09/26/2017

The new focus on randomization inference for experimental findings. Take a look at this very nice blog post by Jason Kerwin from the University of Minnesota on the topic. His post explains how and why researchers are increasingly turning to randomization inference for calculating p-values and conducting tests of statistical significance from experiments. Very approachable and highly likely to be the way all of our experimental papers (whether lab or field experiments) will be reporting results before long. Here also is a link to a new Stata journal article on implementation of the idea in Stata.

10/02/2017

This week's BRITE idea is inspired by a tweet from this week's speaker, Robert Dur , who highlighted this recent working paper . The authors use data on big marketing experiments from Facebook and test whether cutting-edge observational econometric techniques can replicate the treatment effects estimated from the randomized experiments. The punchline is "no" -- even with big data they needed the experiment. This is an interesting way of comparing methods. Justin would personally love to see what happens if you have a similar test comparing "quasi-experimental" observational methods, like diff-n-diff or regression discontinuity.

10/19/2017

This week's BRITE Idea comes from Justin's former coadvisor (Stefano DellaVigna, Berkeley) and frequent coauthor (Devin Pope, U Chicago) who have an interesting paper forthcoming in the Journal of Political Economy. They ran an experiment testing the effects of different monetary and non-monetary incentives on effort in a simple task. They also had a bunch of experts (profs in Econ and psychology) predict the results as well as "non-experts". They find that experts on average predict results quite well, but interestingly it really relies on wisdom of the crowds — basically no individual expert was a particularly good forecaster. Interesting read. Also good to note how they got two top hits out of this study — one on the main results of what motivates and one on how well people can predict the results — clever approach to leveraging experimental data.

10/26/2017

For this week's BRITE ideas we want to highlight a new journal dedicated to replication studies in microeconomics. The link HERE gives all the details. While it is not entirely clear how much focus they will have on experimental findings, as opposed to analysis of other existing data sets, Justin's read is that there will be at least some scope for studies focused on careful replications in experimental lab studies. Importantly, that also seems to include studies that get at the same idea as previously published work but are testing robustness in different populations, etc...

11/01/2017

There is a fascinating new paper accepted at *Management Science* by Matteo Galizzi and Daniel Navarro-Martinez on the external validity of lab-experimental Social Preference Games, such as the dictator, public goods, ultimatum and trust games. Their stated research question is "to what extent do experimental social preference games tap into principles governing social behavior when it is put in context of taken outside the lab?" They are clear that they are not asking the more general question about whether lab experiments have value and, in fact, state that in their view lab experiments are "not only useful but also necessary in the social and behavioral sciences."

They have people do a series of the classic and popular lab experimental games. They then also elicit behavior in some field social contexts related to giving money and helping others and also to self-reports of past measures of pro-social tendencies. Here is their headline conclusion: "none of the behaviors elicited in the field or reported from the past were explained to a significant extent by behavior in the experimental games."

Very important research for those of us who sometimes use these games to consider – important to weigh whether what we are doing is useful only based on the internal validity within the lab environment vs. the extent to which the external validity of these measures is crucial.

11/08/2017

For this week's BRITE ideas we want to point you to a very interesting new paper by Ben Enke from Harvard and Florian Zimmermann at Zurich, who Justin thinks are two of the brightest young minds in behavioral economics. Their paper is called “Correlation Neglect in Belief Formation” and uses a set of laboratory experiments to test the ways in which people neglect the correlation in the signals they get.

The motivation is to think about things like social media where we might here the same “stories” or get the same “signals” from many people in my network, but they are all drawing on the same underlying piece of news. Here is how they describe their contributions:

“ First, we provide clean evidence that in a relatively simple and completely transparent setting people neglect correlations in information sources when forming beliefs, albeit with a strong heterogeneity at the individual level. As a consequence, just like recent models of boundedly rational social learning predict, people’s beliefs are excessively sensitive to well-connected information sources and hence follow an overshooting pattern. Second, we develop a series of treatment variations to uncover that people do in principle possess the mathematical and computational skills that are necessary to process correlated information in our setting. However, when the informational environment is sufficiently complex, many people exhibit conceptual problems in identifying and thinking through the correlation in the first place. As a consequence, exogenously shifting subjects’ focus towards the correlation and the underlying independent signals has large effects on beliefs.”

