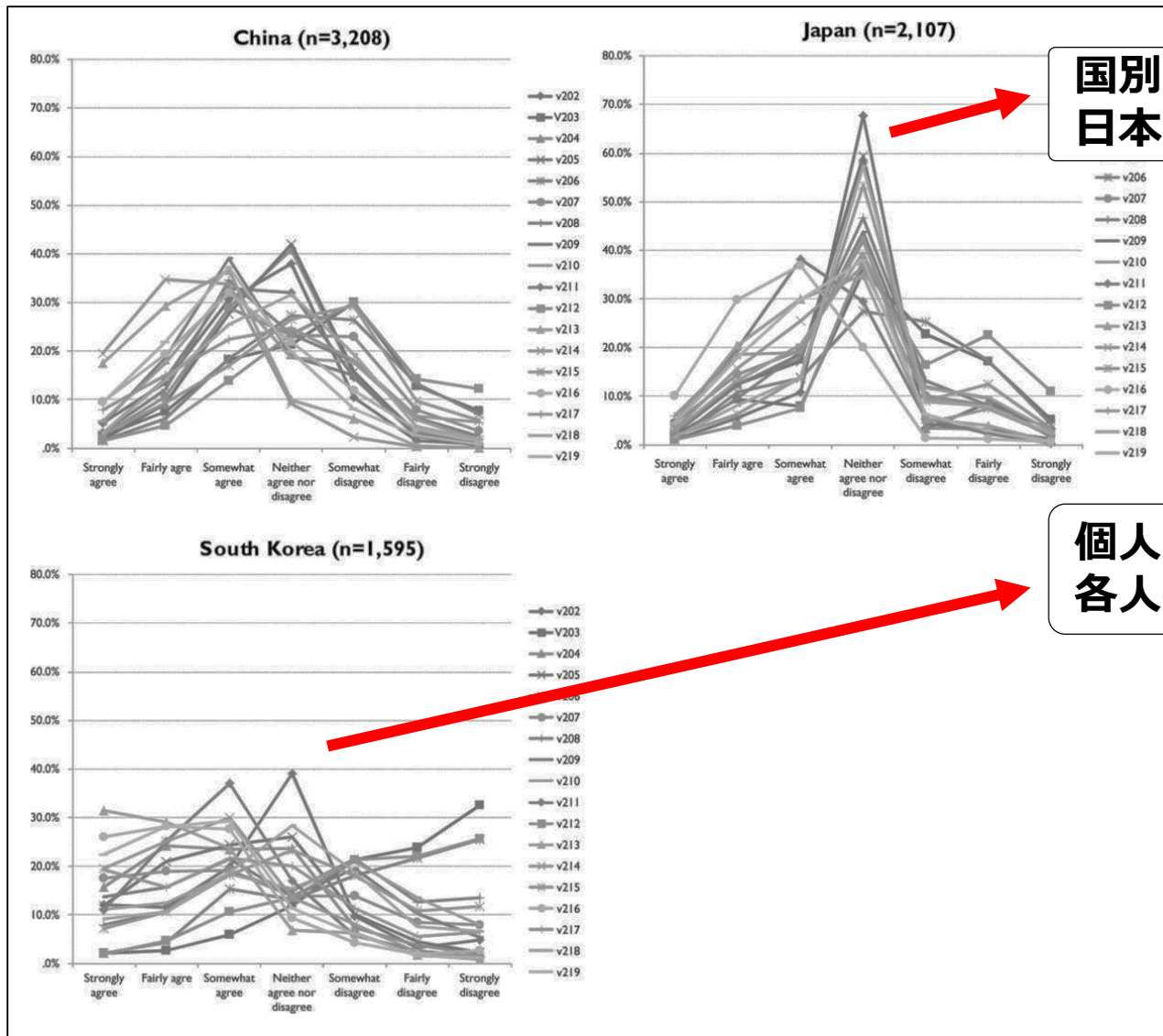


図1 東大朝日調査イデオロギー項目のヒストグラム

図は、各参加者の各評定カテゴリにおける反応頻度を表している(各線が個人を表す)



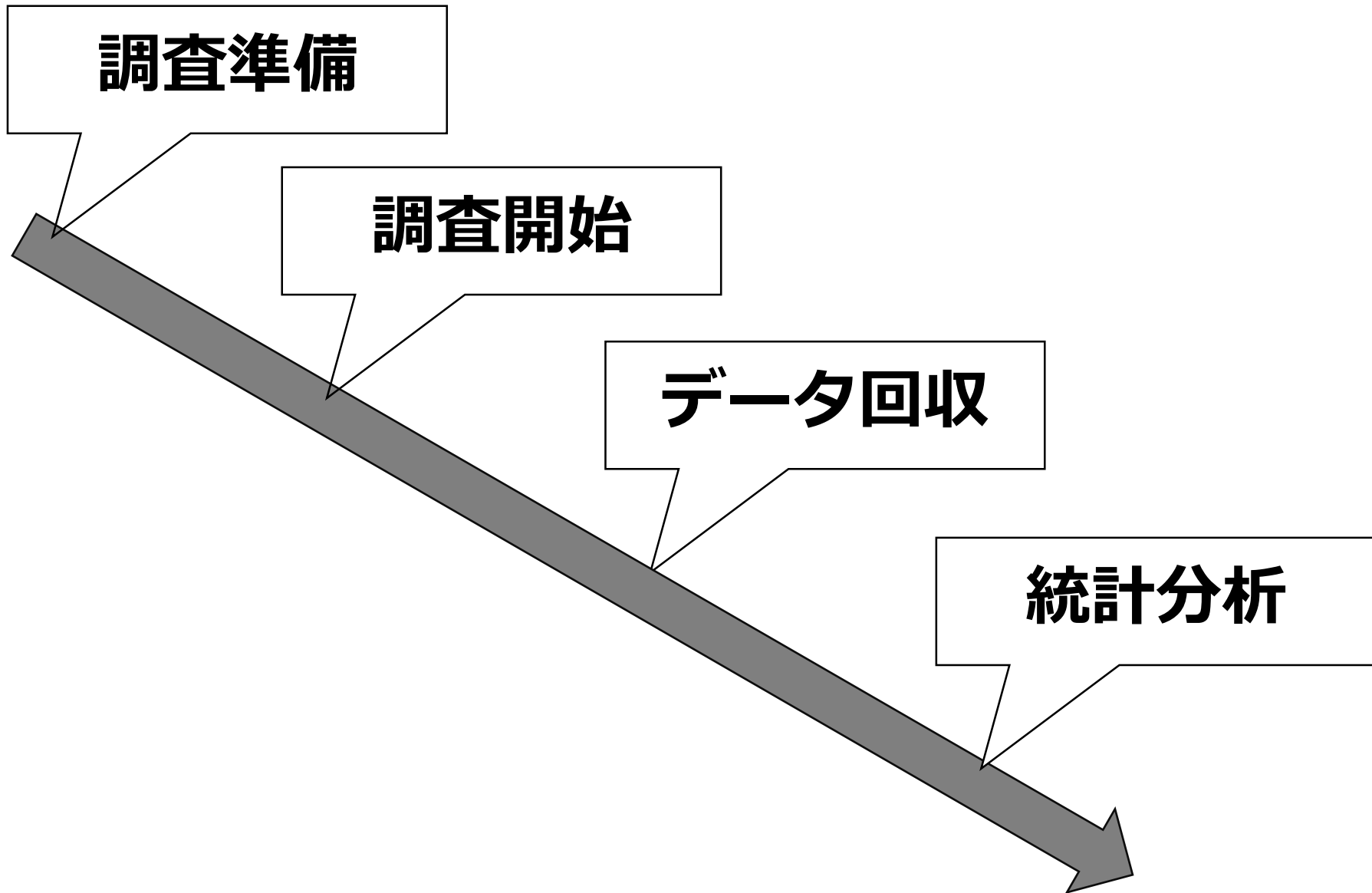
国別で見ると
日本では中間反応傾向が多い

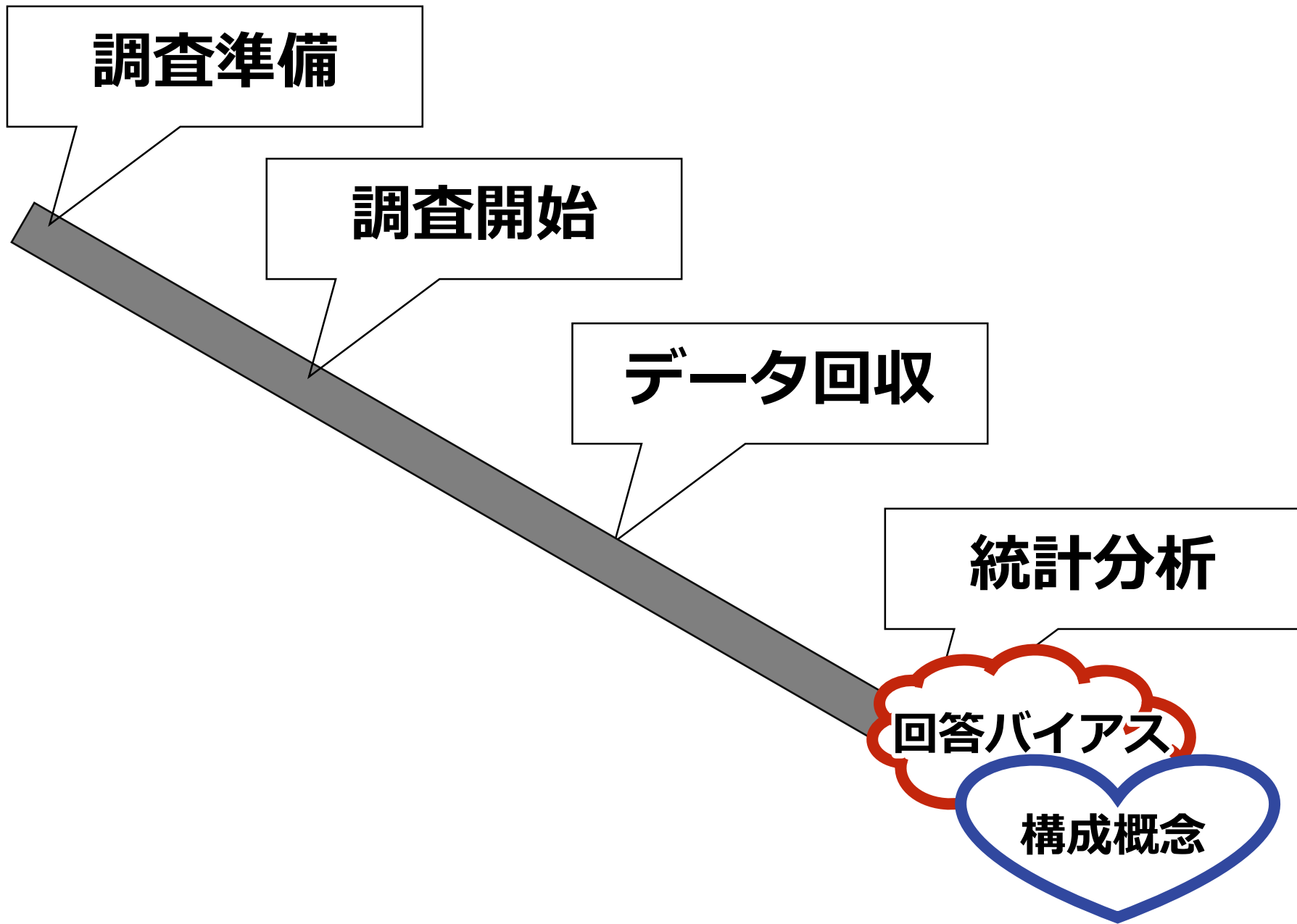
個人別に見ても
各人の反応傾向の特徴は様々

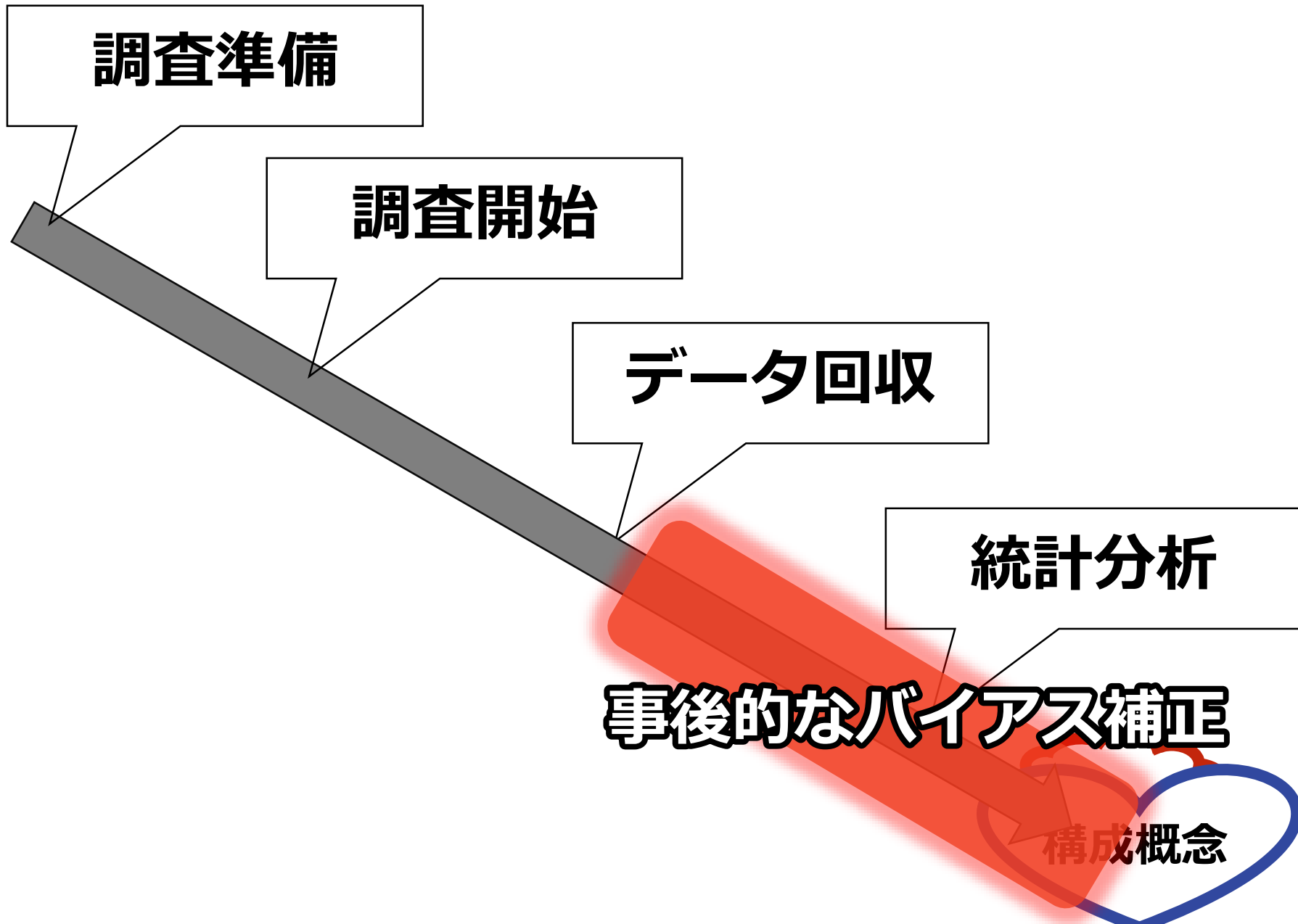
- RS研究の歴史と動向
- RS(Response Style)の種類
- RSと様々な変数との関係性
- RSへの対処アプローチ
 - 統計モデル的アプローチ
 - 方法論的アプローチ
- RSについての3つの疑問
 - **Psychometric level** – RSを測定するには?
 - **Substantive level** – 何でRSが説明できる?何に影響される?
 - **Applied level** – RS除去したい(興味があるのは特性だけ)
- RSに対処するモデル
 - **Item Response Model**
 - **Item Response Tree Model**

RSへの対処アプローチ

- 統計モデル的アプローチ







• 基本的にSEMや項目反応(IRT)モデル

- SEMでは、RIRSMACSモデル(Weijters et al. 2008)がメジャー
- RIRSMACS=(representative indicators response style means and covariance structure)

- たとえば、

- 1. 参加者をランダムに3群に分けて、各グループに異なる特性を測定する3つ尺度の項目をランダムに配る。
- 2. RSの数を下のように計算して、それもデータとする

$$ARS = [f(5) \times 1 + f(6) \times 2 + f(7) \times 3]/k, \quad (1)$$

$$DRS = [f(1) \times 3 + f(2) \times 2 + f(3) \times 1]/k, \quad (2)$$

$$ERS = [f(1) + f(7)]/k, \text{ and} \quad (3)$$

$$MRS = f(4)/k, \quad (4)$$

where $f(o)$ refers to the frequency of response option o .

