

インターネットを利用した
心理学実験・調査に関する
諸問題
より適切なデータを得るためには

三浦麻子（関西学院大学・大阪大学）
小林哲郎（香港城市大学）

オンライン調査・実験におけるSATISFICE

- 定義：オンライン調査・実験において参加者が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
- 例：論理的に不正解とならない中点回答（「どちらともいえない」や「わからない」など）を選択して教示文や調査項目を読み飛ばす
- 研究結果を歪める可能性

オンライン調査・実験におけるSATISFICE

- オンライン調査・実験で特に問題となる理由
 - 研究者が回答行動を直接監督することができない
 - まとまった額の報酬を得るために多くの調査・実験に参加するというインセンティブ
- 本講演の論点
 - Satisficeの発生状況
 - Satisficeの弊害
 - Satisficeへの対処

一連の研究

- 三浦麻子・小林哲郎 (2015a). 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.
- 三浦麻子・小林哲郎 (2015b). オンライン調査モニタのSatisficeはいかに実証的知見を毀損するか. 社会心理学研究, 31(2), 120-127.
- 三浦麻子・小林哲郎 (2016a). オンライン調査における努力の最小限化(Satisfice)傾向の比較:IMC 違反率を指標として メディア・情報・コミュニケーション研究, 1, 27-42.
- Miura, A., & Kobayashi, T. (2016b). Survey satisficing inflates stereotypical responses in online experiment: The case of immigration study. *Frontiers in Psychology*, 7:1563.
- 三浦麻子・小林哲郎 (2016c). オンライン調査におけるSatisficeを検出する技法:大学生サンプルを用いた検討 社会心理学研究, 32(2), 123-132.
- 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), 1-11.
- Miura, A., & Kobayashi, T. (2019; in press). Survey satisficing biases the estimation of moderation effects. *Japanese Psychological Research*, 61(3).

心理学ワールド小特集

[小特集]クラウドソーシング—心理学データの新しい姿	クラウドソーシングを使った研究が気になっているあなたへ	眞嶋 良全（北星学園大学）
	機械仕掛けのトルコ人，人間仕掛けのクラウドソーシング	五十嵐 祐（名古屋大学）
	クラウドソーシングで得られるデータの質	三浦 麻子（関西学院大学）
	日本の研究者がクラウドソーシングを使いこなすには	伊藤 言（株式会社イデアラボ）

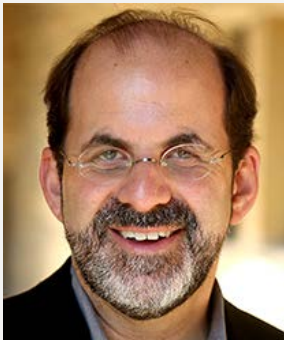
<https://psych.or.jp/publication/world082/>

意思決定において、完全情報、無限の計算能力を備えた全知全能的合理性下であれば、最適な代替案の選択が可能であるが、それらの条件が満たされない限定合理性の下では、要求水準を越えた選択は不可能であるため、限られた代替案の中から選択する行動。



Herbert Simon

satisfy + suffice = satisfice



John Krosnick

調査回答に際して、回答者が応分の注意資源を割かずに、最善の選択肢ではなく満足できる選択肢を求める行動。

教示文を読まない

- 研究者が見たいのは「ビールがどこで売られているか」という状況設定の効果なので、隅々までちゃんと読んでくれないと話にならないのだが、果たしてあらゆる参加者にそんなことを期待してもいいのだろうか？

You are on the beach on a hot day. For the last hour you have been thinking about how much you would enjoy an ice cold can of soda. Your companion needs to make a phone call and offers to bring back a soda from the only nearby place where drinks are sold, which happens to be a **run-down grocery store**. Your companion asks how much you are willing to pay for the soda and will only buy it if it is below the price you state. How much are you willing to pay?

You are on the beach on a hot day. For the last hour you have been thinking about how much you would enjoy an ice cold can of soda. Your companion needs to make a phone call and offers to bring back a soda from the only nearby place where drinks are sold, which happens to be a **fancy resort**. Your companion asks how much you are willing to pay for the soda and will only buy it if it is below the price you state. How much are you willing to pay?