Justin thinks, in particular, that this is an interesting case where forms of “de-biasing” seem to potentially be fairly powerful.

11/27/2017

This week we want to highlight a nice discussion of a recent "Registered Replication Report"; link available HERE . At this point most people are probably aware that there has been a strong push to work on large-scale, multi-lab replications of key findings in psychology and behavioral/experimental economics. This talks about an example of trying to replicate simple "priming" effects. To put it simply -- they really did not replicate. Justin's personal view is that priming effects are one of the weaker effects in the psychology literature and something that BRITE Lab users should be especially wary/careful about going forward.

12/01/2017

This week we want to highlight a new journal launched by the American Economic Review called *American Economic Review: Insights* . This is a brand new journal soliciting for shorter papers, but still of the quality at the *AER*. The length requirement is under 6,000 words, which is about 15 pages or about 1/3rd the length of typical *AER* papers.

Justin's hunch is that the two types of things that will be published a lot here are experimental results and theory. So for those of you doing nice experiments with important insights that can be conveyed succinctly, consider submitting to this new journal. The editor is Amy Finkelstein , who of course is amazing. Best of all, perhaps, is that they are guaranteeing quick turn around and an initial reject or "conditional accept" decision rather than a huge round of revisions.

01/29/2018

For this week's BRITE Ideas we want to highlight a new working paper by Gathergood, Mahoney, Stewart and Weber titled " How Do Individuals Repay Their Debt? The Balance-Matching Heuristic ". The authors analyze a large data set on credit cards and see how people allocate their monthly payments between cards. The financially efficient thing to do is to allocate payments above the minimum-required payment to the card with the highest interest rate. But people do not do that and are not responsive to the interest rates on their cards.

The main contribution of the paper is that the authors explore a few different heuristics people might be using in deciding how to allocate their payments. They find that a "balance-matching heuristic" where you allocate your share of total payments to each card in proportion to the total balance on that card best fits the data. The reason we wanted to highlight this paper is that it is a very nice example of trying to carefully compare the goodness of fit of different possible decision heuristics. Understanding the heuristics people are using can be a difficult task and is not as straightforward as estimating a parameter from an assumed model -- in these cases you are trying to compare *across* models.

The really interesting thing in the paper, at least in Justin's opinion, is the way they use some basic machine learning techniques to help explore the models that best predict choice. Justin's opinion is that this use of machine learning is likely to become very popular in behavioral/experimental research, especially as people try to better understand descriptive models of how people make decisions.

02/02/2018

Last week we shared an interesting paper that used machine learning techniques to help explore which potential behavioral models (using heuristics in that case) best fit the data and argued that could be an increasingly useful approach even in experimental studies.

Machine learning is not at the core of experimental studies, because we are often interested in estimating a treatment effect and the comparison of simple means does that. But when it comes to using measurements from experiments (e.g., preference measurements) and model building, predictive techniques could be useful.

This week we want to highlight some resources for those who are thinking of moving in this direction. Justin is interested in this and thinking of ways to invest in doing it himself, but is not yet up to speed. Anyone interested in possibly creating a reading/practice group on this over the summer, maybe we should discuss?

Okay, the resources: Here is a link to a summer workshop series put on at the NBER in 2015. Here is an approachable paper in the Journal of Economic Perspectives on the topic of machine learning for economists. (Side note: one of the authors Jan Spiess will be giving a job talk in the econ department in a few weeks). Finally, I just saw this page put together with links on places to get started with ML methods.