Weijters et al. (2008)より

• 図のモデルを解く

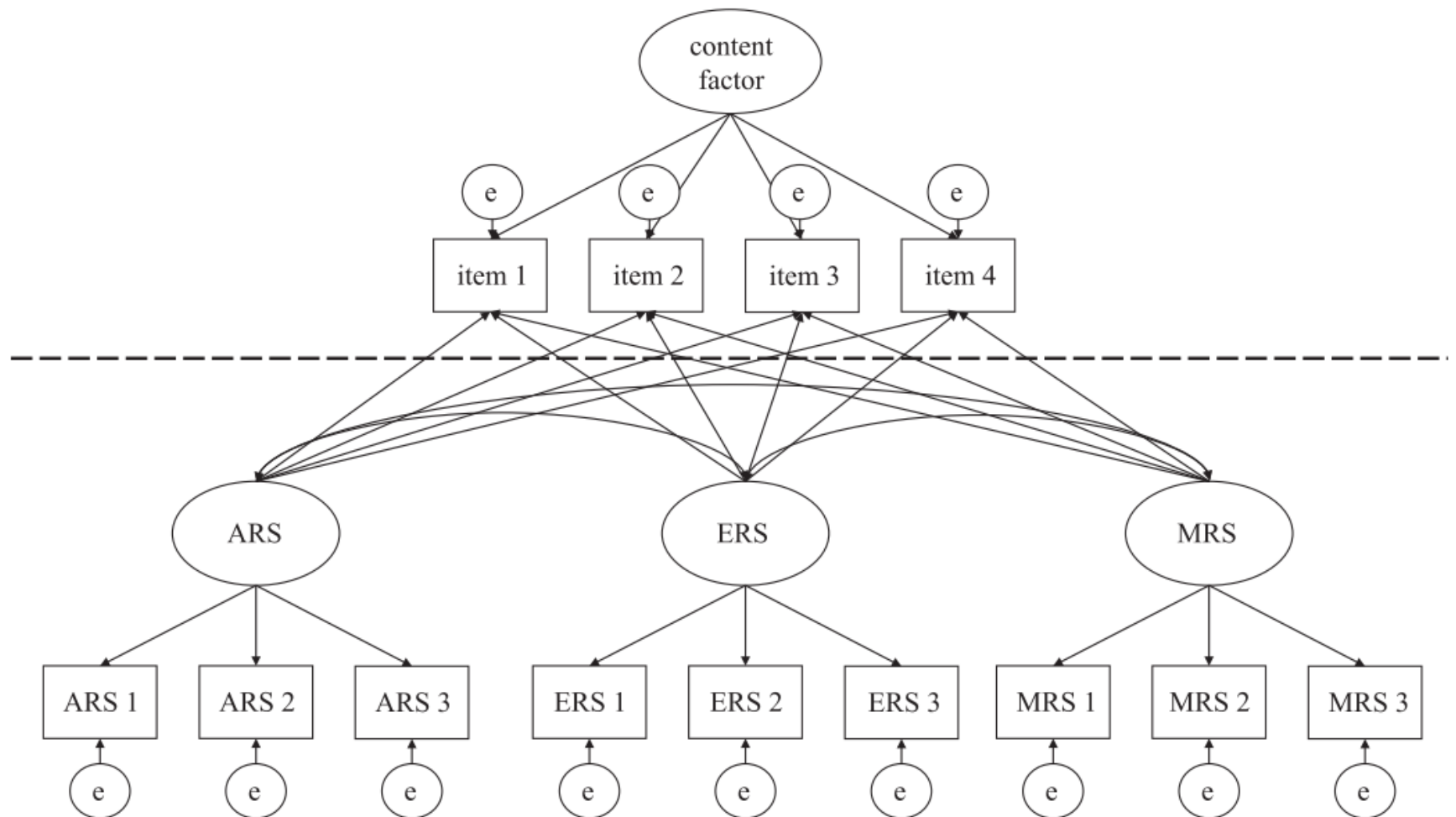


Figure 1. RIRSMACS モデル (Weijters et al., 2008)。

【復習】IRTモデル

必要なければ飛ばします

- 項目回答から**回答者の特性**と**項目の特徴**を推定可能な理論
- IRTを応用することで回答者集団や項目特性に依存しない回答者の能力を測ることが可能になる

- **特性 θ** の回答者が、項目 j に「はい」と答える確率を

$$P_j(\theta) = \frac{1}{1 + \exp(-D a_j (\theta - b_j))}$$

と表現する

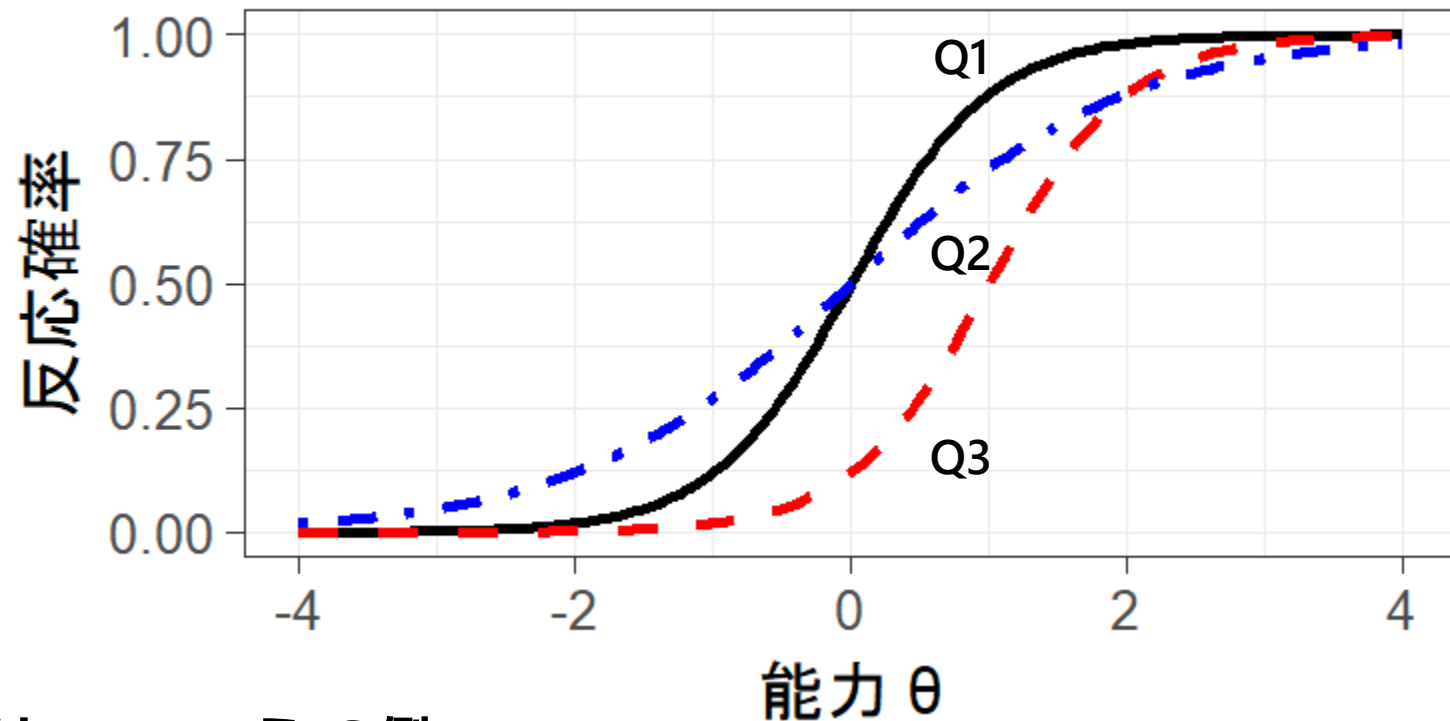
- a_j : 項目 j の識別力
- b_j : 項目 j の困難度
- θ : 回答者の特性
- D : 尺度因子(ここでは無視)

※2値(はい・いいえ)の例

ID	Q1	Q2	Q3
1	1	0	0
2	0	1	1
3	0	0	1

項目特性を推定し、
回答者の能力を推定

項目特性曲線(1つの線は1つの項目を表す)



※2値(はい・いいえ)の例

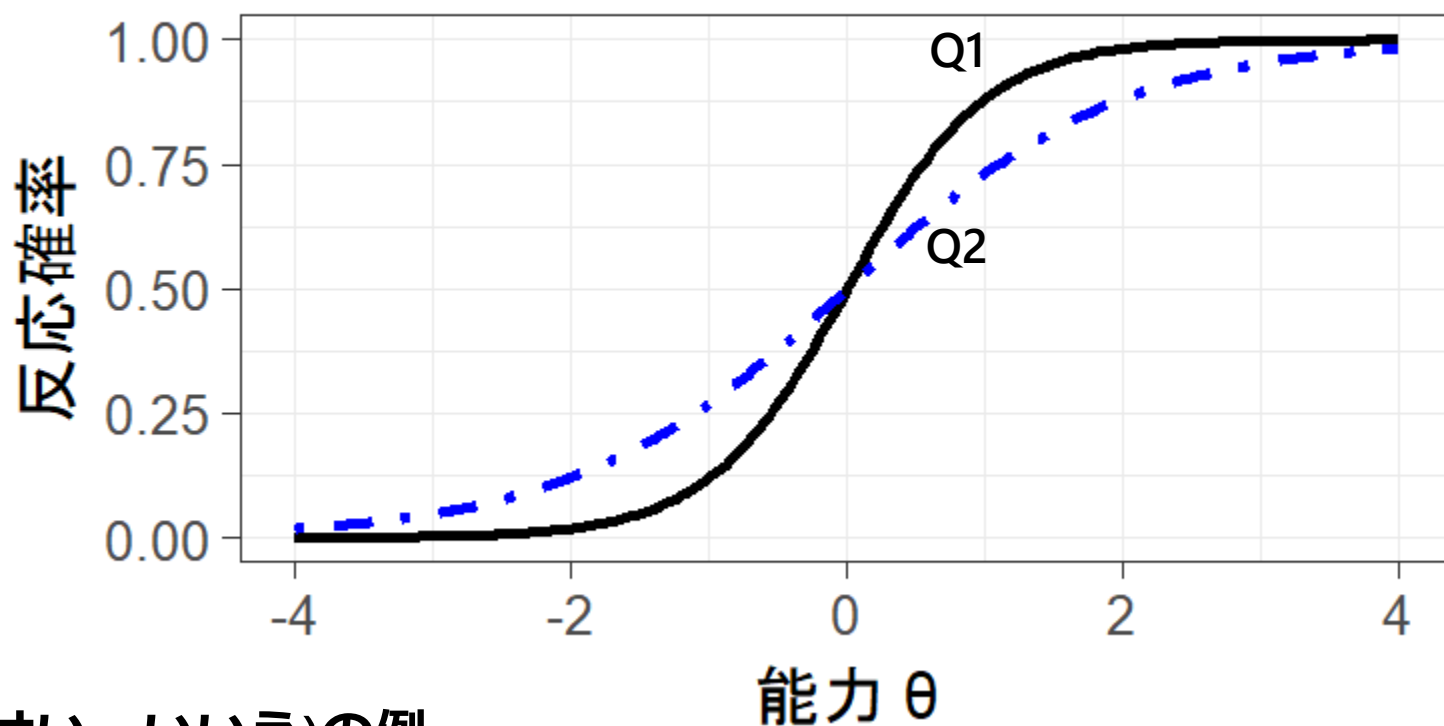
項目の何がわかるのか? (識別力)

53

ID	Q1	Q2	Q3
1	1	0	0
2	0	1	1
3	0	0	1

項目特性を推定し、
回答者の能力を推定

どれだけ 識別できる 項目かを検討できる(識別力)



※2値(はい・いいえ)の例

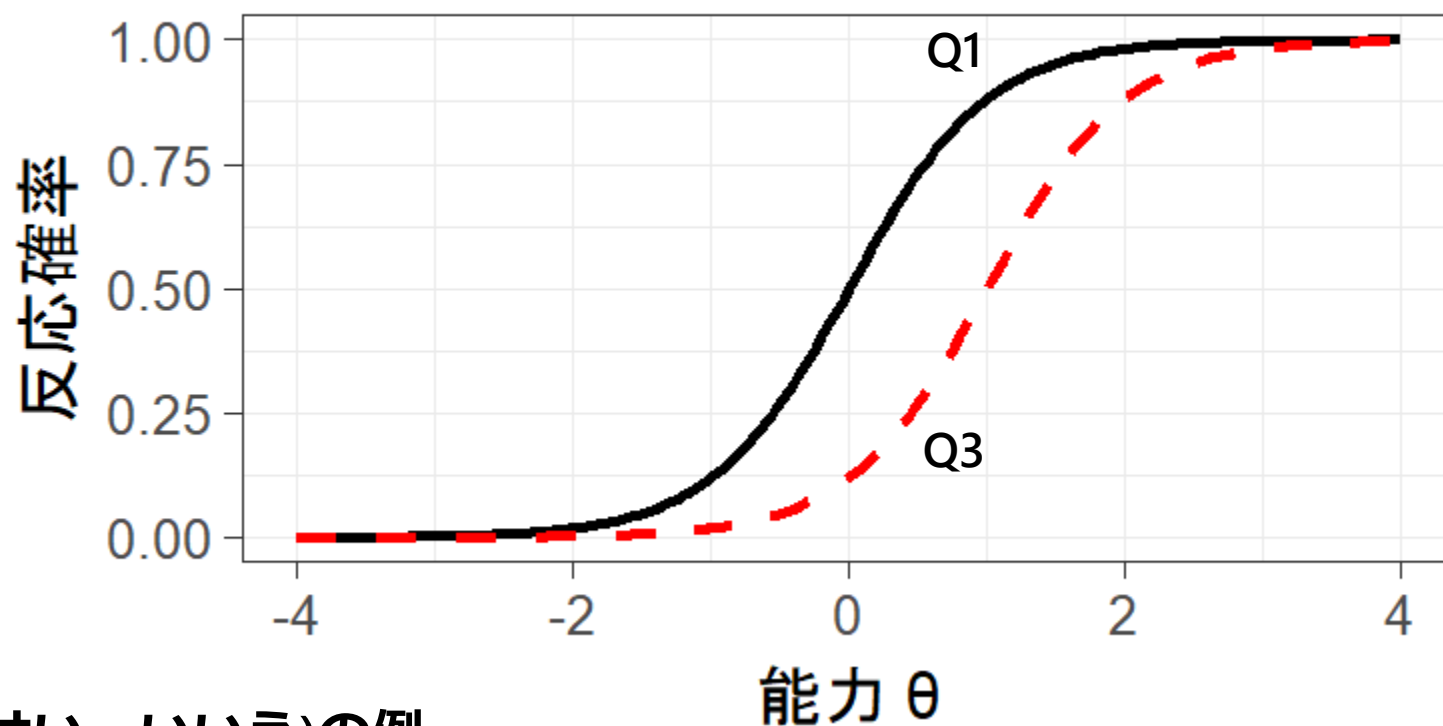
項目の何がわかるのか? (困難度)

54

ID	Q1	Q2	Q3
1	1	0	0
2	0	1	1
3	0	0	1

項目特性を推定し、
回答者の能力を推定

どれだけ 回答しやすい 項目かを検討できる(困難度)

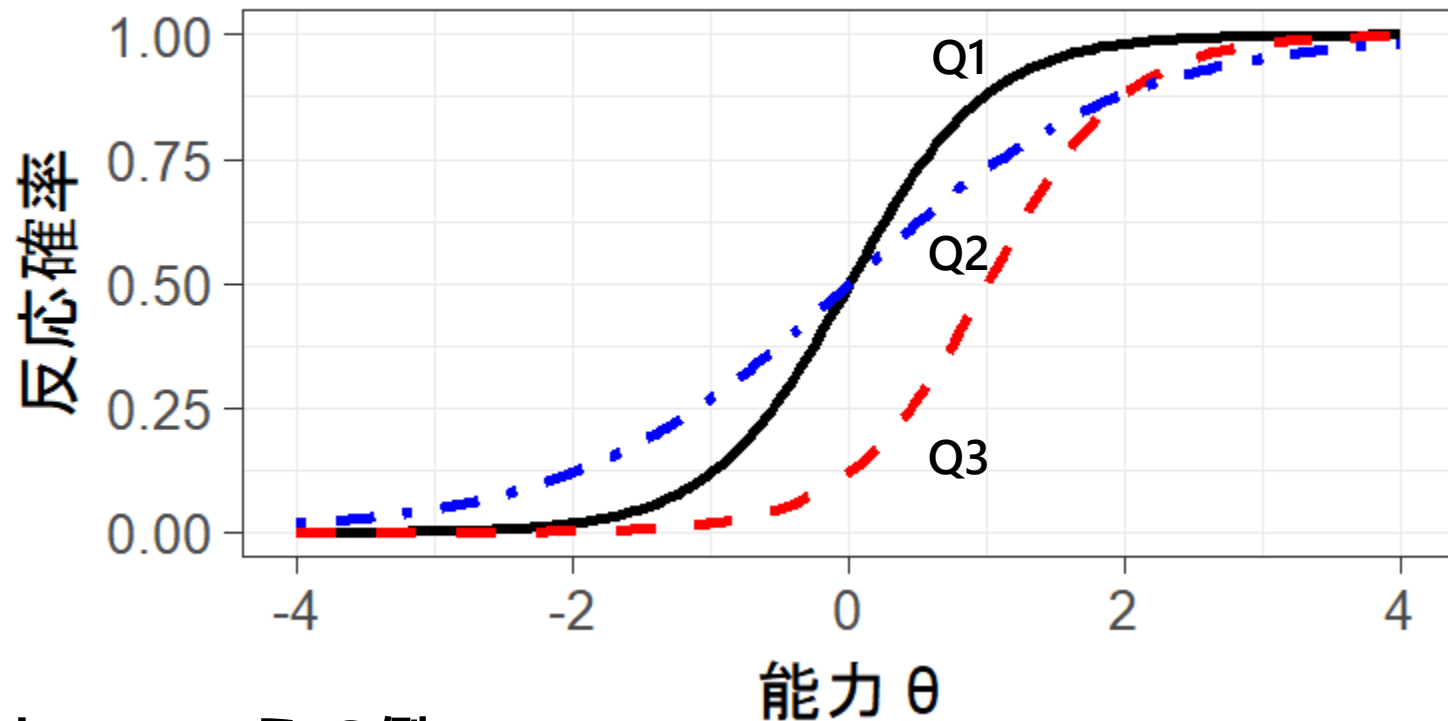


※2値(はい・いいえ)の例

ID	Q1	Q2	Q3
1	1	0	0
2	0	1	1
3	0	0	1

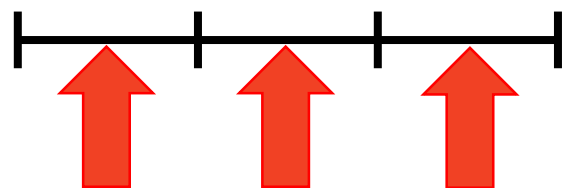
項目特性を推定し、
回答者の能力を推定

項目特性曲線(1つの線は1つの項目を表す)



