Journal of Behavioral and Experimental Finance 17 (2018) 22–27

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Finance

journal homepage: www.elsevier.com/locate/jbef

Full length article

Prolific.ac—A subject pool for online experiments[☆]

Stefan Palan[a,b,*], Christian Schitter[a]

[a] Department of Banking and Finance, University of Graz, Universitätsstraße 15, 8010 Graz, Austria
[b] Department of Banking and Finance, University of Innsbruck, Universitätsstraße 15, 6020 Innsbruck, Austria

ARTICLE INFO

Article history:
Received 7 December 2017
Accepted 14 December 2017
Available online 21 December 2017

JEL classification:
C93
C81
B41

Keywords:
Prolific
Online experiment
Subject pool

ABSTRACT

The number of online experiments conducted with subjects recruited via online platforms has grown considerably in the recent past. While one commercial crowdworking platform – Amazon's Mechanical Turk – basically has established and since dominated this field, new alternatives offer services explicitly targeted at researchers. In this article, we present www.prolific.ac and lay out its suitability for recruiting subjects for social and economic science experiments. After briefly discussing key advantages and challenges of online experiments relative to lab experiments, we trace the platform's historical development, present its features, and contrast them with requirements for different types of social and economic experiments.

1. Introduction

1.1. Online experiments with subjects sampled from crowdworking platforms

The number of online experiments in the social sciences has surged in the last few years. In a recent article in *Science*, Bohannon (2016) reports that the number of published papers reporting social science experiments conducted with participants sourced via the most commonly used Amazon Mechanical Turk (*MTurk*) web platform grew from 61 in 2011 to more than 1,200 in 2015. The success of online experiments is not surprising, as they offer at-scale recruitment of participants in a short time, are generally cheap, and offer access to a broader population – potentially even representative of the internet population – than classical lab experiments with students (Paolacci and Chandler, 2014; Crump et al., 2013; Mason and Suri, 2012). Results from such online experiments also appear to offer reliability, as researchers have successfully replicated a range of well-known lab experiments from economics and psychology using subjects sourced via *MTurk* (Crump et al., 2013; Amir et al., 2012; Horton et al., 2011; Suri and Watts, 2011; Paolacci et al., 2010) and as *MTurk* workers also answer (basic) survey questions relatively consistently across experiments (Rand, 2012). Replication therefore appears to be possible as long as web-based technology is able to provide the accuracy and reliability needed for data collection in the specific task (Crump et al., 2013).

However, as *MTurk* and other crowdworking platforms were not explicitly designed for the scientific community, their use for experimental research entails some challenges. Not taking into account the somewhat "arbitrary" distinction between online subjects and other convenience samples like students (as, e.g., discussed in Landers and Behrend, 2015), the following methodological and technical challenges have recently gained attention:

First, a population of professional survey-takers may be evolving on crowdworking platforms. *MTurk*, for example, claims a pool of more than 500,000 workers, yet Stewart et al. (2015) find that the average lab samples, per quarter, a far smaller number of only 7,300 on the platform,[1] with many participants being sourced by several labs simultaneously. This could lead to loss of naivety. While recent research did not find experienced subjects to be a problem in common lab experiments (Benndorf et al., 2017; Kleinlercher and Stöckl, 2017), the effect of online subjects participating in potentially hundreds of studies remains to be quantified and has the potential to bias results of tasks which suffer from practice effects (Chandler et al., 2014). This problem is exacerbated by the existence of several discussion boards that *MTurk* workers

☆ The authors thank the team of Prolific.ac for providing detailed information about their platform, including utilization data. Funding by the Austrian Science Fund FWF (START-Grant Y617-G11 Kirchler) is gratefully acknowledged. Disclaimer: None of the authors is affiliated with Prolific.ac or received any compensation, financial or otherwise, for authoring this paper.

* Corresponding author at: Department of Banking and Finance, University of Graz, Universitätsstraße 15, 8010 Graz, Austria.
E-mail addresses: stefan.palan@uni-graz.at (S. Palan), christian.schitter@edu.uni-graz.at (C. Schitter).