02/20/2018

For this week's BRITE IDEA we want to highlight a working paper by Goswami and Urminsky with a nice approachable write-up for a public audience available he re . In this paper, the authors conduct simple lab experiments where subjects work for either a flat wage for completing a task or for a payment that is based on the length of time they work. The key focus on the paper is on the decisions by another set of subjects who are the "managers" in the experiment and have to decide whether to offer the flat wage or time-varying contract.

They find that most managers choose the flat wage, but that paying by the minute would actually be much more profitable. They show that the reason for this is that the managers systematically over-estimate how long it will take people to complete the tasks.

The results here are interesting, but we also wanted to highlight it as a BRITE idea for two methodological points. First, this is a nice example of eliciting *beliefs and expectations* of subjects in the experiment as a way of understanding the mechanism behind the main effect. This is quite common in psychology but is probably under-used by those coming from the experimental-econ background.

Second, the way that they sequenced the study is quite interesting. The conducted the first stage having people work on different contracts separately and gathered those results. Then for the managers, they conducted the study online and told people they could hire a randomly selected worker from one of the two conditions with full information about the conditions from that first-stage of the experiment.

This type of structure is a nice, clean way of running an experiment that does not require a complicated interactive programming (i.e., z-tree) or the need to have large session sizes with different people doing different things in the session.

03/02/2018

For this week's BRITE idea we want to highlight an interesting new working paper from Cohn, Gesche and Marechal entitled " Honesty in the Digital Age ". They did a clever experiment here building off a now well-trodden design for detecting cheating. They had participants flip coins in private and then report their number of successes, for which they were paid. There was no way for the experimenters to verify whether they reported accurately and the instructions and nature of the design made that clear to subjects. That basic approach has been used to detect cheating patterns in many studies now, because at a group level success rates deviating above 50% identify likely cheating.

What was interesting in this study was that they varied whether you reported your answers to a human or machine interface and did so in a clever way. In the human conditions the subjects, who did the task remotely (e.g., online participants), placed a skype call or did a skype chat with a member of the research team. In the machine interactions in one condition they just entered into an online form and in another placed a similar skype call, but the answer on the other end was an automated recording instead of a live person.

They find that cheating rates are significantly higher in the two machine interactions than the two human interactions, with subjects reporting about 7 percentage points greater share of successes in the machine interactions. Justin thinks this is an interesting paper on a number of dimensions, but wanted to highlight it for BRITE Lab in part to point out the novel way of using Skype technology to facilitate different treatments remotely.

03/08/2018

This week we want to highlight a blog post by Keith Ericson at Boston University who is a co-editor at the *Journal of Public Economics* (and one of Justin's co-authors). He wrote a post about how he frequently sees experimental economics papers that end up failing to actually randomize their treatments.

Keith's post highlights the subtle cascade of issues that can lead this to happen, especially in experiments like the public-goods game where different treatments are delivered at different session levels. If you are running an experiment like that and having issues like this, please reach out to Justin and the broader BRITE community for feedback.

03/16/2018

For this week's BRITE ideas we want to highlight an article that shows how null results in clinical trials became shockingly more common after they started requiring pre-registration for clinical trials. Take a look at Figure 1 -- what a change.