Thaler's (1985) beer pricing task

質問項目を読まない

- 研究者は「あまり考え込まずに思ったとおりを答えて下さい」と言いつつ行動志向から個人の特徴的な傾性を取り出したいと思っているので、質問項目をちゃんと読んでくれないと話にならないのだが、果たしてあらゆる参加者にそんなことを期待してもいいのだろうか？

	あてはまらない	あてはまらない あまり	どちらとも いえない	あてはまる 少し	あてはまる
人と話すときにはできるだけ自分の存在をアピールしたい					
自分が注目されていないと、つい人の気を引きたくなる					
大勢の人が集まる場では、自分を目立たせようとはりきる方だ					
高い信頼を得るため、自分の能力は積極的にアピールしたい					
初対面の人にはまず自分の魅力を印象づけようとする					
人と仕事をするとき、自分の良い点を知ってもらうように張り切る					
目上の人から一目おかれるため、チャンスは有効に使いたい					
責任ある立場につくのは、皆に自分を印象づけるチャンスだ					
皆から注目され、愛される有名人になりたいと思うことがある					

SATISFICEの発生状況

- 三浦・小林(2015a)：国内のオンライン調査会社2社を対象に、Instructional Manipulation Check (IMC)およびDirected Questions Scale (DQS)を用いてSatisficeの程度を調査

Sports Participation

Most modern theories of decision making recognize the fact that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. In order to facilitate our research on decision making we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, then some of our manipulations that rely on changes in the instructions will be ineffective. So, in order to demonstrate that you have read the instructions, please ignore the sports items below, as well as the continue button. Instead, simply click on the title at the top of this screen (i.e., "sports participation") to proceed to the next screen. Thank you very much.

Which of these activities do you engage in regularly?
(click on all that apply)

skiing

soccer

snowboarding

running

hockey

football

swimming

tennis

basketball

cycling

Continue

- オリジナルのIMC設問 (Oppenheimer et al. (2009) のFig. 1)