1 Amazon only publishes registered, not active users. Therefore, estimating the population size actually available is difficult on this platform.

https://doi.org/10.1016/j.jbef.2017.12.004
2214-6350/

regularly use to discuss, and share information about, tasks. Experimenters can monitor these forums to determine whether their particular study is the subject of discussion, but there is little they can do to prevent such discussion from happening (Mason and Suri, 2012).

Second, there is a lack of control over the environment. This could potentially lead to undesired participant behavior, like devoting limited attention to the task. Chandler et al. (2014) report survey evidence that many participants watch TV or listen to music while working on *MTurk* tasks. Necka et al. (2016) also report that a non-negligible proportion of *MTurk* users self-reports multitasking while participating in experiments. However, Necka et al. (2016) find such self-reports of undesired behavior in *MTurk*, campus, and community samples alike, and, except for the incidence of multitasking being higher among *MTurk* participants, diagnose hardly any substantial differences between the different samples. In any case, researchers must be more cautious with online than with lab experiments when their design requires environmental factors to be stable between subjects (Crump et al., 2013; Horton et al., 2011).

Third, there are no clear standards for payment of workers on *MTurk* and many other crowdworking platforms. This can pose ethical problems in the form of potential exploitation of participants (Shank, 2016; Crump et al., 2013; Mason and Suri, 2012) and can also backfire in the form of low-quality reports (Bohannon, 2016). Payment sizes can also influence the decision of workers to participate, and can thereby bias participant selection, though in general data quality seems to be independent of reward size (Buhrmester et al., 2011). While ethical issues are either dismissed or seen as repairable on *MTurk* by Mason and Suri (2012), clear guidelines for researchers that are common when using university subject pools do not exist. Another issue with participants on crowdworking platforms is that they cannot be sure whether they are subject to deception or not. This can be an issue in fields which consider it crucial that participants are certain of not being deceived. An example would be economic science experiments, where deception is largely considered to be taboo (Cooper, 2014).

Fourth, crowdworking platforms generally do not verify participants' identities. This means that people could create multiple accounts and thereby participate in the same study multiple times. Many crowdworking platforms take measures to prevent participants from having more than one account, but there is always the possibility that such profiles exist. One experiment with *MTurk* workers investigating this aspect hints at the issue being comparably small, though: Less than 3% of participants behaved in a manner suggesting that they may have had more than one account (Chandler et al., 2014).

Fifth, we believe that there may be potential demand effects specifically on *MTurk*, as requesters (i.e., those commissioning work on the platform) can decide not to pay workers if they are unhappy with the delivered results. Rejections of submissions on *MTurk* are discretionary to the requester (Chandler et al., 2014). Rejecting a submission does not only affect the immediate payoff for workers, but also negatively affects workers' reputation via an acceptance score. Not fully knowing or believing the rules of a scientific study, participants may try to anticipate the results they believe requesters expect. They would then choose their actions in the experiment such as to maximize the chance of yielding precisely such results. Necka et al. (2016) find demand effects to exist across *MTurk*, campus and community populations, with up to a third of participants indicating that they tend to respond in a way that helps researchers find support for their hypothesis. However, only among the *MTurk* population do they find a substantial amount of subjects reporting that they try to participate in studies of researchers they already know. This is a sign of a general trust issue towards requesters on this platform, lending support to our reasoning that demand effects may be higher with *MTurk* than with other samples.

Sixth and last, the ease of entering and exiting online experiments could lead to unwanted and potentially selective attrition, both within a session and particularly in longitudinal studies (Horton et al., 2011; Shank, 2016). One remedy for this issue lies in having treatments which are sufficiently similar. In this case, effects of attrition should at least be identical and random across treatments, thereby allowing for the identification of treatment effects (Rand, 2012; Shank, 2016).

