

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、つまり、インターネット上で多数の人々に業務を外注するサービスを利用することで心理学におけるデータの取り方に革命的な変化が起きています。今回の小特集でその一端に触れ、新しい時代の息吹を感じてください。(大久保街重)

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

北星学園大学社会福祉学部 准教授

眞嶋良全 (まじま よしまさ)

Profile—眞嶋良全

2000年、北海道大学大学院文学研究科行動科学専攻博士課程修了。博士(行動科学)。北海道大学助教、北星学園大学社会福祉学部専任講師を経て現職。専門は認知心理学(推論、判断)。著書は『心を測る』(分担執筆、八千代出版)など。



クラウドソーシングは、不特定の人(クラウド)に対して業務を外部委託(アウトソーシング)するという意味の造語で、最近注目を集めている新しい労働市場である。そこでは多種多様な仕事が行き取りされ、短時間で簡単にこなせるスキル要求の低い仕事は、特にマイクロタスクと呼ばれる。

マイクロタスク特化型サービスの代表が、アマゾン社のAmazon Mechanical Turk (MTurk)である。元々は、店舗の紹介写真として最も良いものを選ぶなどの、コンピュータによる自動処理には適さないが人には容易な小規模な作業を、少額の報酬で担ってくれる作業員(ワーカー)を探すという目的で開発されたサービスであったが、学術研究用の参加者プールとして利用する試みが拡大している。ある調査によると、IF=25以上の学術誌の中でMTurkを用いた論文数が、2011年からの5年間で約10倍に増加している(Chandler & Shapiro, 2016)。

読者の中には、何となく気にはなっている、そもそもクラウドソーシングが何なのか、どのように使うのかわからないという方

もいるだろう。本小特集の皮切りとして、クラウドソーシングを使った研究の基本の「き」から始めてみたい。

クラウドソーシング調査の流れ

クラウドソーシング自体は、人材募集から報酬の支払いまでのプロセスをワンストップで実現するサービスであり、タスク、特に調査・実験は、別の環境を用意して行うことが多い。十分なスキルがあれば自前のサーバでプログラムを走らせてもよいし、Qualtrics, SurveyMonkey等の調査環境構築ソフトウェア(リサーチ・スイート)もよく使われる(詳細は、本小特集の伊藤論文を参照)。

クラウドソーシングを用いた調査では、まずタスク募集を公開し、応じたワーカーをオンライン上の調査ページに誘導する。ワーカーには、PCやスマートフォンで回答させ、回答終了時にランダムな完了コードを発行する。最後に、クラウドソーシング側で正しいコードが入力されたことを確認の後、報酬の支払い処理を行う。

所要時間や回答の困難さなどによって報酬額は千差万別だが、およそ数十円~数百円に設定される

ことが多い(報酬額にまつわる問題については五十嵐論文に詳しい)。また、報酬額の20~40%程度の委託手数料が徴収されるのが一般的である。

クラウドソーシングの メリット・デメリット

クラウドソーシングの利点は、大規模で多様なサンプルに対する迅速かつ容易なアクセス、研究遂行の人的・時間的コストの大幅な削減、その結果として最終的な成果発表までの一連の研究サイクルを短時間で完了させることにある(Mason & Suri, 2012)。研究者がデータ収集や報酬の支払いのためにどこかに出向く必要はほぼなく、手元のPCでほぼ全てが完結する(経費精算のための事務処理の手間はもちろんあるが)。また、参加者の匿名性を保ちやすく、従来の対面式調査・実験に比べて個人情報保護のハードルが相対的に低いこと、調査項目や制御プログラムを公開することで追試が容易になることも副次的なメリットとして挙げられる。

一方で、クラウドソーシングにも限界はある。まず、サンプルの多様性である。心理学研究で

よく使われる「大学生」に比べれば、十分に多様であるとはいえ、クラウドソーシングサンプルにも一定の偏りはある。例えば、ワーカーは20～30歳代が圧倒的に多い。これは、経済・教育的理由から、情報端末の利用とネットアクセスに関するハードルが、この年齢層で特に低いからだろう。また、MTurkワーカーの多くは米国在住の白人で、近年はインド系のワーカーも増加しているものの、人口全体の人種・文化の分布を代表しているとは言えない。一方、本邦のクラウドソーシングのワーカーは大多数が日本人である。これは、サービスが提供するインターフェースの言語や、ワーカーに対して報酬を支払う際の決済システムの制約によると考えられる（例えば MTurkでは、米国在住者かインド人ワーカーでないと報酬を金銭で受け取れない）。