あなたの日常的な行動についておたずねします

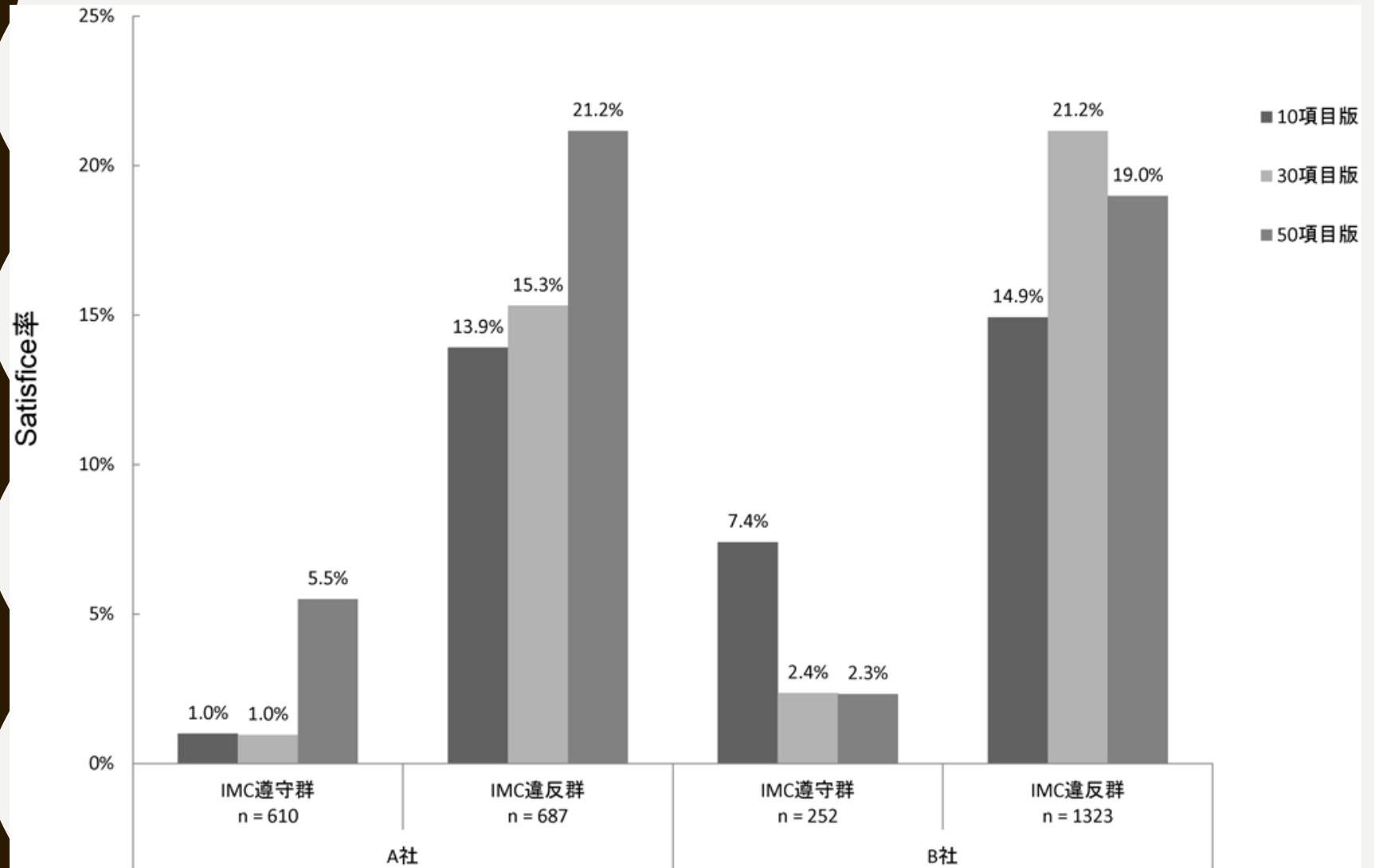
人間の意思決定に関する近年の研究で、人間の決定は「真空」状態でおこなわれるものではないことが知られています。個人の好みや知識、そしてその人がそのときどんな状況にあるかが、意思決定過程に重要な影響を及ぼすのです。われわれはこうした意思決定過程の研究のため、あなたの意思決定者としてのある要素を知りたいと考えています。つまり、あなたがこの指示を時間をかけてよく読んでいるかどうかに興味があるのです。もし誰もこの指示をお読みになっていないとしたら、指示内容を変えることが意思決定に与える影響を見たい、というわれわれの試みは効果を持たないからです。そこで、あなたがこの指示をお読みになったなら、以下の質問には回答せずに（つまり、どの選択肢もクリックせずに）次のページに進んで下さい。よろしくお願いします。

	あてはまらない	あまりあてはまらない	どちらともいえない	ややあてはまる	あてはまる
1. さまざまな意見を聞いたり議論したりすることが楽しい	1	2	3	4	5
2. 政治や経済など、社会の出来事や状況に常に興味を持っている	1	2	3	4	5
3. 自分の知識や経験を社会のために生かしたい	1	2	3	4	5

注：レイアウトやフォント等は調査会社により異なる

SATISFICEの発生状況 (三浦・小林, 2015A)

- IMCでは対象とした2社のうちA社では51.2%、B社では83.8%のサンプルが指示に従わなかった
- DQS：後続のリッカート尺度の測定において「この項目は一番左（右）を選択してください」という項目を挿入
 - 尺度は10項目、30項目、50項目の3水準
 - 回答者を無作為配置



SATISFICEの発生状況 (三浦・小林, 2016)

あなたの日常的な行動についておたずねします。

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。ここからがお願いです。この指示をお読みになったことの証明として、実際のあなたがどうであろうが、以下の質問には「はい」と回答して、次のページに進んで下さい。よろしくお願いします。

私は電子メールを使ったことがない

はい

いいえ

わからない

>>

SATISFICEの発生状況 (三浦・小林, 2016)

調査	調査時期	調査協力依頼対象	IMC設問
1	2014年8月	ネット調査会社A	5件法3項目の無回答
2	2014年8月	ネット調査会社 B	5件法3項目の無回答
3	2015年5月	ネット調査会社 A	3肢1項目の選択肢指定
4	2015年12月	ネット調査会社 A	3肢1項目の選択肢指定
5	2015年12月	クラウドソーシング (ランサーズ)	3肢1項目の選択肢指定
6	2016年1月	クラウドソーシング (クラウドワークス)	3肢1項目の選択肢指定
7	2016年1月	大学生参加者プール (関西学院大学)	3肢1項目の選択肢指定