1.2. Prolific *as a dedicated research subject pool*

In lab experiments at universities, many issues surrounding subject pool management have in the past been thoroughly addressed by professionalizing this task, including the development of dedicated recruitment software (e.g., Greiner, 2015; Bock et al., 2012). Unsurprisingly, the administration of online subject pools now tends in a similar direction. There are, for example, third party services aimed at facilitating organizing the subject pool on *MTurk* (Litman et al., 2016), or techniques to manage the *MTurk* subject pool by individual experimenters (Chandler et al., 2014).

Contrary to these offerings, which try to make general crowdworking platforms better suited for the needs of the scientific community and easier to use, *Prolific*[2] is a recently established platform for online subject recruitment which explicitly caters to researchers. It combines good recruitment standards with reasonable cost, and explicitly informs participants that they are recruited for participation in research. Several thousand researchers have registered with *Prolific* to date, many of whom have successfully used it as a subject pool in different areas, like economics (e.g., Marreiros et al., 2017), psychology (e.g., Callan et al., 2016) or even food science (Simmonds et al., 2018). *Prolific* has detailed rules regarding the treatment of subjects on the platform, has a user-friendly interface, and has functionality that is a superset of *MTurk's*.

In a recent study, Peer et al. (2017) compared *Prolific* to *MTurk* and another crowdsourcing platform (*CrowdFlower*) as well as a university subject pool.[3] While the response rate was slightly lower on *Prolific* than on either *MTurk* and *CrowdFlower*, it was still superior to the university pool. In a range of tasks and experiments, *Prolific* and *MTurk* both managed to replicate existing results and delivered a higher data quality than *CrowdFlower* and the university subject pool. Given their choice of evaluation criteria (response rate, internal reliability, naivety, dishonesty) and explicitly mentioning response time as the main advantage of *MTurk*, Peer et al. (2017, 161) conclude: "... [*Prolific*] provides data quality that is comparable or not significantly different than *MTurk's*, and [*Prolific's*] participants seem to be more naïve to common experimental research tasks, and offer a more diverse population in terms of geographical location, ethnicity, etc. This suggests [...] [that] researchers who prefer naivety and diversity in their sample, could turn to [*Prolific*] if they are willing to wait some more for data collection to complete (depending on sample size)".

Prolific has grown significantly in the last years (see Fig. 1), which is likely to result in shorter waiting times by now. From January to the beginning of November 2017, *Prolific* had about 27,500 participants who had at least one accepted study submission. As of December 22nd, 2017, about 35,600 participants were listed as active on *Prolific's* website.[4] *Prolific's* growth is also reflected

2 www.prolific.ac.

3 The Center for Behavioral Decision Research participant pool managed by Carnegie Mellon University.

4 The number of active participants that is continuously published on *Prolific's* website counts signed-up subjects who have logged on to *Prolific* at least once in the last 90 days.