※2値(はい・いいえ)の例

- 2つの(隣接)カテゴリのどちらが起こりやすいか？
 - たとえば、4件法だと以下のように考える



1	2	3
•	•	•
2	3	4
の	の	の
どちら	どちら	どちら
がお	がお	がお
き	き	き
やす	やす	やす
い？	い？	い？

各カテゴリ k に対する反応確率を $P_{jk}(\theta)$ とすると

$$P_{jk}^*(\theta) = \frac{P_{jk}(\theta)}{P_{jk-1}(\theta) + P_{jk}(\theta)}$$

ここから、整理すると

各区カテゴリの反応確率を以下と表現でき、

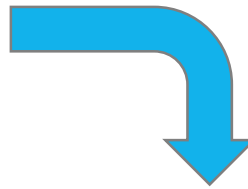
$$P_{jk}(\theta) = \frac{\exp \sum_{l=0}^k (\theta_i - b_{jl})}{\sum_{m=0}^{K_j} \exp \sum_{l=0}^k (\theta_i - b_{jl})}$$

名義反応モデルの特殊形

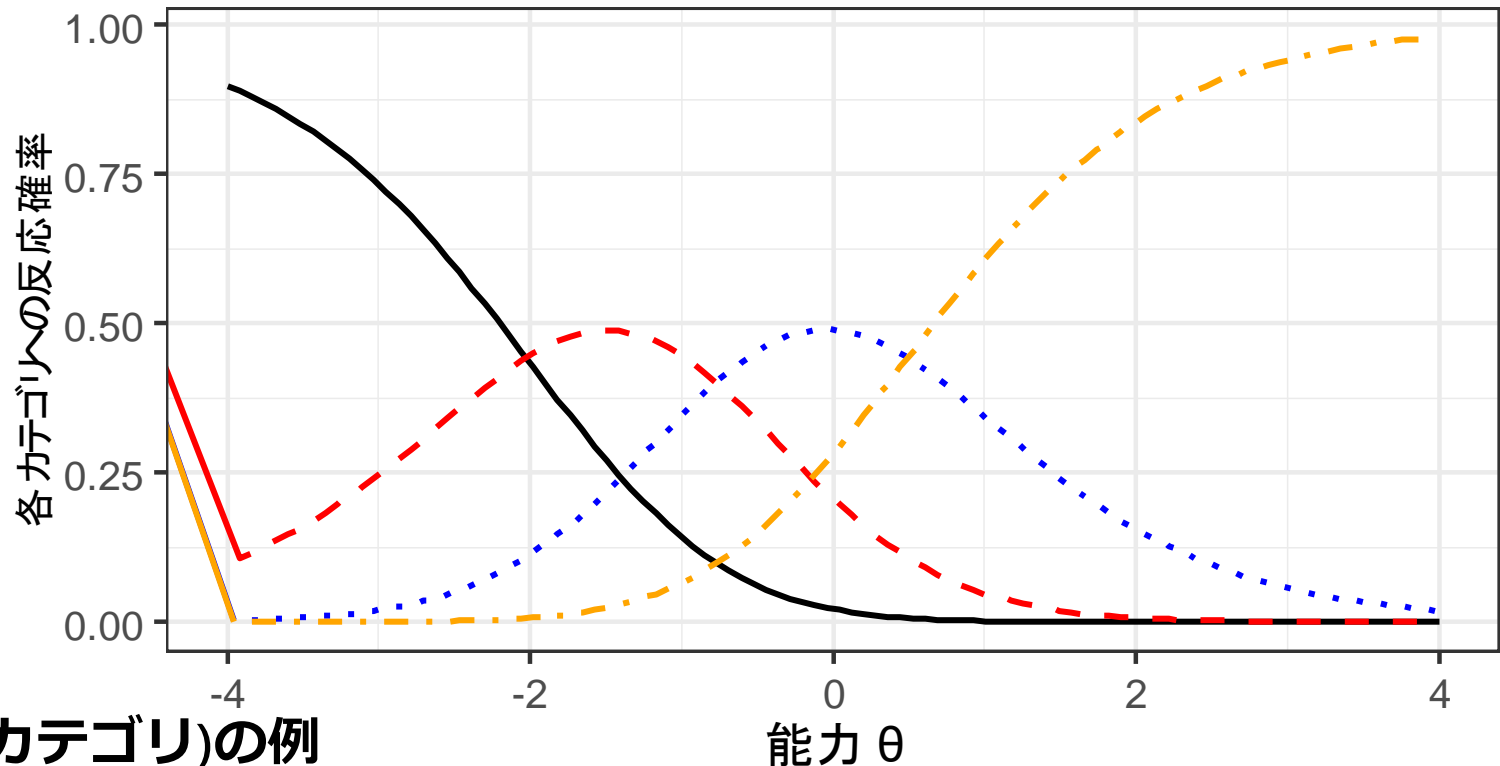
このように2値IRTモデル(1PL)を置く
※2PLを置くと、一般化部分得点モデル

ID	Q1	Q2	Q3
1	1	4	3
2	4	4	3
3	1	2	2

項目特性を把握し、
回答者の能力を測定



項目特性曲線(1つの線は1つのカテゴリを表す)



回答者 i 項目 j カテゴリ k

$$P(y_{ij} = k | \theta_i, a_j, s_j, c_j) = \frac{\exp([a \circ \mathbf{s}_{jk}]' \theta_i + c_{jk})}{\sum_{m=1}^K \exp([a \circ \mathbf{s}_{jm}]' \theta_i + c_{jm})}$$

a ... 識別力 \mathbf{s}_{jk} ... スコアリング関数(カテゴリ間隔)

θ_i ... 特性 (RSの程度) c_{jk} ... 切片

通常の多次元名義反応モデルとの変更点

制約

$$\tilde{a}'_{jk} = a \mathbf{s}_{jk}'$$

$$\sum_{k=1}^K c_{jk} = 0$$

\mathbf{s}_{jk} (スコアリング関数)によって

名義反応・部分得点・**反応スタイル**に対応

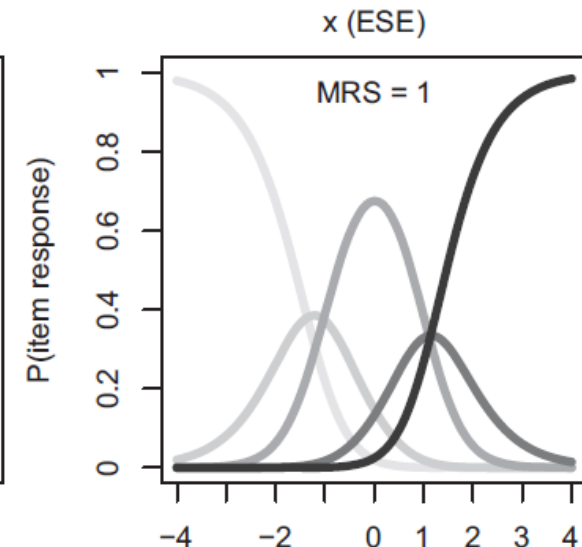
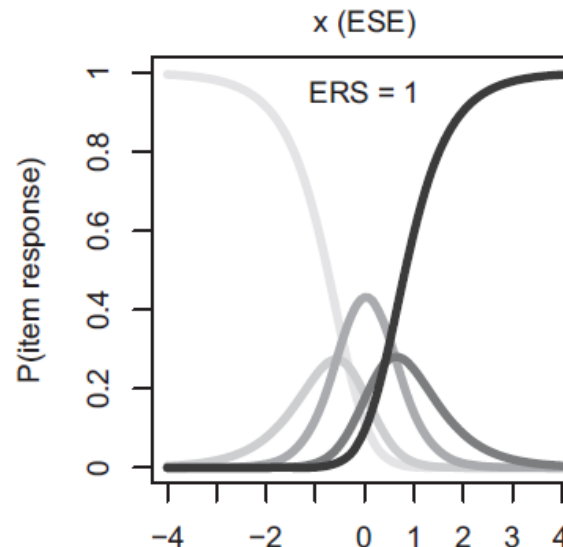
RS統制するためのモデル IRT編(Falk et al., 2016) 59

回答者 i 項目 j カテゴリ k

$$P(y_{ij} = k | \theta_i, \mathbf{a}, \mathbf{s}_j, \mathbf{c}_j) = \frac{\exp([\mathbf{a} \circ \mathbf{s}_{jk}]' \theta_i + c_{jk})}{\sum_{m=1}^K \exp([\mathbf{a} \circ \mathbf{s}_{jm}]' \theta_i + c_{jm})}$$

たとえば、5件法で以下の設定ならば

$\mathbf{s}_j = [0, 1, 2, 3, 4]'$ ➡ 部分得点モデル
 $= [1, 0, 0, 0, 1]'$ ➡ 極端反応スタイル
 $= [0, 0, 1, 0, 0]'$ ➡ 中間反応スタイル



回答者 i 項目 j カテゴリ k

$$P(y_{ij} = k | \boldsymbol{\theta}_i, \boldsymbol{a}, \boldsymbol{S}_j, \boldsymbol{c}_j) = \frac{\exp(a_1 s_{1jk} \theta_{i1} + a_2 s_{2jk} \theta_{i2} + c_{jk})}{\sum_{m=1}^K \exp(a_1 s_{1jm} \theta_{i1} + a_2 s_{2jm} \theta_{i2} + c_{jm})}$$

$\boldsymbol{s}_{1j} = [0, 1, 2, 3, 4]'$ にすると θ_{i1} は通常IRT (特性)

$\boldsymbol{s}_{2j} = [0, 0, 0, 1, 2]'$ にすると θ_{i2} は黙従反応スタイル

$$P(y_{ij} = k | \boldsymbol{\theta}_i, \mathbf{a}_j, \mathbf{s}_j, \mathbf{c}_j) = \frac{\exp([\mathbf{a} \circ \mathbf{s}_{jk}]' \boldsymbol{\theta}_i + c_{jk})}{\sum_{m=1}^K \exp([\mathbf{a} \circ \mathbf{s}_{jm}]' \boldsymbol{\theta}_i + c_{jm})}$$

事前分布

$$\boldsymbol{\theta}_i \sim \text{MultiNormal}(\mathbf{0}, \boldsymbol{\Omega}_{(\text{Cor Matrix})})$$

$$\boldsymbol{\Omega}_{(\text{Cor Matrix})} \sim \text{lkj}(1)$$

$$a_d \sim \text{LogNormal}(0, 1)$$

制約

$$\sum_{k=1}^K c_{jk} = 0$$

(右の制約をみたすためのパラメタリゼーション)

$$\mathbf{c} = \mathbf{T}_c \boldsymbol{\gamma}$$

$$\boldsymbol{\gamma}_{(K-1)} \sim \text{Normal}(0, 2)$$

$$\mathbf{T}_c = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & t_{22} & \dots & t_2(K-1) \\ 2 & t_{32} & \dots & t_3(K-1) \\ \vdots & \vdots & \vdots & \vdots \\ K-1 & 0 & \dots & 0 \end{bmatrix}$$

上のフーリエ基底関数を通すことで次元が落ちる(Thissen, & Cai, in press)。

RSへの対処アプローチ

- 方法論的アプローチ
 1. データの取り方を変えれば、RSの影響は減る？
 2. RSを測定しやすい調査デザインにしてしまう

1. データの取り方を変えれば、RSの影響は減る？ 63

• Likart式だから悪いのか？ RSでるのか？

- Likart vs Drag & Drop (DD)を検討(Böckenholt, 2017)
- Drag & Dropは、Thorstone法の拡張（次スライド）

• 結果

- Drag & Dropは、RS軽減する(Böckenholt, 2017)
- そんなことなかった(Plieninger, 2017)
- 今後さらなる検討が必要だが、
Likartのほうが調べつくされているし、
開発も、計算も色々楽だという主張(Plieninger, 2017)
- 僕個人の感想としては、Web調査が増えていくなかDD式の可能性もあるのかなと思います。

Please move the following items to the best-fitting response.

Items	Strongly Disagree	Strongly Agree
<p>I enjoy being spontaneous.</p> <p>I find that a well-ordered life with regular hours makes my life tedious.</p> <p>I don't like situations that are uncertain.</p> <p>I hate to change my plans at the last minute.</p> <p>I hate to be with people who are unpredictable.</p> <p>I find that a consistent routine enables me to enjoy life more.</p> <p>I enjoy the exhilaration of being in unpredictable situations.</p> <p>I become uncomfortable when the rules in a situation are unclear.</p>		
	<p>Disagree</p> <p>It upsets me to go into a situation without knowing what I can expect from it.</p>	<p>Agree</p> <p>I'm not bothered by things that interrupt my daily routine.</p>
	<p>Somewhat Disagree</p>	<p>Somewhat Agree</p> <p>I enjoy having a clear and structured mode of life.</p>

Figure 7. Instructions and items for drag-and-drop response format.

- 選択肢を選ぶ
- ↓
- 項目を選択カテゴリの枠内に移動

• Plieninger (2017)では、**カテゴリの列数の影響も含め検討**

Ich bin ungern mit Leuten zusammen, deren Verhalten nicht vorhersehbar ist.

Ich finde, dass eine gewisse Routine es mir ermöglicht, mein Leben mehr zu genießen.

Ich genieße die Herausforderung, mich in unvorhersehbaren Situationen zu befinden.

Ich fühle mich unwohl, wenn die Regeln in einer Situation unklar sind.

stimme voll und ganz zu

Ich mag es, wenn alles seinen Platz hat und alles an seinem Platz ist.

stimme zu

Ich genieße es, spontan zu sein.

stimme eher zu

Ich finde, dass ein wohlgeordnetes Leben mit regelmäßigen Abläufen langweilig ist.

stimme eher nicht zu

Ich hasse es, meine Pläne in der letzten Minute zu ändern.

stimme nicht zu

stimme überhaupt nicht zu

(A) Response format DnD-I.

Ich bin ungern mit Leuten zusammen, deren Verhalten nicht vorhersehbar ist.

Ich finde, dass eine gewisse Routine es mir ermöglicht, mein Leben mehr zu genießen.

Ich genieße die Herausforderung, mich in unvorhersehbaren Situationen zu befinden.

Ich fühle mich unwohl, wenn die Regeln in einer Situation unklar sind.

stimme überhaupt nicht zu

Ich mag es, wenn alles seinen Platz hat und alles an seinem Platz ist.

stimme voll und ganz zu

stimme nicht zu

Ich genieße es, spontan zu sein.

stimme zu

stimme eher nicht zu

Ich finde, dass ein wohlgeordnetes Leben mit regelmäßigen Abläufen langweilig ist.