SATISFICEの発生状況 (三浦・小林, 2016)

調査	調査協力依頼対象	違反率(%)	所要時間 (秒)		
			違反者	遵守者	差分
1	ネット調査会社 A	59.9	90.28(132.59)	134.45(158.88)	44.16**
2	ネット調査会社 B	82.5	91.78(136.16)	146.31(101.21)	54.53**
3	ネット調査会社 A	42.3	423.74(337.76)	495.73(303.90)	71.99**
4	ネット調査会社 A	44.2	251.02(221.55)	321.31(268.03)	70.29**
5	クラウドソーシング (ランサーズ)	19.5	388.70(277.65)	419.20(235.02)	30.50*
6	クラウドソーシング (クラウドワークス)	21.7	312.43(191.75)	378.19(200.42)	65.76**
7	大学生参加者プール	20.3	335.32(103.31)	373.61(139.58)	38.28+
カッコ内数値はSD					
** $p < 0.01$, * $p < 0.05$, + $p < 0.10$					

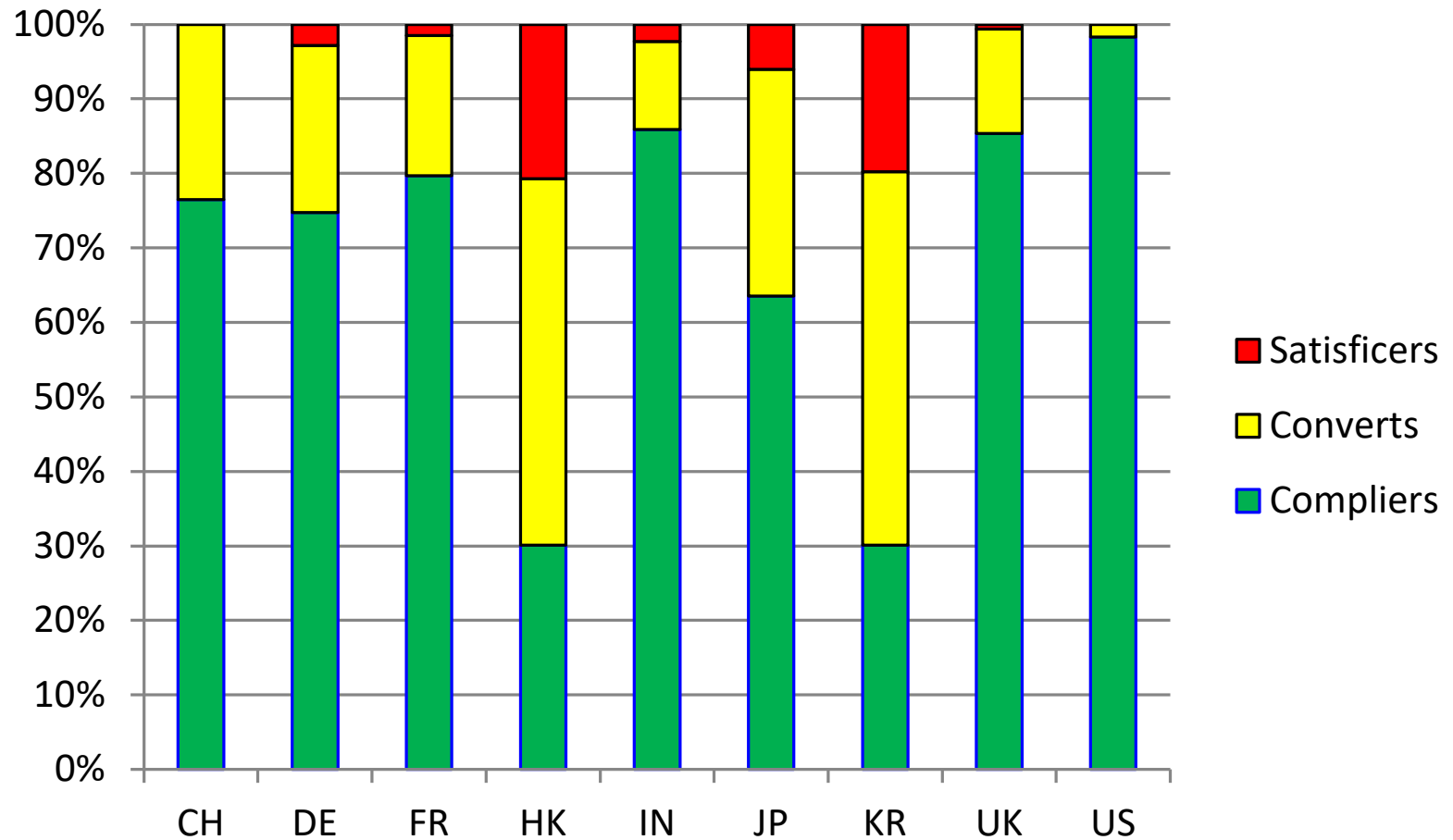
SATISFICEの国際比較

	対象	N
中国	オンライン調査会社	800
ドイツ	クラウドソーシング	819
フランス	クラウドソーシング	803
香港	オンライン調査会社	821
インド	クラウドソーシング	830
日本	クラウドソーシング	812
韓国	オンライン調査会社	800
イギリス	クラウドソーシング	807
アメリカ	クラウドソーシング	822

SATISFICEの国際比較

- IMCを実験の冒頭で最大2回繰り返して測定
 - Compliers : 1回目のIMCに対して適切に回答
 - Converts : 1回目のIMCでSatisficeするが、2回目では適切に回答
 - Satisficers : 2回ともSatisfice

SATISFICEの国際比較



調査会社とクラウドソーシングの違い