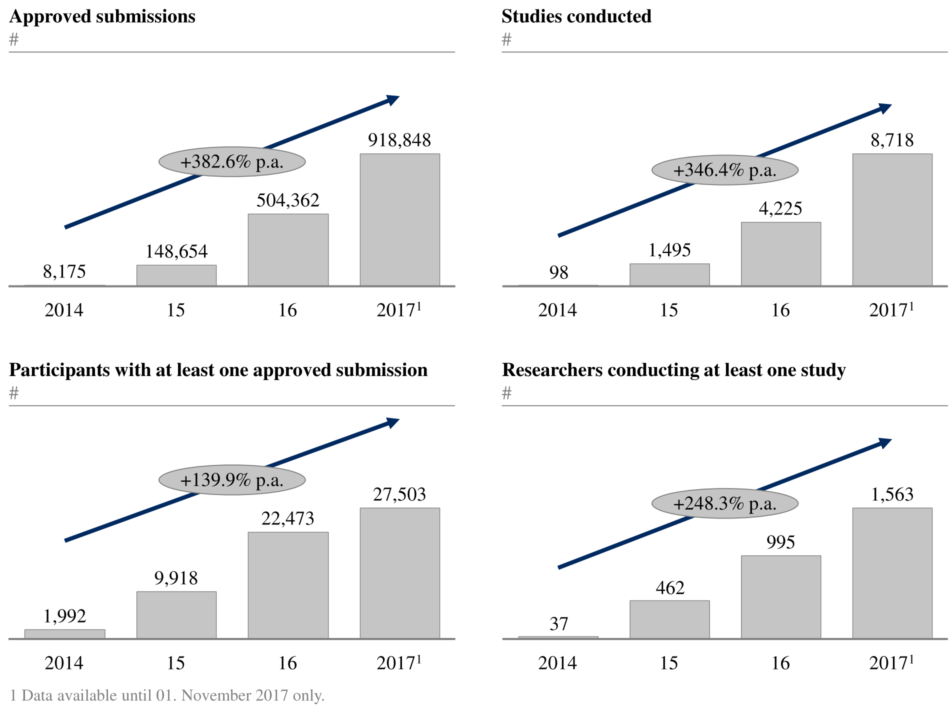

Fig. 1. Selected utilization statistics for *Prolific*.
Source: Prolific, private communication

in its increase in the number of researchers conducting at least one study, which was greater than 1,500 in 2017. The significant growth comes at a potential cost in terms of naivety, though: Fig. 1 shows that the number of submitted responses increases about 3 times as fast annually as the number of active participants. While a participant in 2014 only participated in 4.10 studies on average, this number grew to 33.41 in 2017. As discussed earlier, this is not necessarily a problem, but important to be aware of when designing a study.

We believe it to be helpful for other researchers to learn about this platform either as a replacement, or as a potential addition to *MTurk* and other crowdworking platforms in their toolkit. In the following, we therefore address key features of *Prolific* in terms of subject management, and discuss its suitability for different types of online surveys and experiments.

2. *Prolific* subject pool management features

2.1. Payment rules, return option and rejection guidelines

Clarity about rights, obligations and compensation of participants in scientific studies is a requirement both for ethical research and for the validity of results. As discussed in Section 1.1, using *MTurk* or other conventional crowdworking platforms might therefore be troubling from a research perspective, as these platforms' guidelines on these topics to both researchers and participants are often limited. This might lead to unreasonably low subject payment and create experimenter demand effects due to uncertainty about the expectations of, or due to mistrust towards, the experimenter.

Prolific, on the other hand, sets out very clear guidelines for the handling of submissions, and defines a minimum fixed payment per unit of time required to complete an experiment that is also communicated to participants when they sign up to the platform.[5] The time required for an experiment is initially estimated by the experimenter, but is then updated with the actual time taken once participants make submissions. The (continuously updated) minimum payment per unit of time is communicated to participants before they agree to participate in an experiment.

As is the case on most crowdworking platforms, rejections of a submission on *Prolific* negatively affect the *acceptance score*, a reputation score of the participant. Researchers have the option to filter for participants with a high acceptance score. A high rate of rejections therefore can mean that subjects cannot participate in many subsequent studies. However, rejections on *Prolific* have to be reasonable and can be overruled if a participant objects to a rejection and convinces *Prolific* that the rejection was unjustified.[6] Subjects on *Prolific* also have several options for terminating a study without negatively affecting their reputation score: They can decide not to finish the study at all by letting it time out (that is, not reporting completion to *Prolific*), or they can *return* their submission, thereby indicating that they wish the researcher not to use their data. In both cases, participants can, but do not have to be paid by the researcher. In either case (timing out or returning a submission), the participant's reputation remains unaffected. This way, participants on *Prolific* have a quick and risk-free option for withdrawing their consent at any time during the study.

Where the acceptance score is intended to deter participants from delivering submissions of low quality, *Prolific* also offers means to encourage participants to deliver submissions of particularly high quality. Researchers have the possibility to award up to 5% of participants a *star* when they deliver excellent submissions. Participants with many stars are then recognized and can win prizes.