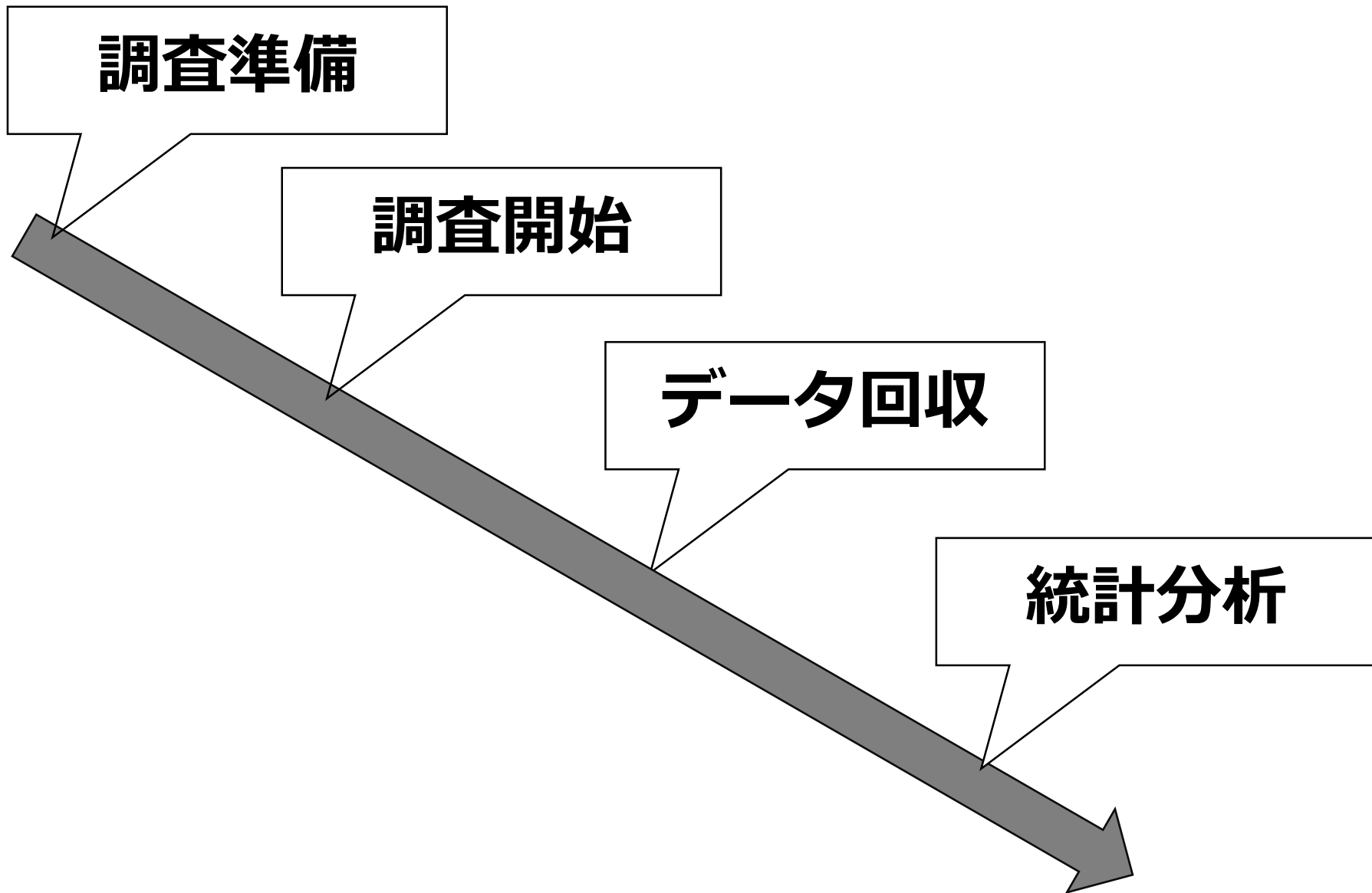
stimme eher zu

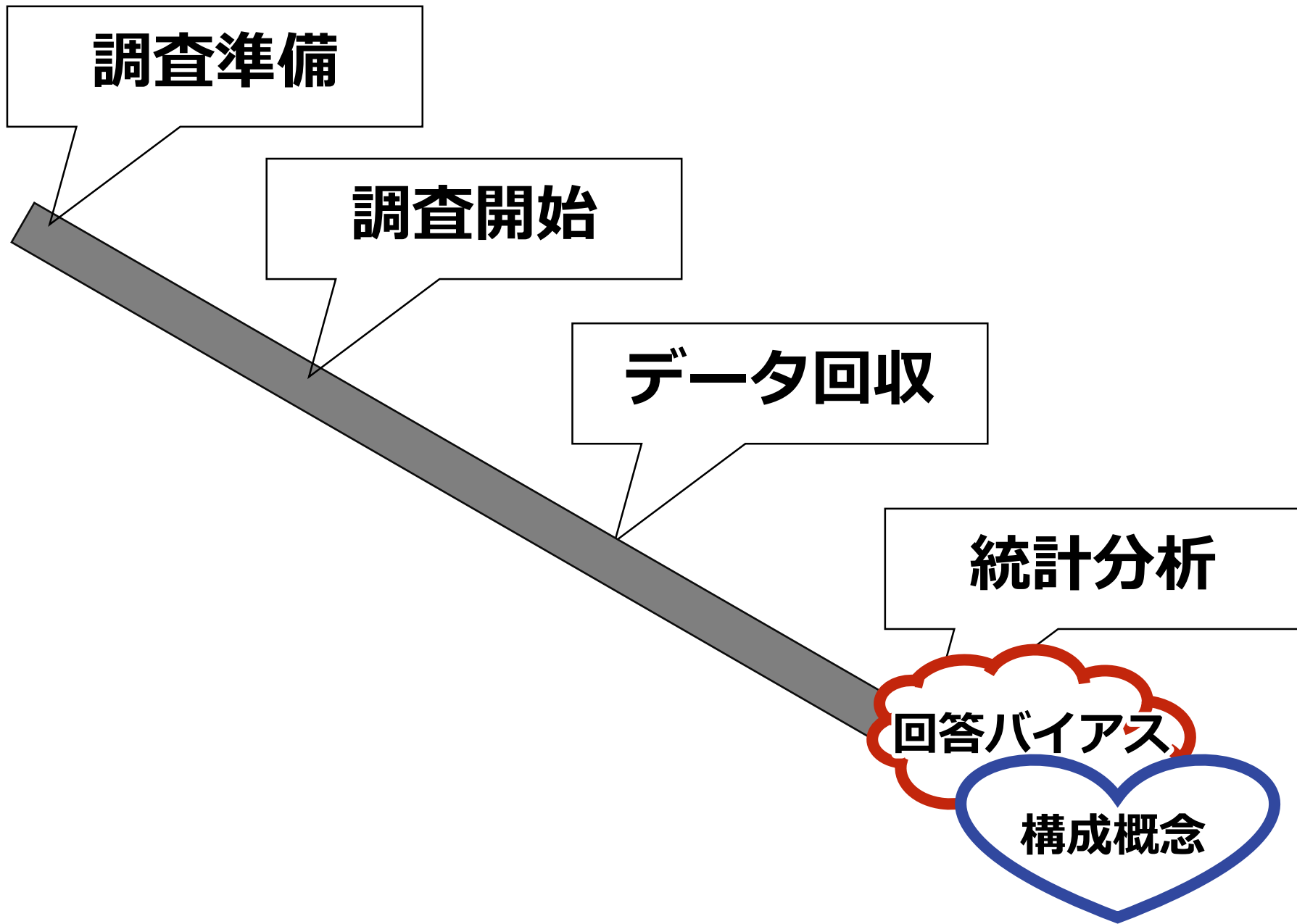
Ich hasse es, meine Pläne in der letzten Minute zu ändern.

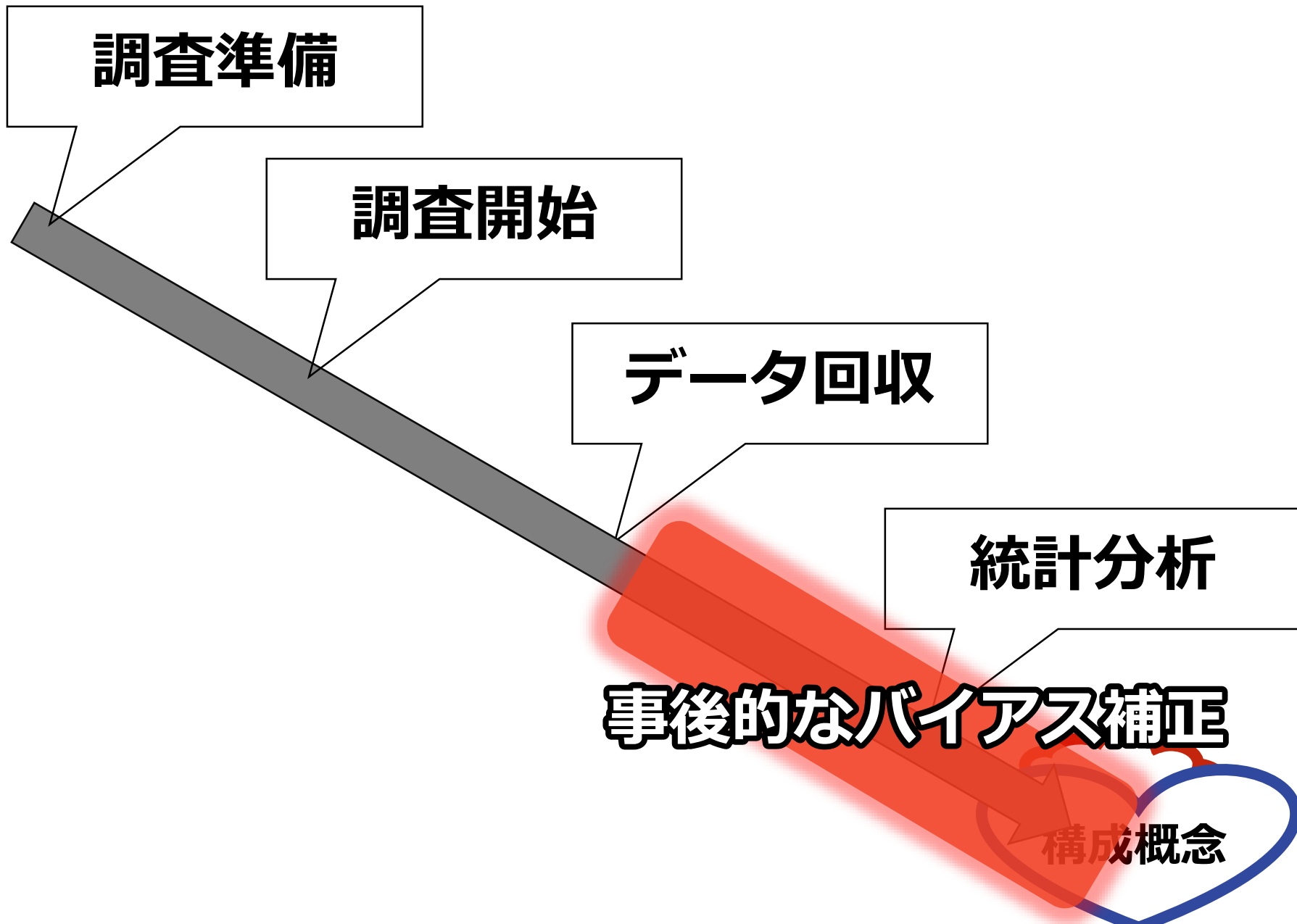
(B) Response format DnD-II.

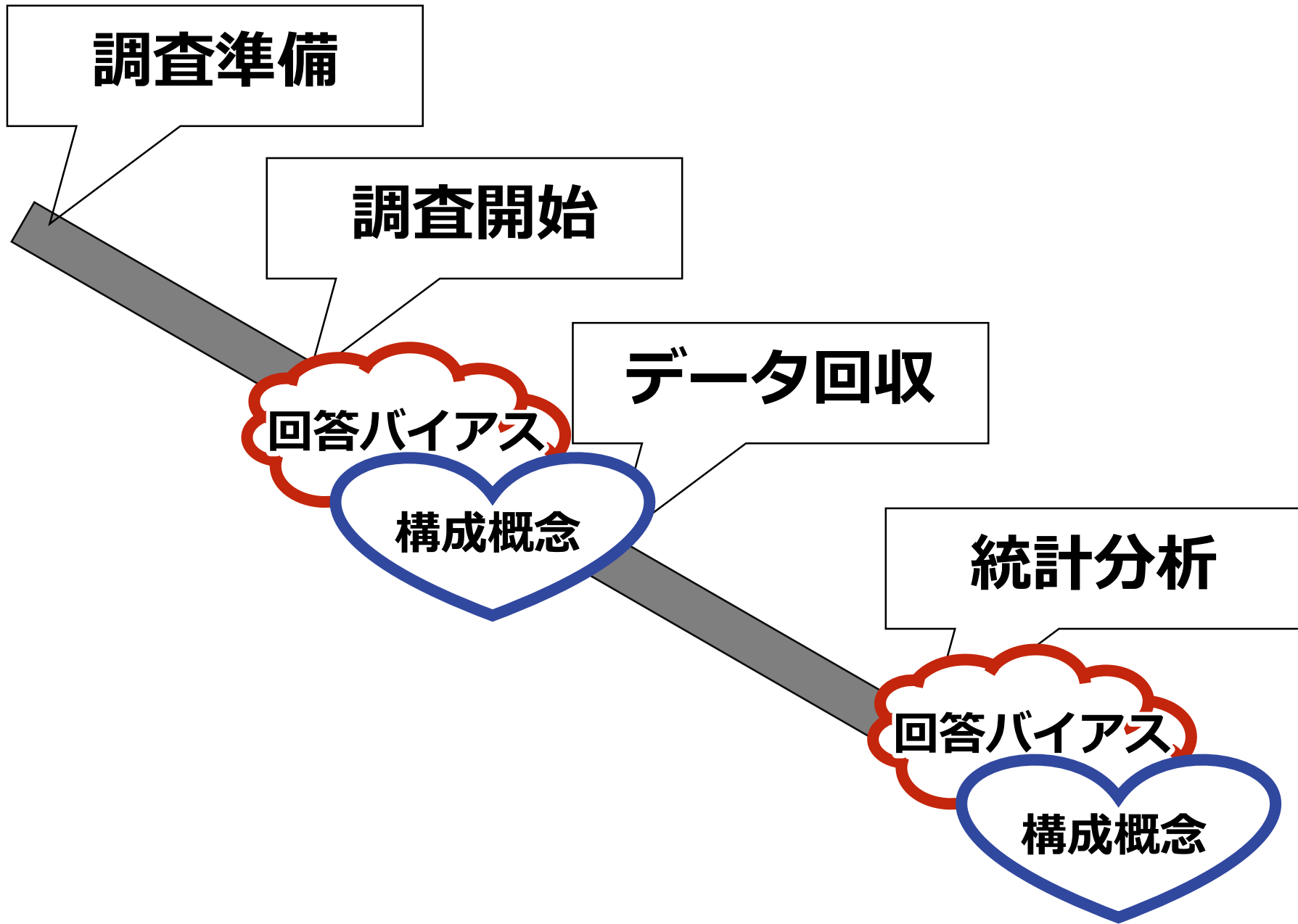
FIGURE 4.1: Drag-and-drop formats used in Study 2 of Plieninger et al. (2017). To-be answered items appear on the left, and already answered items appear in the chosen category on the right.

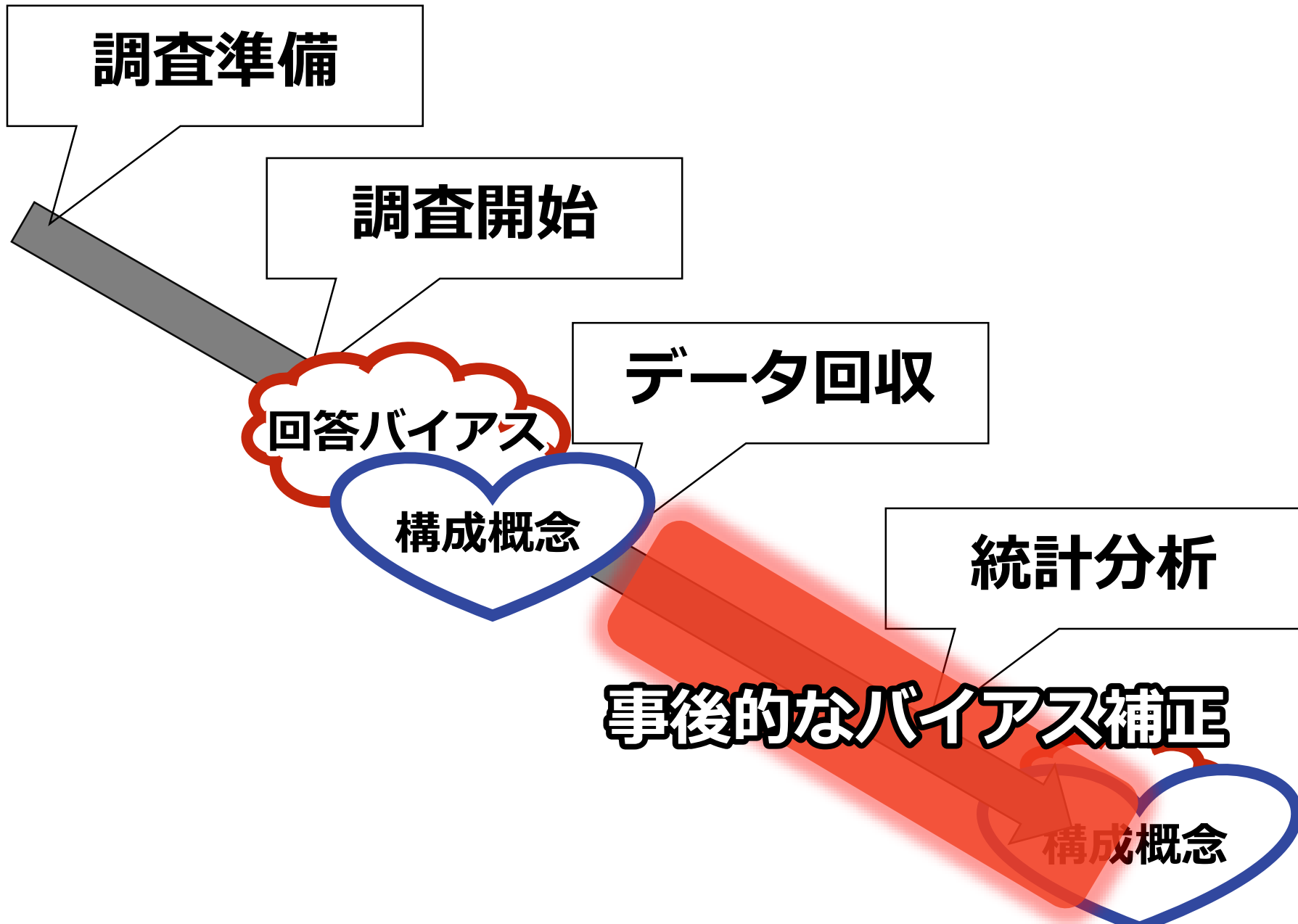
- Anchoring Vignettes method (King et al., 2004)
 - アンカリングビネット・係留ビネット・係留寸描法とも
 - 何をするのか？
 - 通常の項目回答に加えて、ビネット(短い記述)を読んで回答する項目(ビネット項目)を追加し、これを利用（アンカリング）して、RSの測定につなげる方法（詳細は後述）
 - もうすこしイメージで...(次スライド)











調査準備

調査開始

反応スタイルを直接
測定可能な調査デザイン

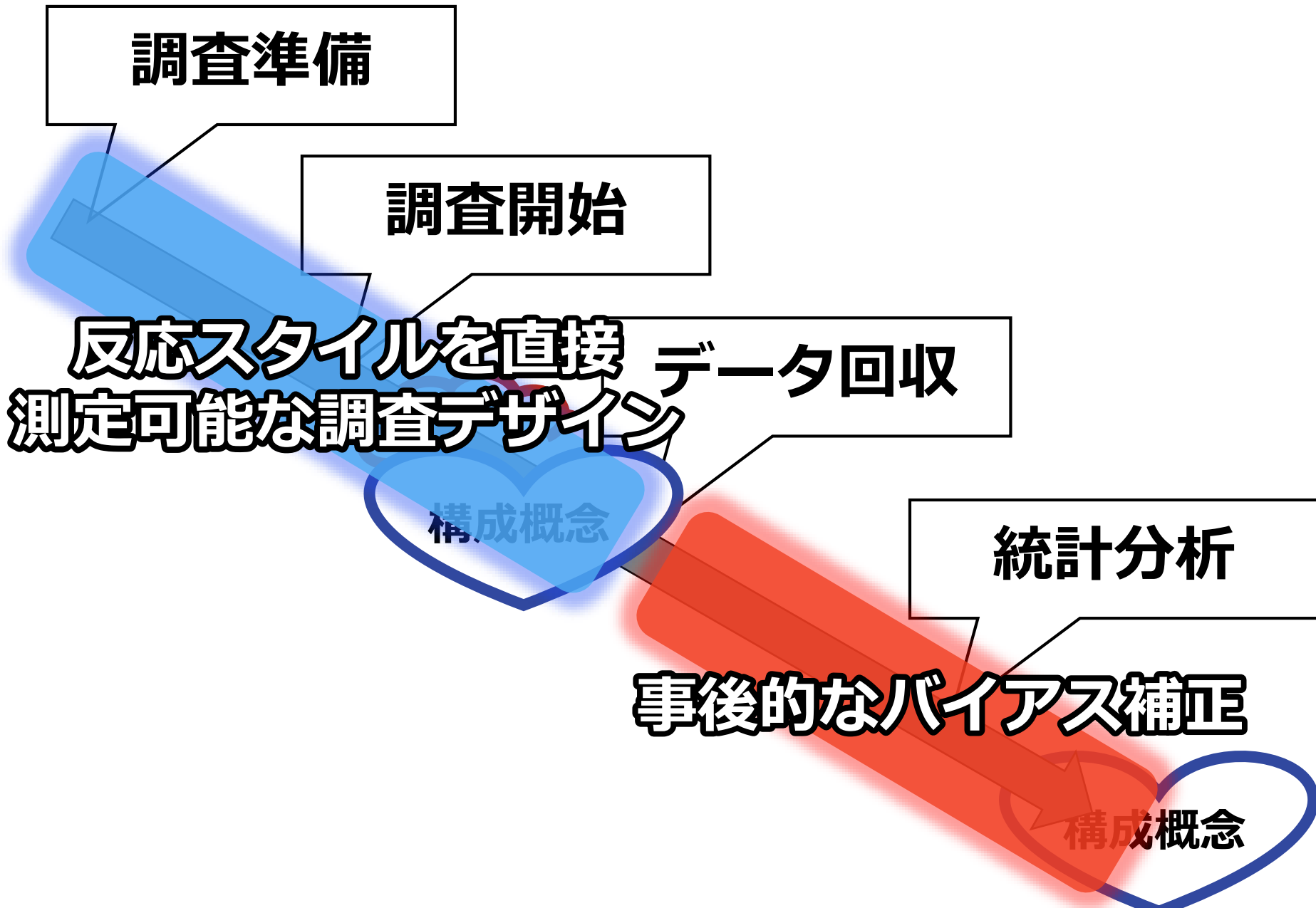
データ回収

構成概念

統計分析

事後的なバイアス補正

構成概念



調査準備

調査開始

けいりゅう
係留ビネット法

データ回収

統計分析

ベイズ多次元IRTモデル

調査準備

調査開始

Drag & Drop法

データ回収

統計分析

構成概念

DD法でRSを軽減できるなら、
のちの統計分析で補正いらなくなるかも！？
というのを検討していた

(Böckenholt, 2017)
(Plieninger, 2017)