- 研究者が謝礼の支払をコントロールできるかどうか
- 研究者がSatisficerを直接punishできるかどうか
- 評判が将来的なタスクの請負に影響するかどうか

SATISFICEの弊害： 処置効果

- 三浦・小林（2015b）
 - 議題設定効果
 - プライミング効果
 - 動画を用いたオンライン実験
 - Satisficerは動画を「見飛ばす」ので処置効果が予測される形で検出されない可能性

従属変数：「政治とカネの問題」争点の重要度認知（処置後）	全体	見飛ばし群	優良群	最優良群
	Coef. (B)			
映像刺激の種類（処置）	-0.178 (0.143)	-0.037 -0.315	-0.089 (0.244)	0.109 (0.284)
報酬の種類（抽選）	0.187 (0.143)	0.182 -0.303	0.325 (0.245)	0.240 (0.288)
「政治とカネの問題」争点の重要度認知（処置前）	12.571** (0.468)	10.024** -0.929	13.866** (0.824)	14.283** (0.961)
ブロック変数 （1：1波IMC遵守ブロック、2：1波IMC違反ブロック）	-0.087 (0.144)	0.267 -0.387	-0.064 (0.250)	-0.083 (0.294)
カットポイント1	0.928** (0.249)	1.006 [†] -0.611	1.207** (0.430)	1.467* (0.677)
カットポイント2	4.179** (0.261)	2.952** -0.619	4.840** (0.468)	5.202** (0.718)
カットポイント3	7.671** (0.333)	5.966** -0.747	8.753** (0.615)	9.056** (0.857)
カットポイント4	10.952** (0.425)	8.808** -0.892	12.243** (0.781)	12.747** (1.034)
N	946	189	365	286
擬似決定係数	0.479	0.382	0.539	0.554