While these points in our view support valid results and ethical research on *Prolific*, we find one potential caveat: The minimum hourly payment applies to the full subject pool, independent of the

5 At the time of writing, the minimum payment per hour was 5 GBP or 6.50 USD, with fractions of hours requiring proportionally smaller payments. An experiment taking 6 min would thus require a minimum payment of 0.50 GBP or 0.65 USD.

6 http://help.prolific.ac/managing-and-reviewing-submissions/reviewing-submissions/reviewing-submissions-how-do-i-decide-who-to-acceptreject, retrieved 22.12.2017.

cultural and regional background of subjects. 5 GBP per hour seems reasonable for the current user base which originates exclusively from OECD countries, but it might be a substantial payment for participants from some developing countries. Guidelines on ethical research explicitly rule out the use of excessive incentives which lure participants into studies they would otherwise prefer not to participate in. Researchers should take this into account when deciding whether to run studies with low-income populations on *Prolific* if the subject pool broadens to include such populations in the future.

2.2. Transparency about the population: Pre-screening

One of the key advantages of *Prolific* over other platforms is that researchers can pre-screen participants based on pre-screening questions used in earlier studies. If a study for example requires female subjects who have previously invested in the stock market, obtaining such subjects is not trivial on most platforms. One option is to advertise this criterion in invitation emails and consent forms. However, this carries significant risk of pollution in the sample stemming from dishonest participants. Another option is to elicit a subject's gender and stock market experience during the experiment. Subjects who do not fulfill the criteria can then be informed that the study does not require their participation, or can be excluded post-hoc. The downside to this strategy is that it is costly both in terms of subjects' time and in terms of researchers' funds, since subjects need to be paid at least a small participation fee. Furthermore, since many members of online platforms are networked and participate in online discussion groups regarding their platform, news can spread quickly, leading again to the risk of dishonest participants polluting the final sample.

Prolific solves this problem by eliciting subject characteristics independently of specific studies. When participants sign up to the platform, they are asked to provide such basic information as their gender and age. Both during the sign-up phase and afterwards, participants are also offered the chance to answer additional questions, which may make them eligible to participate in more studies. Taking the earlier example as an illustration, subjects have an incentive to answer the question regarding their stock-market participation because that makes them eligible to participate in studies which require an answer to this question. Since it is not a priori clear whether researchers will be more likely to look for subjects with or without stock market experience, there is little incentive to answer the question in such a way as to maximize one's chances of fitting into the recruiting pattern of future studies.[7]

Researchers can propose their own questions to *Prolific*, which allow them to condition on the answers in their subsequent recruiting. Since the answers to all of these questions remain on record at *Prolific*, they can furthermore be used by other researchers in the future, obviating the need to re-ask questions which subjects answered previously. This cuts down on the time subjects spend on answering similar questions and allows researchers to quickly run studies relying on subject characteristics already elicited before.

Researchers can also choose to only invite participants who have answered some pre-screening questions at all, independently of the answer they gave. This allows the researchers to then download these answers for all subjects who participated in a study, thereby reducing the time requirement for typical exit questionnaires (particularly if standard fields like age, gender or student status are concerned).

2.3. Options for exclusion of individual participants

Prolific only handles participant recruitment and payment. The researcher can use any web-based software to collect the actual data. The way a study is run on *Prolific*, no participant can make a submission to any one study that is running on *Prolific* twice.[8] However, if a study has to be repeated, or different treatments are to be run sequentially or cannot be implemented in the same survey or experiment, participants may need to be excluded from later participation. If experiments are run using the same account, there is a screener allowing for the exclusion of subjects who participated in specified previous studies. Furthermore, it is possible to create a blacklist, containing a list of participant IDs. Each of these IDs is unique to the specific *Prolific* user and does not change over time. These two options offer simple but effective tools to control participation in an experiment. Using these screening methods is also in line with *Prolific's* rules, which allow for barring subjects from participating in future experiments based on experiments they have participated in previously.