調査準備

調査開始

けいりゅう
係留ビネット法

データ回収

統計分析

ベイズ多次元IRTモデル

- 回答者(あなた)の回答項目とは別に
仮想人物(A,B...)についての複数のビネットを用意し、
回答者(あなた)に読ませ、その仮想人物への回答も得る手法

ビネット項目×3

反応スタイルのみ反映

Aさんは、仕事を楽しみ、

日々の活動に大きな支障を来してはいない。

Bさんは、仕事に不満で、

Q-1 ときどき活動に支障を来している。

Cさんは、仕事ができないほど具合が悪いことがあり、

Q-2 日々の活動に大きな支障を来している。

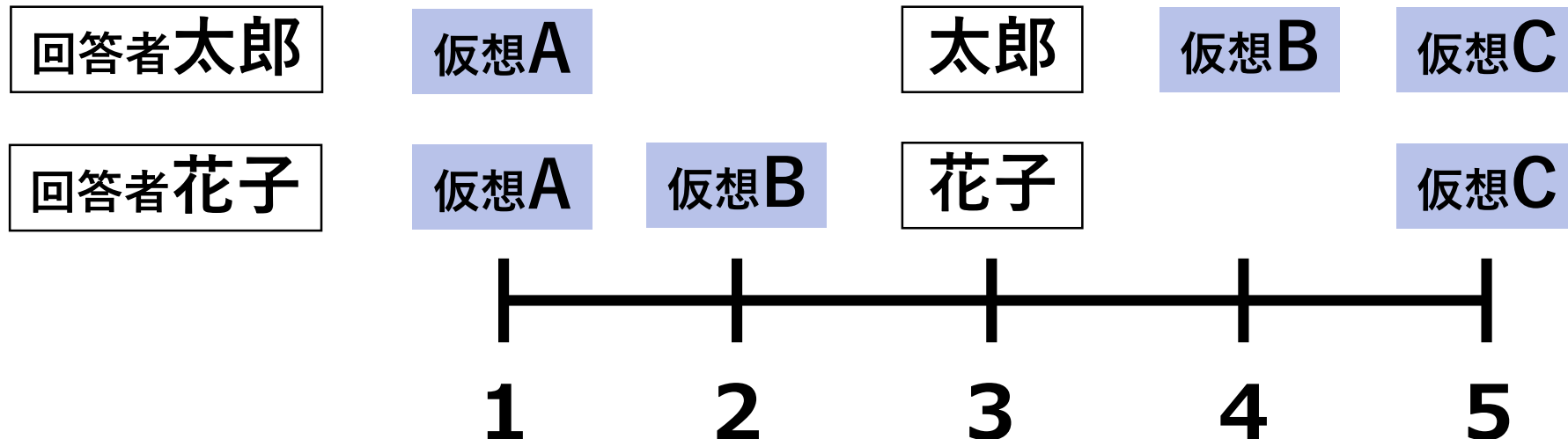
Q-3 あなたはCさんがどれだけ落ち込んでいると感じますか？

1 2 3 4 ⑤

自己評価項目 構成概念+反応スタイルを反映

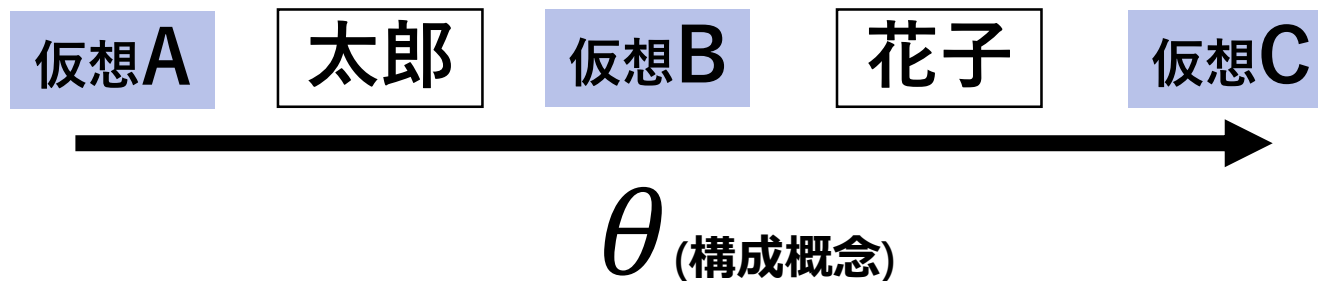
Q-4 あなたは自分がどれだけ落ち込んでいると感じますか？

1 2 3 ④ 5



モデル

多次元IRT型モデル



- **Response consistency (反応一貫性)**
 - 全てのビネットに同じように（同じバイアスのもとで）回答しているという仮定
 - この項目にはバイアスA、別の項目にはバイアスBではだめ
- **Vignettes equivalence (ビネット等価性)**
 - 全ての回答者がビネットの内容を“ちゃんと”理解できているという仮定
 - ビネットの仮想人物にどう思うかではなく、ビネットの記述を理解できているかということ
- この2つが満たされないと、うまくいかない

通常の調査方法

I	自己	自己	自己	自己
D	Q1	Q2	Q3	Q4

係留ビネット法

I	自己	自己	自己	自己	ビ	ビ
D	Q1	Q2	Q3	Q4	ネ	ネ
					Q1	Q2

もしくはこんな感じでも良いかも？

80

通常の調査方法

I	自己	自己	自己	自己
D	Q1	Q2	Q3	Q4

係留ビネット法

I	自己	自己	ビ	ビ
D	Q1	Q2	ネ	ネ
			Q1	Q2

もしくははこんな感じでも良いかも？

81

通常の調査方法

I	自己	自己	自己	自己
D	Q1	Q2	Q3	Q4

係留ビネット法

I	自己	自己	ビ	ビ
D	Q1	Q2	ネ	ネ
			Q1	Q2

このように収集された
自己評定項目データとビネット項目データを
統計モデルにかけることで、
反応スタイルを測定しやすくする

係留寸描法データのための統計モデル

CHOPIT法 (Compound Hierarchical Ordered Probit Model, King et al., 2004)

- 反応スタイルの**集団差**を検討するために用いられる
 - 項目反応理論(IRT)に基づいていない
 - 実用上、データへの当てはまりがよくない(van Soest & Vonkova, 2014)
 - 説明（顕在）変数が多いとき有効（心理きつい）

ノンパラメトリック法 (King et al., 2004)

- モデルベースではないため、結果の一般化可能性に問題あり

ベイズ多次元IRT型モデル

- 反応傾向バイアスの**個人差**も検討することができる
 - 現代テスト理論であるIRTに基づいたモデル
 - ベイズアプローチを用いているため、**複雑な応用も可能**
 - **反応スタイルを自由推定可能**(Bolt et al., 2014)
 - **近年、段階反応型モデルも提案** (Jonas & Markon, 2018)

RSについての3つの疑問

- RS研究の歴史と動向
- RS(Response Style)の種類
- RSと様々な変数との関係性
- RSへの対処アプローチ
 - 統計モデル的アプローチ
 - 方法論的アプローチ
- RSについての3つの疑問
 - **Psychometric level** – RSを測定するには?
 - **Substantive level** – 何でRSが説明できる?何に影響される?
 - **Applied level** – RS除去したい(興味があるのは特性だけ)
- RSに対処するモデル
 - **Item Response Model**
 - **Item Response Tree Model**

- RSの起源を辿り、
- RSが様々な要因と関係していそうなのがわかった
- そして、方法とモデルの二側面から主流アプローチの紹介。
- ここからは、目標に応じてRSをどうしたいのか考える

- **心理測定レベル(Psychometric Level)**
 - RSを測定するのが最終目標
 - 方法やモデル作り
- **本質レベル(Substantive Level)**
 - RSを説明するものは何か？
 - 個人差レベル？
- **実務レベル(Applied Level)**
 - RSにはあまり興味ない
 - 調査から目的とする特性を取り出したい

- **心理測定レベル(Psychometric Level)**
 - RSを測定するのが最終目標
 - 方法やモデル作り
- **本質レベル(Substantive Level)**
 - RSを説明するものは何か？
 - 個人差レベル？
- **実務レベル(Applied Level)**
 - RSにはあまり興味ない
 - 調査から目的とする特性を取り出したい

- RSを測定することが目標
- 近年、方法・モデルでこれが可能に
 - RIRSMACSや多次元名義反応モデル
 - 係留ビネット法
- RSを測定可能になったのであれば...
- RSの影響力を調べてみた (Plieninger, 2017b)

the Impact of Response Styles (Plieninger, 2017b) 89

- 多次元名義反応モデル(Wetzel, & Carstensen, 2015)を真のモデルとしてデータ生成
- これを上記モデルで分析したときのズレ(Bias)をチェックした
 - Cronbach α (信頼性の確認)
 - 別尺度との相関 (妥当性の確認)
 - 真値との相関
- **結果**
 - **特性とRSの相関が高い時にBiasが極端に大きくなる**
 - 逆に言えば、相関が大きくなければRSの影響は小さい
 - 今後はRSの効果よりも、
 - 何がRSに影響しているのか (**本質レベル**)が、大事になるだろう
 - **実務レベル**は、既存の多次元名義反応モデルを使えばよいだろう

- **特性と黙従反応スタイル(ARS)を弁別が難しい**
- **なぜ？**
 - **特性**が高い → 項目得点が高くなる
 - **黙従**が高い → 項目得点が高くなる
 - この両者が似通ってしまうから(相関が高くなるから)
- **どうすればいい？**
 - **係留ビネット法**を適用したモデルで解決できるかも

- 回答者(あなた)の回答項目とは別に
仮想人物(A,B...)についての複数のビネットを用意し、
回答者(あなた)に読ませ、その仮想人物への回答も得る手法

ビネット項目×3

反応スタイルのみ反映

Aさんは、仕事を楽しみ、

日々の活動に大きな支障を来してはいない。

Bさんは、仕事に不満で、

Q-1 ときどき活動に支障を来している。

Q-2 Cさんは、仕事ができないほど具合が悪いことがあり、

日々の活動に大きな支障を来している。

Q-3 あなたはCさんがどれだけ落ち込んでいると感じますか？

1 2 3 4 ⑤

自己評価項目 構成概念+反応スタイルを反映

Q-4 あなたは自分がどれだけ落ち込んでいると感じますか？

1 2 3 ④ 5

- **特性と黙従反応スタイル**を弁別することが難しかったが
 - **特性**が高い → **自己評定**項目得点が高くなる
 - **黙従**が高い → **自己評定・ビネット**項目得点が高くなるとなるので、
- 既存のモデルを係留ビネットデータ用に改良するなり、
- 係留ビネット法で用いられるモデルを用いれば解決かな？

そこで...

RSの**実務**と**本質**レベルに 対処するモデル

Item Response Model

Item Response Tree Model

紹介モデルのパラメータ推定法は基本ベイズです。

ベイズは基本です。

- 係留ビネット法を使わない
 - Falk & Cai (2016) **多次元名義反応モデル**
- 相関が高いと困る
- 係留ビネット法を使う場合
 - Falk & Cai (2016) **の拡張** (Hojo & Okada, in prep)
 - Bolt et al. (2014) **の反応スタイル自由推定型モデル**
 - Jonas & Markon (2018) **の段階反応型モデル**

回答者 r が項目 i に対して k 番目のカテゴリを選択する確率を

$$P(U_{ri} = k | \theta_r, s_r) = \frac{\exp(A_{ik}\theta_r + s_{rk} + c_{ik})}{\sum_{h=1}^K \exp(A_{ih}\theta_r + s_{rh} + c_{ih})}$$

where $\sum_h s_{rh} = 0$ and $\sum_h c_{ih} = 0$.

$A_i = \{0, 0, 0, 0, 0\}$ or $\{-2, -1, 0, 1, 2\}$.

モデルの肝 ↘

$$A_{ik}\theta_r + s_{rk} + c_{ik}$$

カテゴリ間隔(固定値) × 構成概念 + 反応スタイル + 切片

回答バイアスのみ

ビネット項目

$$s_{rk} + c_{ik}$$

構成概念+回答バイアス

自己評定項目

$$A_{ik}\theta_r + s_{rk} + c_{ik}$$

• 反応スタイルを量的に測定可能

- Falk らのモデルでは、反応スタイルはスコアリング関数だった(例えば $ERS = [2, 1, 0, 1, 2]$)
- しかし、Boltらのモデルでは、**反応スタイルを個人ごとに直接推定可能**(Aさんの反応スタイル = $[-2.3, 0, 0, 0, 0]$)

• 反応スタイルの解釈

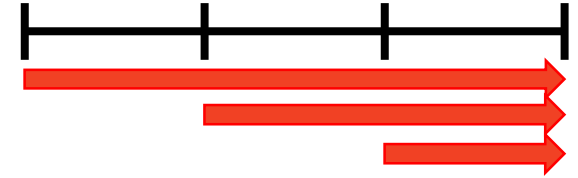
- **正(+)** = そのカテゴリに**答えやすい**バイアス
- **0** = そのカテゴリへの**反応スタイルはない**
- **負(-)** = そのカテゴリに**答えにくい**バイアス

このモデルを用いた実データ分析は、第3部にて

- カテゴリk以上となる確率(累積確率)を考える

- 4件法(1,2,3,4)のとき

- 1以上の確率= $A = 1$ (全員あてはまるので)
- 2以上の確率= B
- 3以上の確率= C
- 4以上の確率= D
- 5以上の確率= $E = 0$ (全員ありえないので)
と考えると



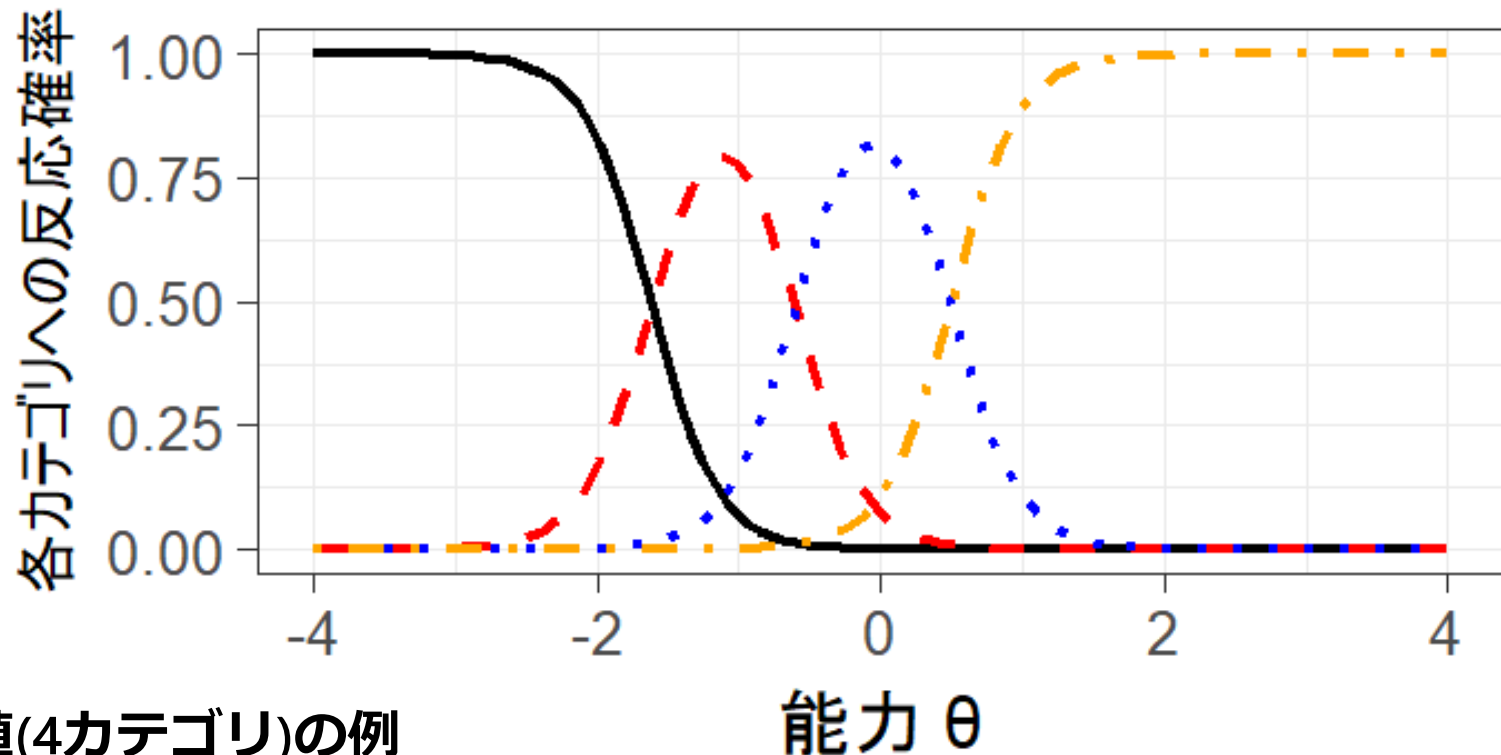
- 1と回答する確率= $A-B$ (回答が1以上ではあるが、2以上ではない)
- 2と回答する確率= $B-C$
- 3と回答する確率= $C-D$
- 4と回答する確率= $D-E$