カッコ内数値は標準誤差

** $p < 0.01$, * $p < 0.05$, [†] $p < 0.10$

共変量は群ごとに異なるので省略した。

従属変数：現内閣の全体的評価（処置後）

	全体	見飛ばし群	優良群	最優良群
	Coef. (B)			
「政治とカネの問題」争点の評価（処置前）	5.477** (0.383)	2.493** -0.922	4.971** (0.671)	4.405** (0.781)
映像刺激の種類（処置）	0.151 (0.196)	0.759 [†] -0.452	-0.450 (0.316)	-0.515 (0.362)
「政治とカネの問題」争点の評価（処置前）×処置	-0.572 (0.484)	-1.666 -1.095	1.509 [†] (0.839)	1.962* (0.948)
報酬の種類（抽選）	-0.094 (0.120)	0.142 -0.28	-0.382 [†] (0.196)	-0.267 (0.229)
ブロック変数 （1：1波IMC遵守ブロック、2：1波IMC違反ブロック）	0.196 (0.120)	0.565 -0.364	0.182 (0.197)	0.181 (0.229)
カットポイント1	-0.448** (0.165)	0.817* -0.4	-0.587* (0.275)	-0.166 (0.578)
カットポイント2	0.972** (0.166)	2.938** -0.462	0.681* (0.276)	1.025 [†] (0.576)
カットポイント3	2.614** (0.184)	5.460** -0.567	2.210** (0.300)	2.509** (0.591)
カットポイント4	5.052** (0.238)	8.215** -0.719	4.913** (0.394)	5.190** (0.674)
N	946	189	365	286
擬似決定係数	0.134	0.296	0.156	0.156

カッコ内数値は標準誤差

** $p < 0.01$, * $p < 0.05$, [†] $p < 0.10$

共変量は群ごとに異なるので省略した。

SATISFICEの弊害：測定

- 三浦・小林 (2018)
 - Directed Questions Scale (DQS; Maniaci & Rogge, 2014)
 - 「この項目は「1. まったくあてはまらない」を選択して下さい」

因子	項目	小塩ら (2012)	調査 A		調査 B	
			遵守 (<i>n</i> =5318)	違反 (<i>n</i> =842)	遵守 (<i>n</i> =1173)	違反 (<i>n</i> =58)
外向性	活発で、外向的だと思う	−0.59	−0.44	0.08	−0.53	−0.05
	ひかえめで、おとなしいと思う					
協調性	他人に不満をもち、もめごとを起こしやすいと思う	−0.22	−0.20	0.17	−0.19	0.16
	人に気をつかう、やさしい人間だと思う					
勤勉性	しっかりしていて、自分に厳しいと思う	−0.38	−0.37	0.20	−0.42	−0.14
	だらしなく、うっかりしていると思う					
神経症傾向	心配性で、うろたえやすいと思う	−0.28	−0.30	0.18	−0.31	0.15
	冷静で、気分が安定している					
開放性	新しいことが好きで、変わった考えをもつと思う	−0.39	−0.28	0.17	−0.37	0.19
	発想力に欠けた、平凡な人間だと思う					

SATISFICEの弊害：測定

- 三浦・小林 (2018)
 - Satisficerは尺度項目を十分に注意して読んでおらず、そのことが解釈不能な項目間相関や因子構造をもたらしている
 - こうした信頼性の低い回答を含めて分析した場合に、本来検出されるべき因子構造が得られなかったり、分析に用いたとしても測定誤差が大きいために相関の希薄化が生じて仮説が適切に検証されないといった弊害があり得る

SATISFICEへの対処

- Satisficeの測定
 - IMCやDQSを用いて注意のレベルを測定
 - Satisficeのレベルごとに分析することが可能になる
 - 回答デバイスやページ滞在時間などのメタデータをバックグラウンドで測定
 - Qualtricsを用いると容易

SATISFICEへの対処

- Satisficeの抑制
 - 知見は一貫していない
 - Satisficerに注意のレベルを高めるように促してもデータの質は変わらない(e.g. Berinsky et al., 2016)
 - 日本ではどうか？