2.4. Options for longitudinal studies

The opposite to the blacklist discussed in Section 2.3 is the whitelist screener. When the researcher uses a whitelist, only participants with IDs entered into the whitelist are invited to participate in a study. This allows gathering information from the same subjects at different points in time. While we are not aware of attrition rates over longer time horizons, we have ourselves run an (as yet unpublished) experiment where we sampled 160 inexperienced (0 or 1 previous participations on *Prolific*) and 160 experienced ($\geq$60 participations on *Prolific*) subjects and invited them back 2 weeks later. We experienced 23.75% attrition for subjects with low experience, and 6.88% attrition for subjects with high experience. Although very thin evidence, this data suggests an overall acceptable attrition rate. This should hold true particularly in experiments without experience restrictions, as in this case more experienced than inexperienced subjects are likely to participate, which should consequently result in a relatively low attrition rate.

Note that participants themselves are not made aware of whether they were invited based on a whitelist or not. The information from *Prolific* to participants is the same as for any other study, as are invitation mails and reminders to eligible participants, which are sent every 48 h. Researchers who so desire can inform participants about the study being longitudinal in nature via the study description.

3. Conducting different types of experiments on *Prolific*

3.1. (Longitudinal,) individual, unincentivized experiments and surveys

The simplest way of using *Prolific* is to ask participants to conduct individual tasks and paying them a fee which is fixed in relation to the average time taken to complete the tasks. This is possible without limitations on *Prolific*. Payment handling is very easy in this case, as only the flat fee payments need to be approved. Funds to pay participants have to be topped-up before running the study, and remaining funds can be transferred back from *Prolific* if they are not needed anymore. Longitudinal surveys are also made possible through the use of participant IDs in the whitelist screener when setting up the study.

---

7 Note, however, that completely ruling out opportunistic answering of screening questions would require finding a way to verify subjects' answers, which is usually impractical or impossible in online subject pools.

8 Participants can theoretically fill in the survey more than once via the supplied link, unless the experimenter rules that out on the survey website. In any case, they will only be paid once by *Prolific*.

3.2. (Longitudinal,) individual, incentivized experiments

In line with *Prolific's* policy, participants always have to be paid a fixed fee per unit of average time subjects need for completing a study. However, it is easy to pay individualized bonuses. Thus, it is easy to run incentivized experiments as long as the variable payment only relies on individual decisions or results, as in honesty experiments with dice (e.g., Fischbacher and Föllmi-Heusi, 2013), or real effort tasks (e.g., Abeler et al., 2011). Bonuses are determined in the experimenter's software outside of *Prolific* and are then sent to participants via a dedicated function in *Prolific*.

3.3. Single-stage, interactive, incentivized experiments

Running interactive experiments (such as a public goods or a prisoner's dilemma game) is harder on crowdworking platforms in general than in in-person settings, because participants' decisions need to be matched. However, if the interaction only consists of one stage that can be evaluated asynchronously (such as in a one-stage dictator game), such studies can be conducted by matching participants and evaluating the result of the interaction ex-post. Payments from the interaction can then be paid via the bonus function.

3.4. Sequential, interactive, incentivized experiments

Interactive experiments with repeated interactions are more difficult than single stage interactive experiments. However, as long as interactions can be evaluated sequentially and, again, not simultaneously, the whitelist and the option of bonus payments can be jointly used to conduct such experiments on *Prolific*. Running such an experiment can be very tedious and time-consuming, though. Every stage of the interaction has to be run as a single experiment and can only be started after all participants have finished the previous session. Such a setup may therefore introduce a bias, as the time between stages allows participants to reflect on, or to forget about, earlier tasks. Such a study may also suffer from potentially increased attrition, since subjects are more likely to choose not to return to a new session of a multi-session study than to drop out during the single session of a conventional study. *Prolific* policy, however, allows for not paying participants bonuses if they have been made aware that they need to return to future parts of the experiment to cash in on these bonuses. While certainly not a perfect remedy, this provision increases the motivation to finish all stages eventually.