- **k以上なら1・k-1以下なら0**という2値に落とし込み
- A~Eそれぞれの累積確率に2値IRT(2PL)モデルを置く

ID	Q1	Q2	Q3
1	1	4	3
2	4	4	3
3	1	2	2

項目特性を把握し、
回答者の能力を測定

項目特性曲線(1つの線は1つのカテゴリを表す)



- Falk、Boltモデルとの大きな違い

➡ Jonasのモデルは段階反応モデル

- このモデルでは、反応スタイルを次のように考える
 - ERS ➡ 人よりもカテゴリの間隔が**広い**
 - MRS ➡ 人よりもカテゴリの間隔が**狭い**
 - ARS ➡ 人とカテゴリの**位置が正にずれている**
- つまり、項目反応モデル的にいえば、
 - ERS&MRSは、**識別力のscaleの個人差**
 - ARSは、**困難度(閾値)の個人差(ずれ)**

という風に定義できる

- 通常の段階反応モデル

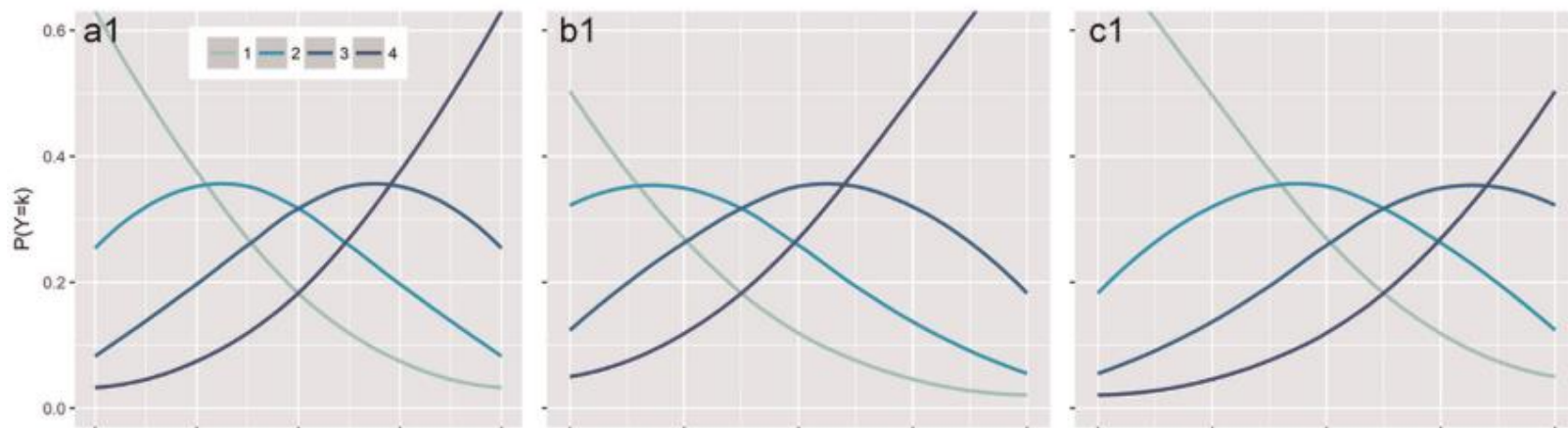
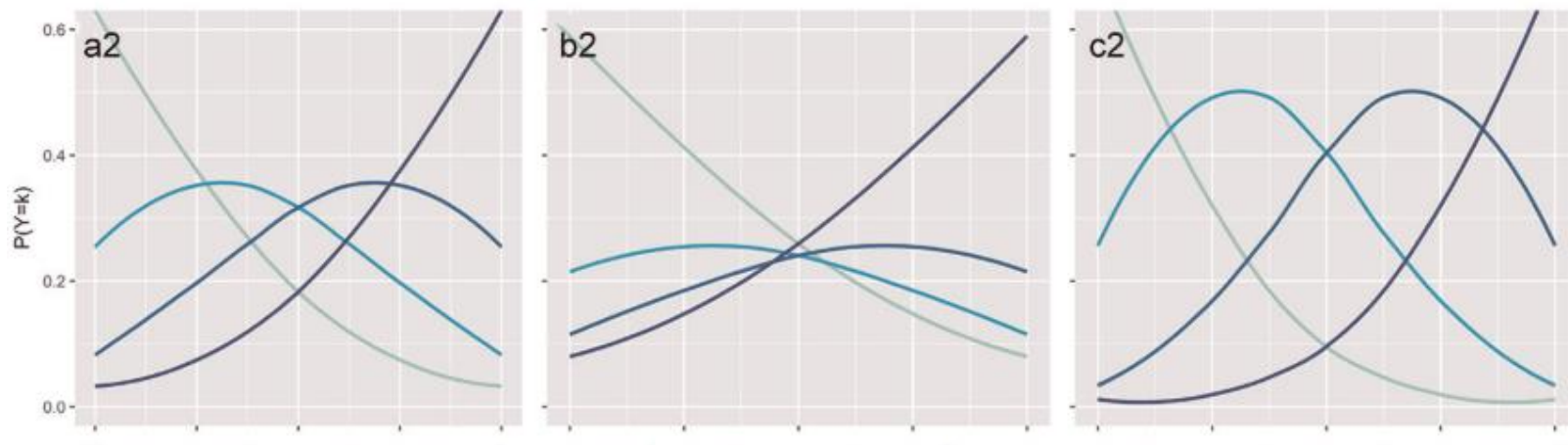
$$P(Y_{ij} > k) = \frac{\exp(a_j \theta_i - b_{jk})}{1 + \exp(a_j \theta_i - b_{jk})}$$

- Jonasらのモデル(a+bモデル)

$$P(Y_{ij} > k) = \frac{\exp\left(a_j \mathbf{a}_i \theta_i - (b_{jk} + \mathbf{b}_i)\right)}{1 + \exp\left(a_j \mathbf{a}_i \theta_i - (b_{jk} + \mathbf{b}_i)\right)}$$

- \mathbf{a}_i : ERS, MRSに影響, 反応確率のscaleの個人差
- \mathbf{b}_i : ARSに影響, 反応確率の位置(閾値)の個人差

- a_i , b_i の変化と項目反応曲線の変化

 b_i  a_i 

- 係留ビネット法を使わない
 - Falk & Cai (2016) **多次元名義反応モデル**
- 相関が高いと困る
- 係留ビネット法を使う場合
 - Falk & Cai (2016) **の拡張** (Hojo & Okada, in prep)
 - Bolt et al. (2014) **のモデル**
 - Jonas & Markon (2018) **のモデル**
- 紹介したモデルが今後主流になる可能性

Response Style Modeling

RSの**本質**を検討したい

- 以下のような項目がある

Q あなたは幸せですか？

1. とてもあてはまらない
2. あまりあてはまらない
3. すこしあてはまる
4. とてもあてはまる

Q あなたは幸せですか？

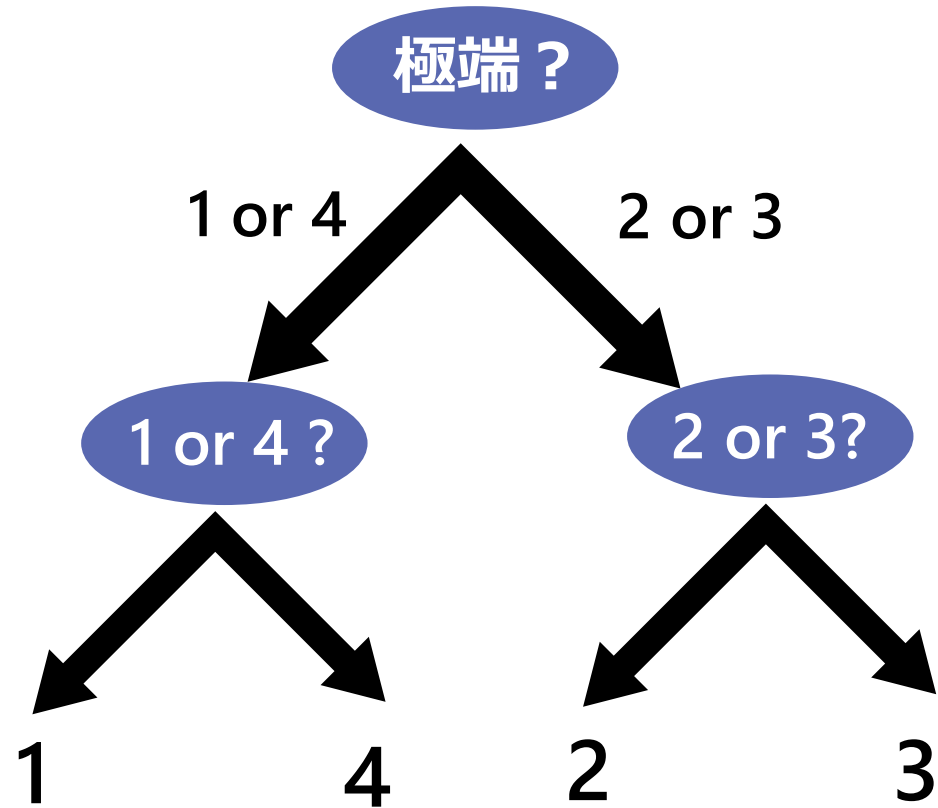
1. とてもあてはまらない
- 2.
- 3.
4. とてもあてはまる

※2, 3はラベルに文字がない

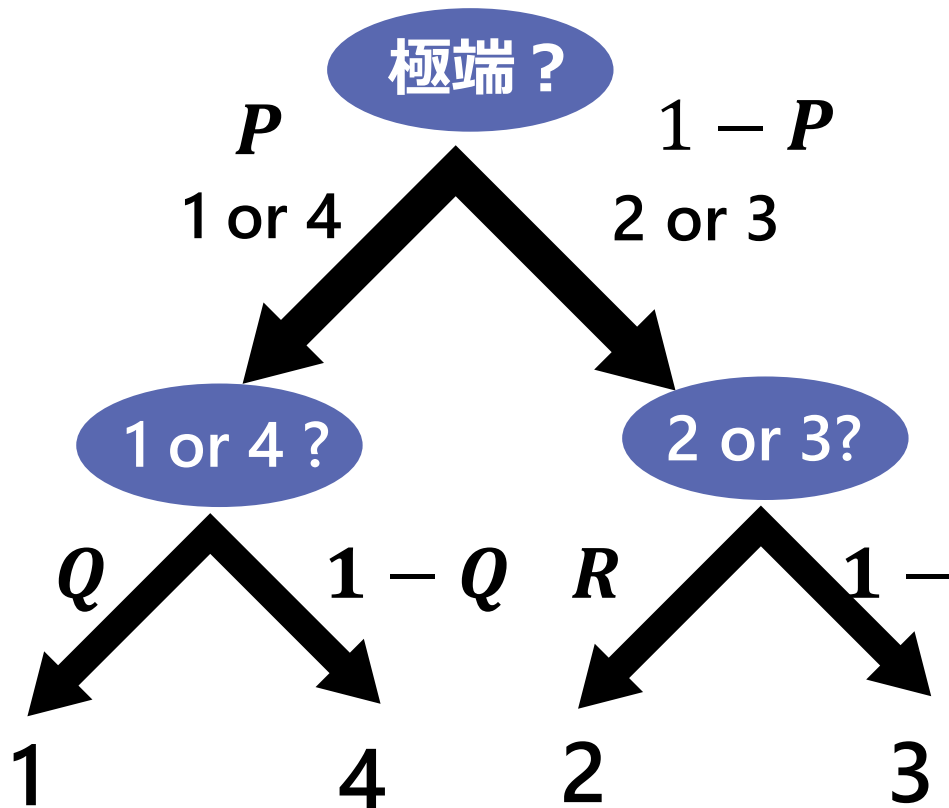
- ラベルの抜け具合で反応過程に影響は？

Q あなたは幸せですか？

1. とてもあてはまらない
- 2.
- 3.
4. とてもあてはまる



- 上の場合、
 - まず、1, 4に目が行くかもしれない
 - 次に、2, 3の選択肢に目が行くかもしれない



P = 極端なカテゴリを選択する確率
 $1 - P$ = 極端でない方を選択する確率
として、他も同様に考えてあげると、

回答者が
1と答える確率は、

$$P * Q$$

4と答える確率は

$$P * (1 - Q)$$

2と答える確率は

$$(1 - P) * R$$

3と答える確率は

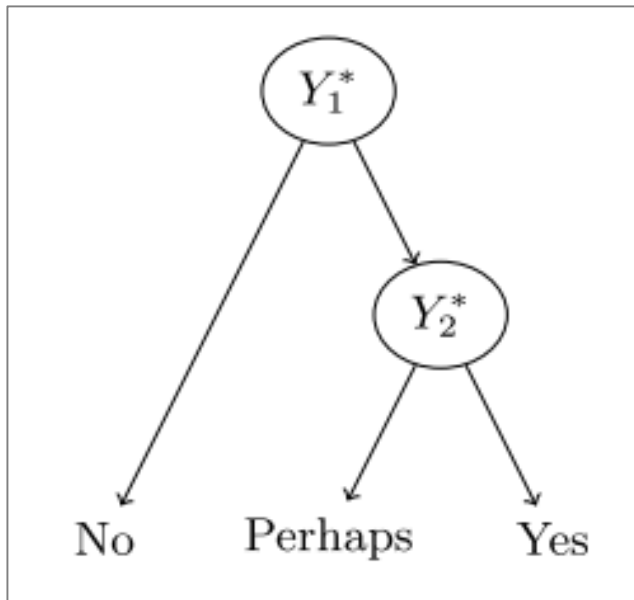
$$(1 - P) * (1 - R)$$

で表せる

• IRTree (項目反応木) モデル

(Boeck & Partchev, 2012; Böckenholt, 2017)

- 回答過程をモデリングできる
- 多値回答を以下の樹形図のように表現
- 各ノードごとに二値IRTを実施



左図の回答過程

- Noかどうか判断(ノード Y_1^*)
- 確信度を判断(ノード Y_2^*)