また、参加者のナイプさの問題もある。例えば、ワーカーは、より多くの報酬を得ようと同一の課題に何度も参加しようとするかもしれない。リサーチ・スイートやクラウドソーシングのチェック機能で一定程度は防げるが、同じワーカーが異なるIDで参加した場合や、類似課題を含む別の研究に参加したワーカーを除くことは難しい。さらに、ワーカー同士の情報共有がナイプさを損なう可能性もある。しかし、報酬に関する情報は共有されても、課題内容がリークされることは稀だと指摘もある（Chandler et al., 2014）。

技術面では、回答で使うブラウザの制約が挙げられる。例えば、認知実験に必要なミリ秒単位の刺激呈示・反応時間の取得は可能だが、100ms以下の高い精度が必要な場合は注意が必要である。また、リサーチ・スイートではキー

押しの反応時間を測定する機能までは持っていないことが多く、jsPsych (<http://jpspsych.org>) のようなライブラリを使ってプログラミングするか、Inquisit Webのような有償のサービスを利用する必要がある。

他にも、データ・セキュリティや、インフォームド・コンセント、デブリーフィングの難しさなど、対面型の調査には見られない新たな配慮事項も存在する（本小特集の五十嵐論文も参照のこと）。

クラウドソーシングデータの信頼性と本邦の現状

クラウドソーシングを使った研究が増えているとはいえ、そこで得られたデータの信頼性に疑問を持つのは当然である。しかし、MTurkを中心とした検証研究では、大学生や年齢がマッチするオフラインサンプルと同様の結果を得る、あるいは先行研究の結果を再現できたことが報告されており、十分に実用に耐えることが確認されている（Chandler & Shapiro, 2016）。さらに、教示操作チェック（IMC, 本小特集の三浦論文も参照されたい）への回答はワーカーのほうが優れているとする結果もある（Hauser & Schwartz, 2016）。

他のクラウドソーシングサービスについては、CrowdFlower, ProlificがMTurkと同等かそれ以上に質の高いサンプルであること（Peer et al., 2017）、我々の調査でも、国内のクラウドソーシングワーカーが、MTurkと同様の傾向を示すことが示されている（Majima et al., 2017）。一方で、MTurkで増加しているインド系のワーカーは、教育経験や誠実さという点では米国人ワーカーと遜色ないものの、IMCや反転項目の回答の質が低下することや、反

応の質が報酬額に影響されることなどが指摘されている。また我々の調査から、PCで回答するほうが携帯端末よりもIMCの成績が良いことも明らかになっている（Majima et al., 2017; 図1参照）。

クラウドソーシングを使った研究は、まだ歴史が浅く、未解決の問題が多数あることは事実であるが、この新しいツールを使うことで、本邦の心理学研究がさらに加速することを期待している。

文献

- Chandler, J., Mueller, P., & Paolacci, G. (2014) *Behav Res*, 46, 112-130. doi:10.3758/s13428-013-0365-7
- Chandler, J. & Shapiro, D. (2016) *Annu Rev Clin Psychol*, 12, 53-81. doi:10.1146/annurev-clinpsy-021815-093623
- Hauser, D. J. & Schwarz, N. (2016) *Behav Res*, 48, 400-407. doi:10.3758/s13428-015-0578-z
- Majima, Y., Nishiyama, K., Nishihara, A., & Hata, R. (2017) *Front Psychol*, 8:378. doi: 10.3389/fpsyg.2017.00378
- Mason, W. & Suri, S. (2012) *Behav Res*, 44, 1-23. doi:10.3758/s13428-011-0124-6
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017) *J Exp Soc Psychol*, 70, 153-163. doi:10.1016/j.jesp.2017.01.006

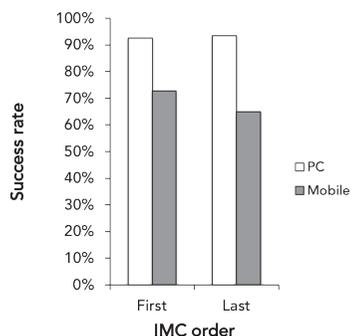


図1 PCと携帯端末によるIMC通過率の違い。PCのほうが一貫して高い。