3.5. Simultaneous, interactive, incentivized experiments

Interactive experiments which require simultaneous decision-making (like, e.g., continuous-time asset markets) can be conducted using subjects from crowdworking platforms (and therefore also *Prolific*) by using a method explained in Mason and Suri (2012) and used, among others, in Suri and Watts (2011) and Arechar et al. (2017). Following this method, a researcher has to set up a panel of participants for the specific reason of running an interactive experiment. First, a short study is run in which participants are only asked whether they would like to take part in such a panel for simultaneous experiments. The result of this session are the *Prolific* IDs of interested participants. When an experiment session comes up, the panel is informed in advance via the *Prolific* messaging service that such an experiment is about to be conducted, including the starting time. At the set time, the study is started with a whitelist of the *Prolific* IDs in the panel and participants can select into the experiment, which starts in a virtual waiting room, already in the experimenter's software outside of *Prolific*. There, participants wait until enough subjects have shown up to start a session. While all of that is doable, *Prolific's* requirement to pay participants with respect to the time invested in the study can make such a waiting room costly, as substantial time may be required for enough participants to assemble.

Simultaneous experiments are currently not a focus of *Prolific* (and neither of other crowdworking platforms, for that matter). However, we see potential for *Prolific* to create a clear process for such online experiments (screeners to select into such a type of experiment, clear rules for the use of waiting rooms administered by *Prolific*, etc.) to extend the possibilities it provides.[9] Adding such a functionality would add easy-to-use opportunities for online social and economic experiments which we have not yet found elsewhere. In our view, this might be an opportunity for *Prolific* and the scientific community, as new software like *oTree* (Chen et al., 2016) makes browser-based experiments ever more easy and feasible.

4. Conclusion

With clear rules for both participants and researchers, *Prolific* is a valuable alternative to other crowdworking platforms in our view. *Prolific* benefits from transparency in several ways: On the one hand, subjects know that they are recruited to participate in research. They are aware about expected payments, treatment, rights and obligations in such an environment. On the other hand, researchers have higher transparency about the subject pool than on other platforms, and can screen it in a range of dimensions before inviting subjects.

Comparing *Prolific* to other platforms in terms of functionality, we find no shortcomings. In terms of usability, *Prolific* is in our view even superior to the most commonly used platform, *MTurk*. The only concern we have is connected to the minimum hourly payment, which is independent of the participant group and may be unreasonably high for low-income populations (if any become available), and may potentially make simultaneous experiments costly when using the "waiting room" method as described in Section 3.5.

In our view, one key asset for the scientific community of having *Prolific* is that providing a subject pool for research is the core of *Prolific's* business. While *MTurk* is not a focus-product of Amazon and has not seen much development in recent years, *Prolific* is constantly evolving and – if continuously successful – will probably not only expand its reach further but will also continue to implement functionality requested by its users. This is true for both participants and researchers. Its existence should also diversify the risk of the (hypothetical) possibility of a shutdown of *MTurk*, as mentioned by Bohannon (2016).

Summing up, we see that *Prolific* cannot solve all potential problems in online surveys and experiments with participants sourced via crowdworking platforms (particularly the question of verifying identities and environmental control), but that it is a large step towards a dedicated online participant pool for sound scientific research, with good prospects of further expanding its functionality.

References

Abeler, B.J., Falk, A., Goette, L., Huffman, D., 2011. Reference points and effort provision. Amer. Econ. Rev. 101 (2), 470–492.
Amir, O., Rand, D.G., Gal, Y.K., 2012. Economic games on the internet: The effect of $1 stakes. PLoS ONE 7 (2), 1–4.
Arechar, A.A., Gächter, S., Molleman, L., 2017. Conducting interactive experiments online. Experimental Economics. Forthcoming.

9 *Prolific* noted in private communication with us that they plan to provide such functionality in the future.