(Boeck & Partchev, 2012) Fig1より

- RSのためのIRTree モデル
- とりあえず、**灰色部分のみ**に注目してください

Plieninger, H. (2017). Fig3.1 より

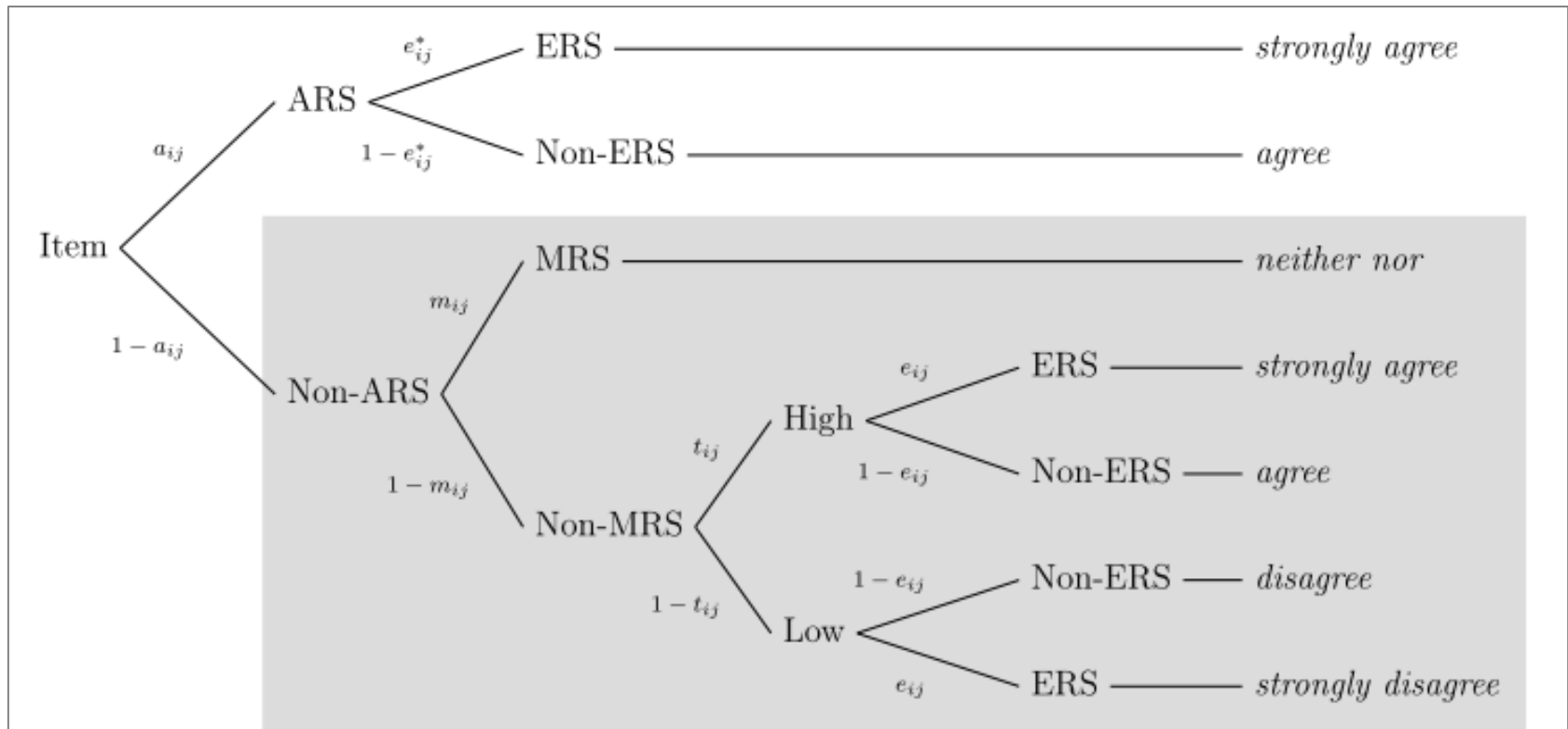


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

- RSのためのIRTree モデル
- **今回は、5件法の項目回答の状況を考える**

Plieninger, H. (2017). Fig3.1 より

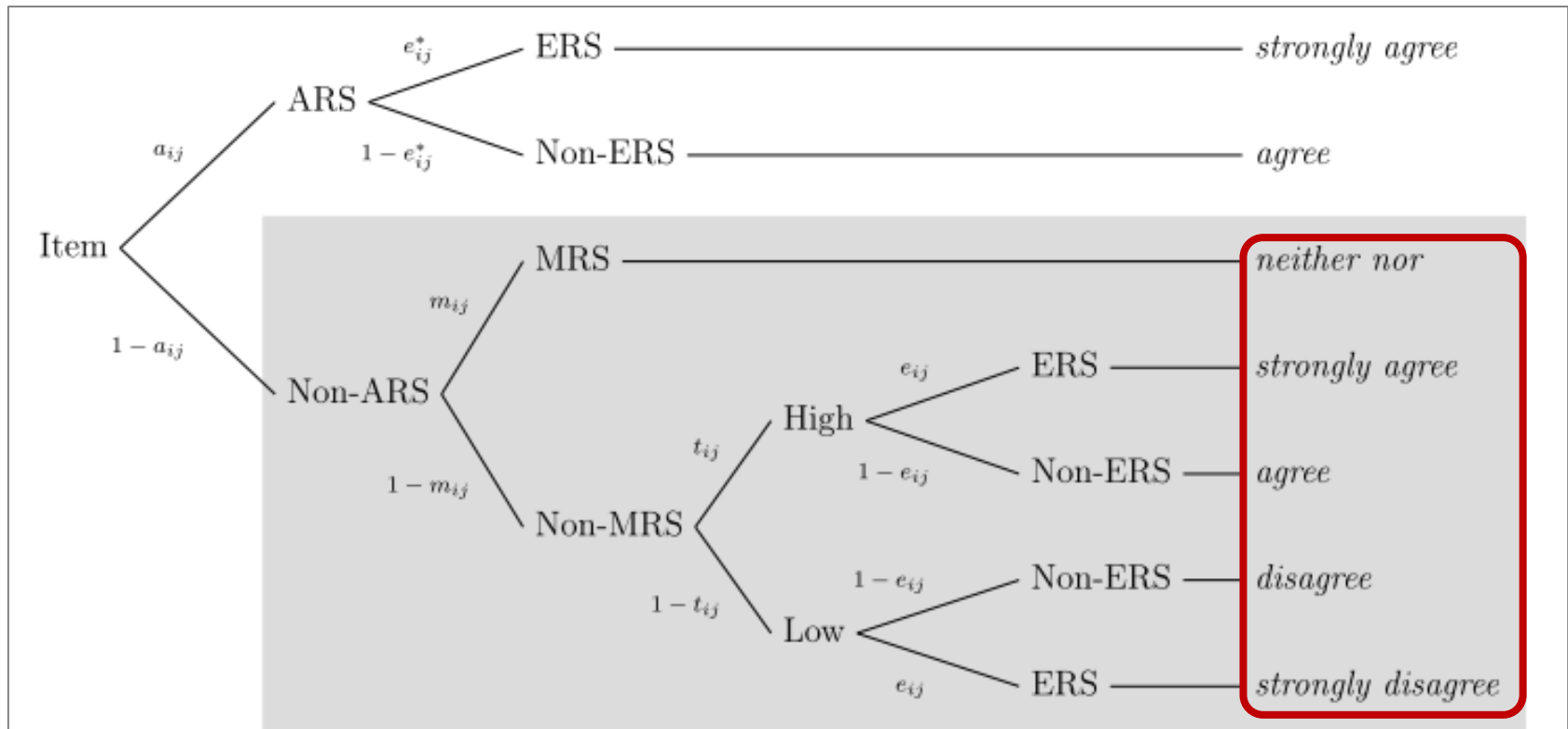


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

• RSのためのIRTree モデル

Plieninger, H. (2017). Fig3.1 より

まずは中間（3）かそれ以外か（1,2,4,5）

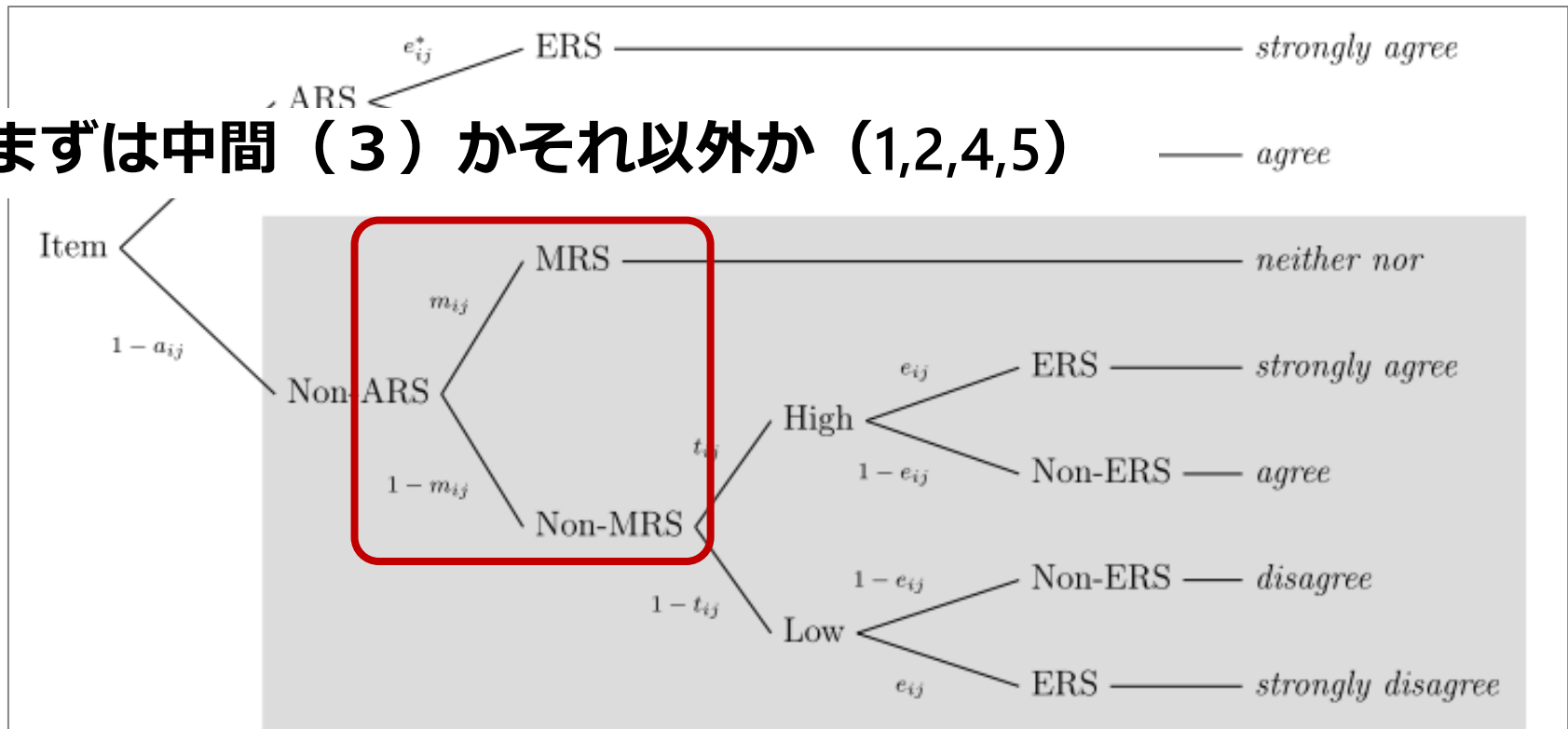


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

• RSのためのIRTree モデル

Plieninger, H. (2017). Fig3.1 より

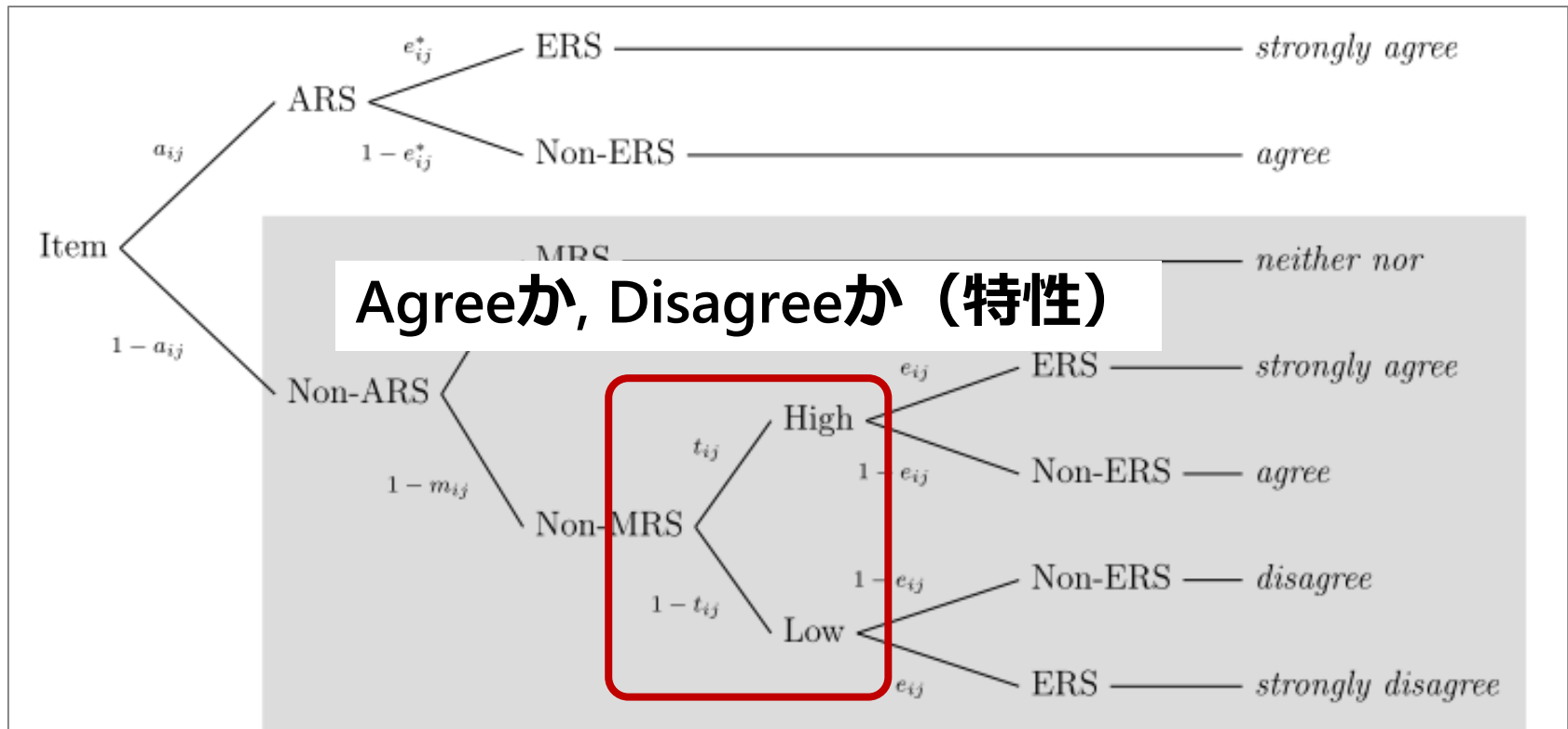


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

• RSのためのIRTree モデル

Plieninger, H. (2017). Fig3.1 より

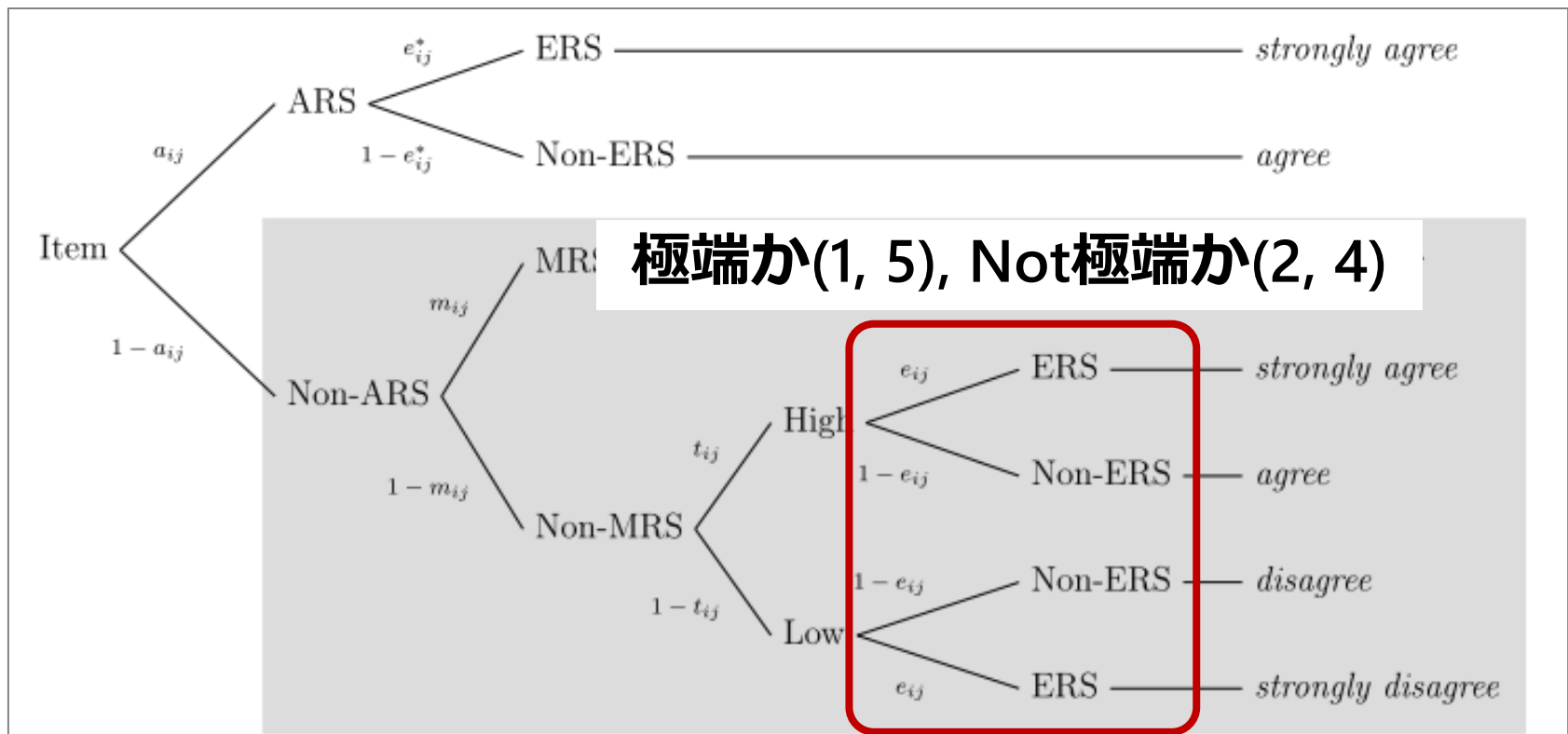
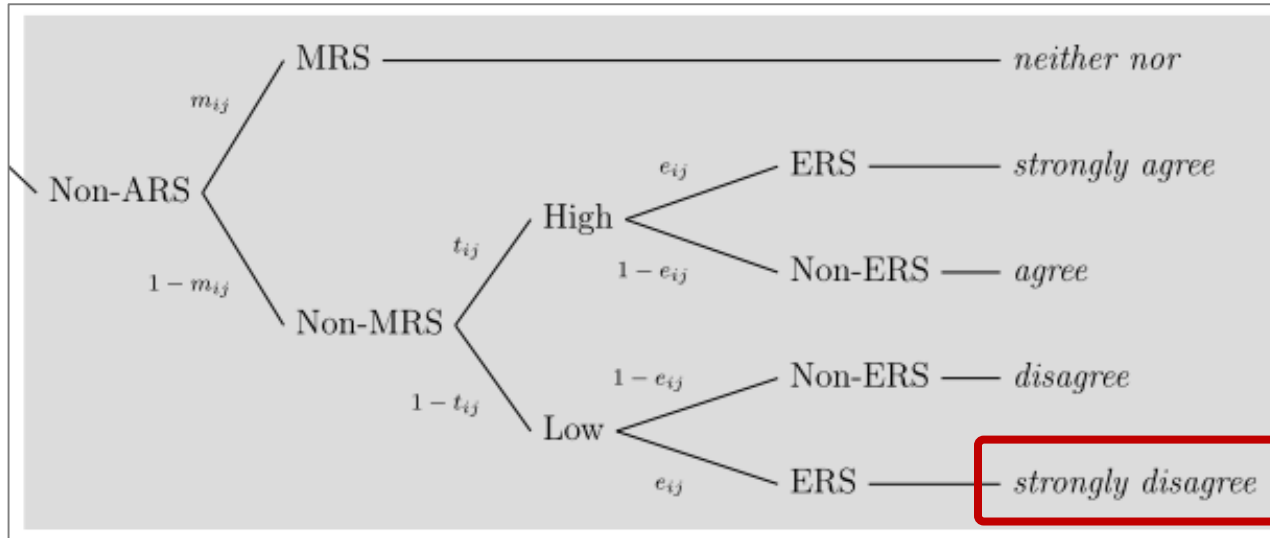