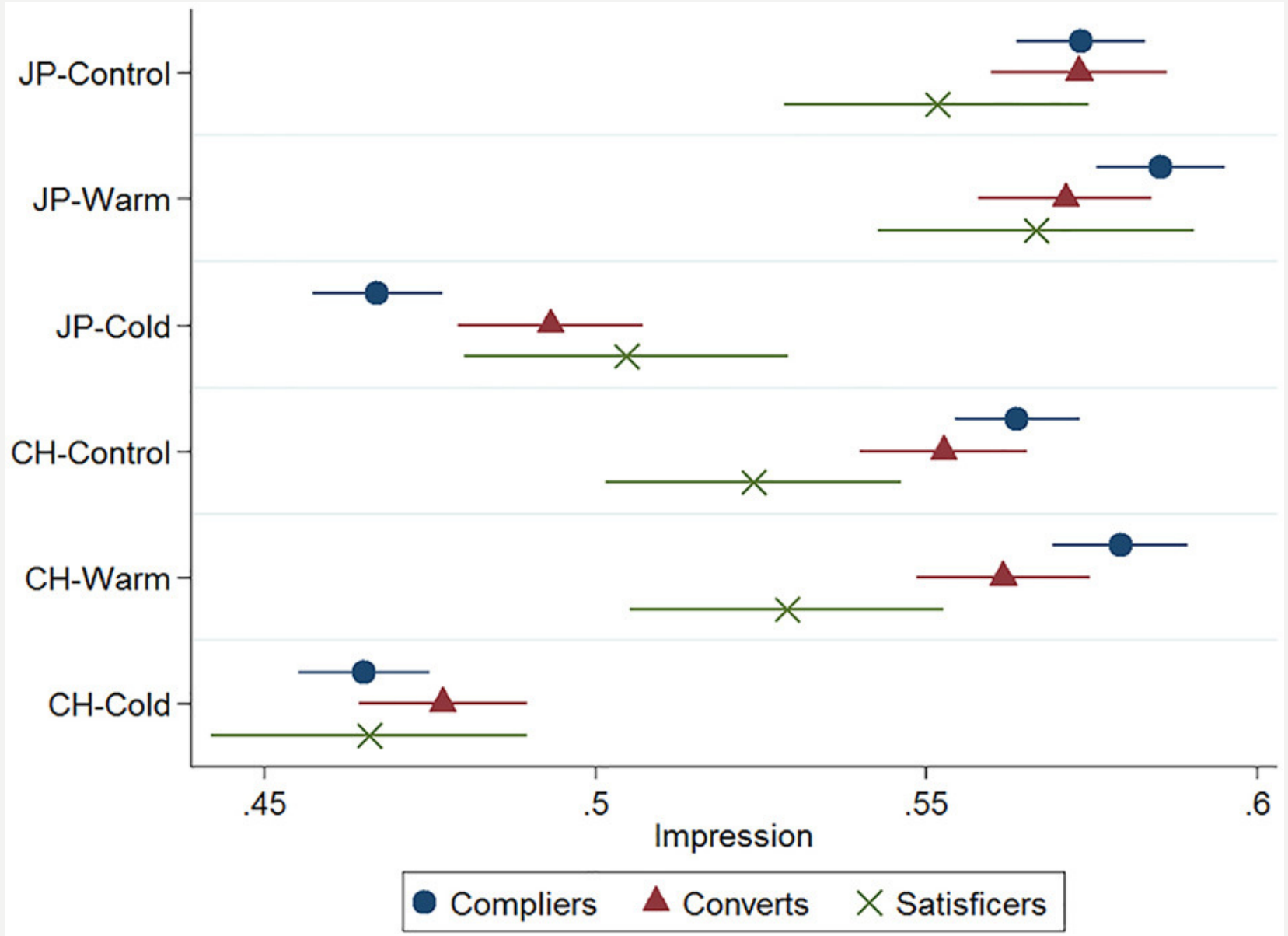
SATISFICERの「矯正」

- Miura & Kobayashi (2016)
 - オンライン調査会社のパネル
 - IMCを実験の冒頭で最大2回繰り返して測定
 - Compliers : 1回目のIMCに対して適切に回答
 - Converts : 1回目のIMCでSatisficeするが、2回目では適切に回答
 - Satisficers : 2回ともSatisfice

Liu Jian Ming is a Chinese person from Shanghai, who is aged 18. He is a freshman at the university in your local area and has just moved in next door to you, which means you will see him regularly.

He belonged to a brass band club when he was at school. According to his close acquaintances, he is a warm, industrious, intelligent, determined, and promising person.





SATISFICERの「矯正」

- IMCを繰り返すことで注意の低さを自覚した潜在的なSatisficersは注意のレベルを引き上げ、完全ではないがCompliersに近いレベルで後続の実験刺激にも注意を払う
- IMCはSatisfice傾向を測定するだけでなく、繰り返すことで潜在的なSatisficeを抑制することができる

結論

- 日本でもオンライン調査会社のパネルを中心にSatisficeは広範に生じている
- Satisficeは処置効果の推定を歪めたり尺度の測定に無視できないノイズを混入させる危険性がある
- 適切な測定項目やメタデータを用いることでSatisficeの発生を把握し、部分的には抑制できる