Benndorf, V., Moellers, C., Normann, H.-T., 2017. Experienced vs. inexperienced participants in the lab: Do they behave differently? J. Econ. Sci. Assoc. 3, 12–25.

Bock, O., Nicklisch, A., Baetge, I., 2012. hroot: Hamburg registration and organization online tool. WiSo-HH Working Paper.

Bohannon, J., 2016. Mechanical Turk upends social sciences. Science 352 (6291), 1263–1264.

Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's Mechanical Turk. Perspect. Psychol. Sci. 6 (1), 3–5.

Callan, M.J., Kim, H., Gheorghiu, A.I., Matthews, W.J., 2016. The interrelations between social class, personal relative deprivation, and prosociality. Soc. Psychol. Pers. Sci. 8 (6), 660–669.

Chandler, J., Mueller, P., Paolacci, G., 2014. Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. Behav. Res. Methods 46 (1), 112–130.

Chen, D.L., Schonger, M., Wickens, C., 2016. oTree - An open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Finance 9, 88–97.

Cooper, D.J., 2014. A Note on deception in economic experiments. J. Wine Econ. 9 (02), 111–114.

Crump, M.J.C., McDonnell, J.V., Gureckis, T.M., 2013. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. PLoS ONE 8 (3).

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguise-an experimental study on cheating. J. Eur. Econom. Assoc. 11 (3), 525–547.

Greiner, B., 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1, 114–125.

Horton, J.J., Rand, D.G., Zeckhauser, R.J., 2011. The online laboratory conducting experiments in a real labor market. Exp. Econ. 14 (3), 399–425.

Kleinlercher, D., Stöckl, T., 2017. On the provision of incentives in finance experiments. Exp. Econ. Forthcoming.

Landers, R.N., Behrend, T.S., 2015. An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. Ind. Organ. Psychol. 8 (2), 142–164.

Litman, L., Robinson, J., Abberbock, T., 2016. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. Behav. Res. Methods 49 (2), 433–442.

Marreiros, H., Tonin, M., Vlassopoulos, M., Schraefel, M.C., 2017. Now that you mention it: A survey experiment on information, inattention and online privacy. J. Econ. Behav. Organ. 140, 1–17.

Mason, W., Suri, S., 2012. Conducting behavioral research on Amazon's Mechanical Turk. Behav. Res. Methods 44 (1), 1–23.

Necka, E.A., Cacioppo, S., Norman, G.J., Cacioppo, J.T., 2016. Measuring the prevalence of problematic respondent behaviors among MTurk, campus, and community participants. PLoS ONE 11 (6), 1–19.

Paolacci, G., Chandler, J., 2014. Inside the Turk: Understanding mechanical Turk as a participant pool. Curr. Dir. Psychol. Sci. 23 (3), 184–188.

Paolacci, G., Chandler, J., Ipeirotis, P., 2010. Running experiments on Amazon Mechanical Turk. Judgm. Decis. Mak. 5 (5), 411–419.

Peer, E., Samat, S., Brandimarte, L., Acquisti, A., 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. J. Exp. Soc. Psychol. 70, 153–163.

Rand, D.G., 2012. The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. J. Theoret. Biol. 299, 172–179.

Shank, D.B., 2016. Using Crowdsourcing Websites for Sociological Research: The Case of Amazon Mechanical Turk. Am. Sociol. 47 (1), 47–55.

Simmonds, G., Woods, A.T., Spence, C., 2018. 'Show me the goods': Assessing the effectiveness of transparent packaging vs. product imagery on product evaluation. Food Qual. Pref. 63, 18–27.

Stewart, N., Ungemach, C., Harris, A.J.L., Bartels, D.M., Paolacci, G., Chandler, J., 2015. The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. Judgem. Decis. Making 10 (5), 479–491.

Suri, S., Watts, D.J., 2011. Cooperation and contagion in web-based, networked public goods experiments. PLoS ONE 6 (3).