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

- 今回考えた回答過程モデルでは、



- 「strongly disagree」と回答するのは、
 - 中間 or Not のノードで **Not** を選ぶ確率 $(1 - MRS_{ij})$
 - Yes or No のノードで **No** を選ぶ確率 $(1 - trait_{ij})$
 - 極端 or Not のノードで **極端** を選ぶ確率 ERS_{ij}

を掛け合わせた確率

$$\text{strongly disagree} = (1 - MRS_{ij}) \times (1 - trait_{ij}) \times ERS_{ij}$$

- 全カテゴリで前スライドのように尤度を表現可能
- **IRTreeはMPT**(multinomial processing tree model)**の特殊形**
 - MPTは、項目・回答者の効果を分離できない
 - **IRTreeモデルは、MPTモデルの項目・回答者毎の解釈ができるモデル**と考えられる(MPTをよりIRTらしくしたモデル)
 - ただ、最近MPTも上記効果を検討可能に
 ➡階層MPTモデル
- **IRTree modelの制約**
 - 各カテゴリに許容されるパスは一つ
 - ERS or MRS **のみのモデル**しかなかった

• Plieninger, H. (2017). による拡張 • ARS対応

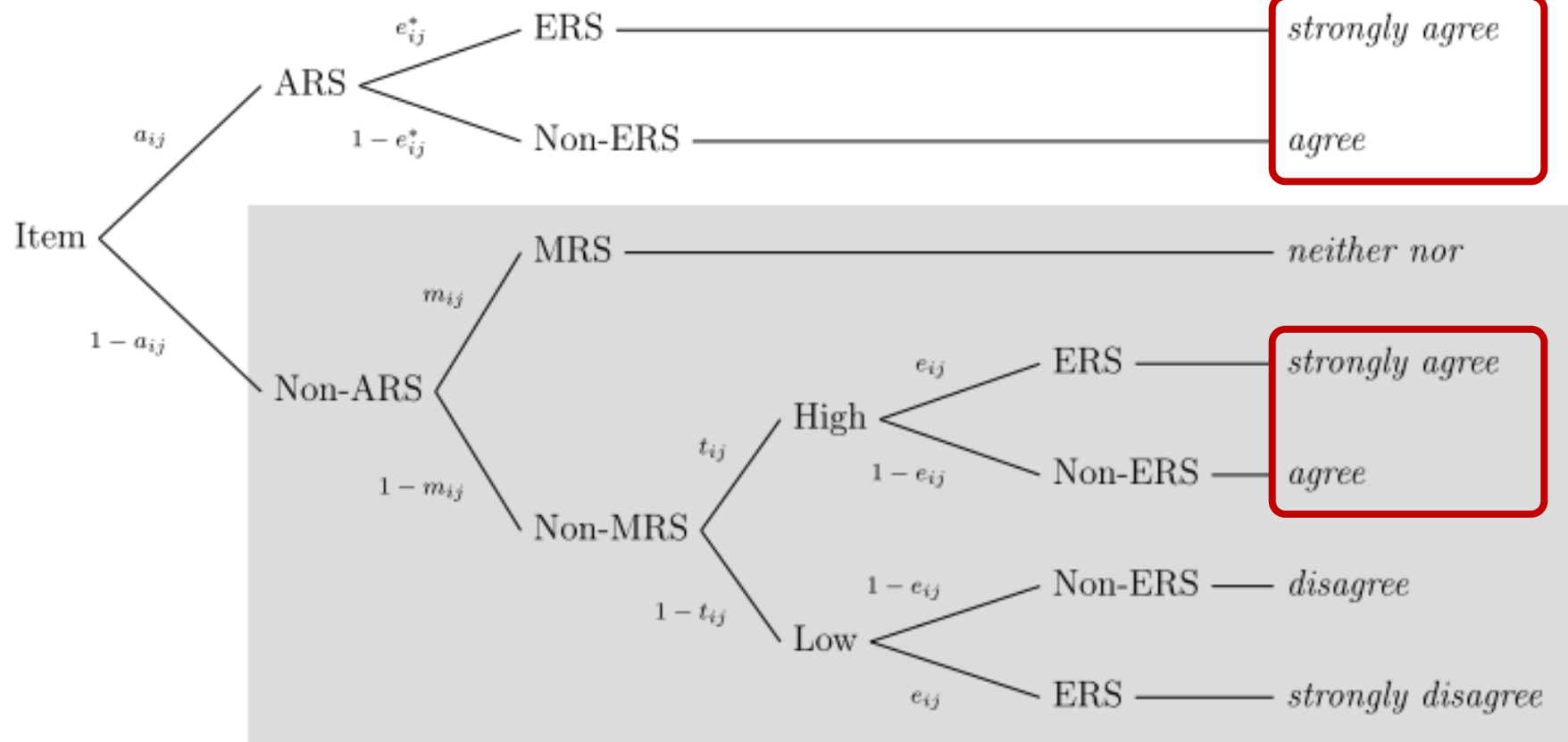


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

ARSの人、(項目全体で)閾値がシフトする

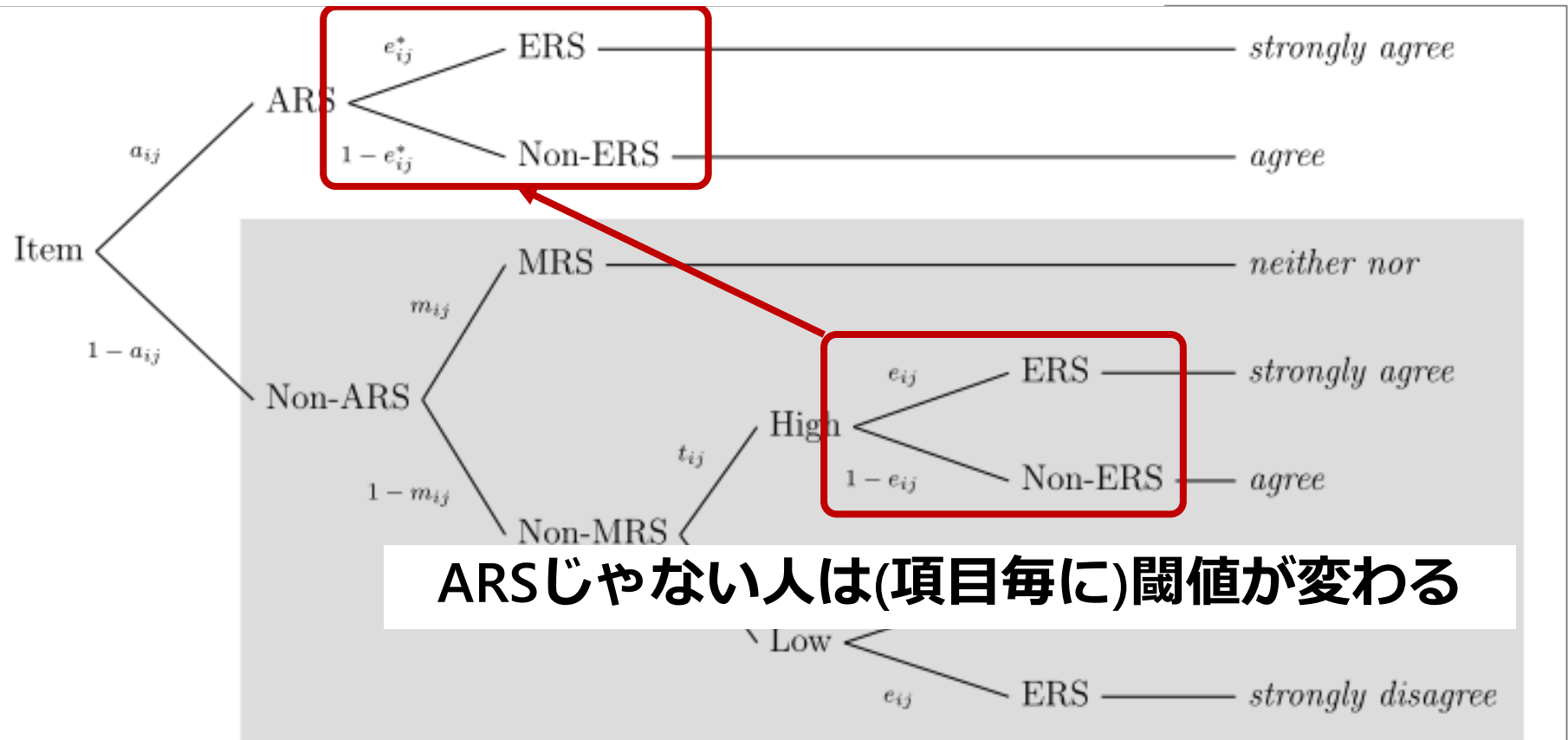
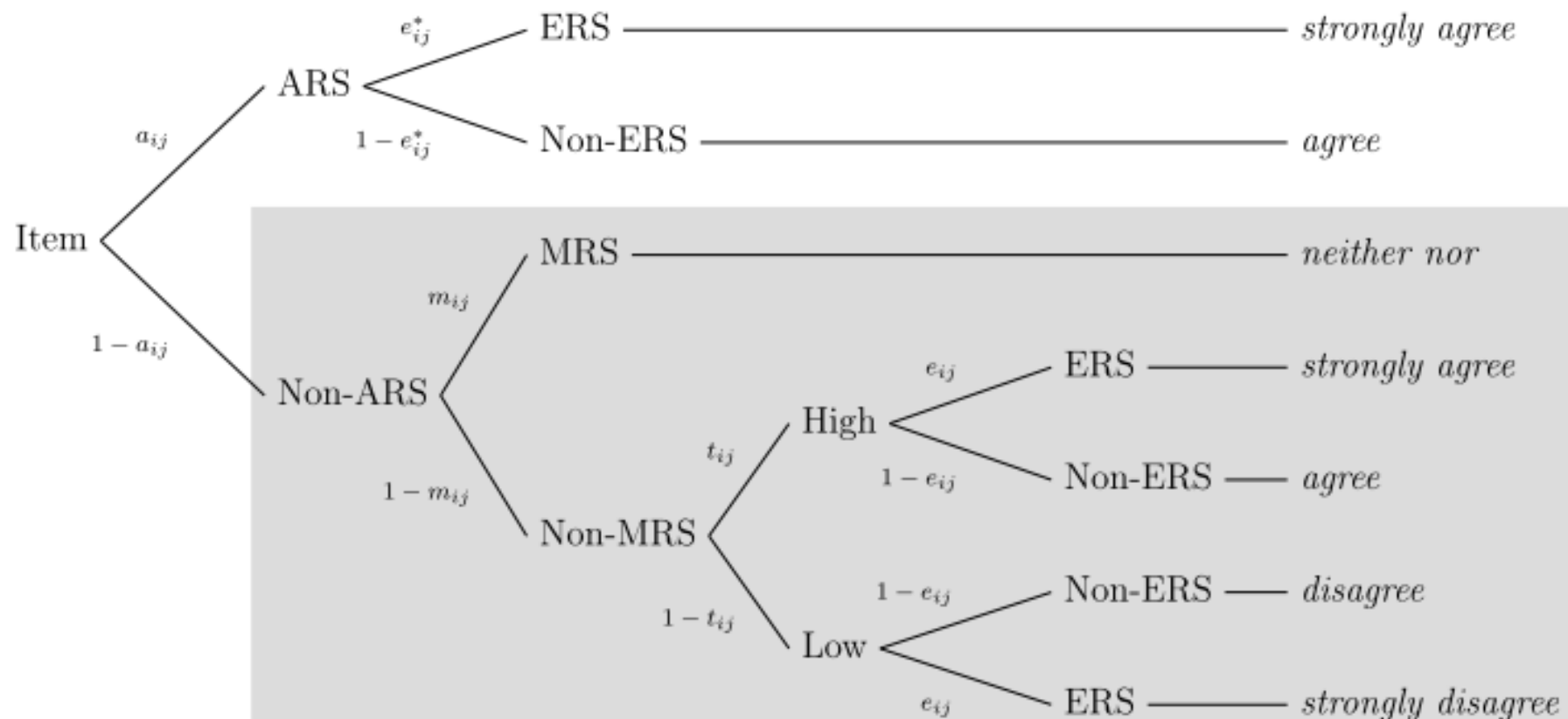


FIGURE 3.1: Tree diagram of the Acquiescence Model. The model includes the Böckenholt Model (depicted in gray) as a special case if $a_{ij} = 0$.

• そのため、

- ARS: $e_{ij}^* = \Phi(\theta_{ei} - \beta_{e^*})$
- Not ARS: $e_{ij} = \Phi(\theta_{ei} - \beta_{ej})$



- 次式のような形にもできるのでは？

$$f(x_{ij}) = \lambda_j \theta_{ti}^* + \theta_{ai}^* - \beta_j$$

係留ビネットのJonasの段階反応の考え方に近い

• IRTreeモデルについて

- 複数の回答過程を作って、比較する必要あり
- BridgeSamplingによるBayesFactor比較で構造探索できる？
Gronau, Wagenmakers, Heck, & Matzke, (2018)
- まだまだ、わかっていないこともたくさん
- だが、今後発展していくモデルの一つ
- 日本では、まだ誰もやっていないはず
- 僕もそのうちやりますw

- RS研究の歴史と動向
- RS(Response Style)の種類
- RSと様々な変数との関係性
- RSへの対処アプローチ
 - 統計モデル的アプローチ
 - 方法論的アプローチ
- RSについての3つの疑問
 - **Psychometric level** – RSを測定するには?
 - **Substantive level** – 何でRSが説明できる?何に影響される?
 - **Applied level** – RS除去したい(興味があるのは特性だけ)
- RSに対処するモデル
 - **Item Response Model**
 - **Item Response Tree Model**

- 広義のResponse Styleの歴史、これまでの研究を振り返った
- **方法からのアプローチ(DD/Likart/Likart with 係留ビネット法)**
- **モデルからのアプローチ(RIRSMACS/多次元名義反応モデル)**
- **そして、3つの目標・目的にどう対応していくかをまとめた**
 - 特性とRSの相関高い問題
 - ARSと特性の相関高くなってしまう問題
 - **この2つをどうにかするならば係留ビネットありかも**
- **Response Styleのモデリングしたいなら、IRTtreeは考えても良いかも！**

- Behr, D., & Shishido, K. (2016). 19 The Translation of Measurement Instruments for CrossCultural Surveys. The SAGE handbook of survey methodology.
- Böckenholt, U. (2017). Measuring response styles in likert items. *Psychological Methods*, 22(1), 69–83. <https://doi.org/10.1037/met0000106>
- Boeck, P. De, & Partchev, I. (2012). IRTrees: Tree-Based Item Response Models of the GLMM Family. *Journal of Statistical Software*, 48(Code Snippet 1). <https://doi.org/10.18637/jss.v048.c01>
- Bolt, D. M., Lu, Y., & Kim, J. S. (2014). Measurement and control of response styles using anchoring vignettes: A model-based approach. *Psychological Methods*, 19(4), 528–541. <https://doi.org/10.1037/met0000016>
- Cronbach, L. J. (1946). Response sets and validity. *Educational and Psychological Measurement*, (6), 672–683. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/001316444600600405>
- Gronau, Q. F., Wagenmakers, E.-J., Heck, D. W., & Matzke, D. (2018). A SIMPLE METHOD FOR COMPARING COMPLEX MODELS: BAYESIAN MODEL COMPARISON FOR HIERARCHICAL MULTINOMIAL PROCESSING TREE MODELS USING WARP-III BRIDGE SAMPLING. <https://doi.org/10.1007/s11336-018-9648-3>
- Jonas, K. G., & Markon, K. E. (2018). Modeling Response Style Using Vignettes and Person-Specific Item Response Theory. *Applied Psychological Measurement*, 014662161879866. <https://doi.org/10.1177/0146621618798663>
- 加藤健太郎, 山田剛史, & 川端一光. (2014). R による項目反応理論. 株式会社 オーム社.
- King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review*, 97(04), 567–583. <https://doi.org/10.1017/S0003055403000881>
- Okada, K., Hojo, D., Takahashi, Y. (revised) ...
- Plieninger, H. (2017). Towards a Deeper Understanding of Response Styles Through Psychometrics. Retrieved from [https://ub-madoc.bib.uni-mannheim.de/44325/1/Plieninger-2017-Towards a Deeper Understanding of Response Styles Through Psychometrics.pdf#page=87](https://ub-madoc.bib.uni-mannheim.de/44325/1/Plieninger-2017-Towards%20a%20Deeper%20Understanding%20of%20Response%20Styles%20Through%20Psychometrics.pdf#page=87)
- Plieninger, H. (2017b). Mountain or Molehill? A Simulation Study on the Impact of Response Styles. *Educational and Psychological Measurement*, 77(1). <https://doi.org/10.1177/0013164416636655>
- Plieninger, H., & Heck, D. W. (2018). A new model for acquiescence at the interface of psychometrics and cognitive psychology. *Multivariate behavioral research*, 1-22.
- Schneider, S. (2018). Extracting Response Style Bias From Measures of Positive and Negative Affect in Aging Research. <https://doi.org/10.1093/geronb/gbw103>
- 清水・稲増(in press). 政治的態度の母集団分布形状を推定する 統計モデリングアプローチ, 理論と方法.
- Tasaki, K., & Shin, J. (2019). Response styles as a threat to cultural equivalence. *The Japanese Journal of Psychology*. <https://doi.org/10.4992/jjpsy.90.18302>
- Thissen, D., & Cai, L. (in press). Nominal categories models. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (2nd ed.). New York, NY: Chapman & Hall.
- Vaerenbergh, Y. Van, & Thomas, T. (2012). Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion*, 25(2). Retrieved from <https://ijpor.oxfordjournals.org/content/early/2012/06/18/ijpor.eds021.full>
- Van Soest, A., & Vonkova, H. (2014). Testing parametric models using anchoring vignettes against nonparametric alternatives.
- Weijters, B. (2006). Response Styles in Consumer Research.
- Weijters, B., Schillewaert, N., & Geuens, M. (2008). Assessing response styles across modes of data collection. *Journal of the Academy of Marketing Science*, 36(3), 409–422. <https://doi.org/10.1007/s11747-007-0077-6>
- Wetzel, E., & Carstensen, C. H. (2015). Multidimensional modeling of traits and response styles. *European Journal of Psychological Assessment*.