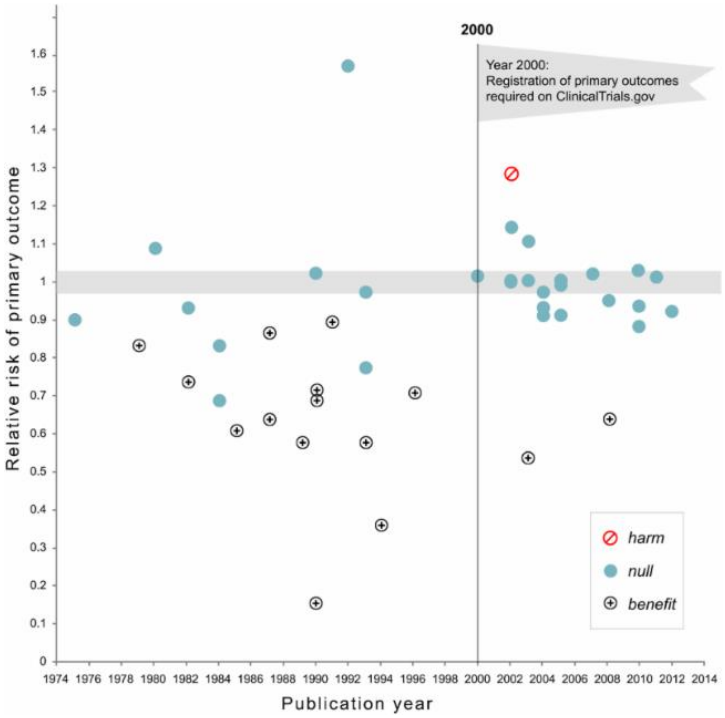


Fig 1. Relative risk of showing benefit or harm of treatment by year of publication for large NHLBI trials on pharmaceutical and dietary supplement interventions. Positive trials are indicated by the plus signs while trials showing harm are indicated by a diagonal line within a circle. Prior to 2000 when trials were not registered in clinical trials.gov, there was substantial variability in outcome. Following the imposition of the requirement that trials preregister in clinical trials.gov the relative risk on primary outcomes showed considerably less variability around 1.0.

We have posted before about the rise of preregistration for experimental studies. Justin has started doing this for his studies and is seeing a few clear benefits. First, it disciplines your study so that you plan it out better. Having to preregister is an up-front pain, but makes the study better. Second, there is a rapid increase in referees and journals looking for preregistration. If you have an interesting result, especially relying on interaction or heterogenous effects, you really need a pre-registration to be credible. Finally, good science should have us sometimes finding null results. If you pre-register, it is a little harder for refers to argue the study should not be published because it doesn't find anything -- you force them to consider the design and not just the results.

09/07/2018

A new academic year brings a new installment of the BRITE Ideas part of the news blast. This week Justin wants to highlight an interesting study by Chapman, Dean, Ortoleva, Snowberg, and Camerer that came out this summer.

The paper looks at willingness to pay (WTP) and willingness to accept (WTA) gaps. The common finding is that WTP is less than WTA, which is the classic "endowment effect". However, in this paper the authors establish a new pattern -- namely that WTP and WTA for the same individual evaluating the same financial lottery not only have this gap but also show very low correlation with each other. So in some sense it appears the two measures are simply capturing distinct and different things about attitudes toward risk.

The authors argue that theories of reference dependence could help rationalize this low correlation, but that it takes some pretty unique parameterizations of reference-dependent models to do so. Ultimately they argue that there is a need for "more theories and empirical studies of the process of buying and selling" to better understand why these measures show such low correlation.

09/14/2018

This week's BRITE Ideas comes from a paper that Jordan Tong (a frequent BRITE user and one of the faculty on the advisory board for BRITE Lab) sent along. The paper is by Klaus Fiedler from the journal Perspectives on Psychological Science (link here).

Fiedler makes an argument that more psychological (and I'd read that also more generally as behavioral) research should be based on "a priori theorizing". His broad point is that we should not be running experiments just to test for interesting phenomena. If we are very phenomena focused, it leads (in part) to the problem that many seemingly interesting findings might simply not be true. Instead, he argues that we should be developing generalizable theories about the world, individual behavior, etc... and designing experiments that shed light on those theories.

I've seen the power of this in my own work. A number of my field experiments were conducted with only vague ideas in mind and sometimes we got something interesting. But the best things I've done have been the cases where we actually wrote down a simple model and figured out what our theory had to say in conjunction with designing our experiment. Jordan's work uniformly reflects this perspective (which is part of why he sent it along). There are, of course, counter-arguments and other ways of doing things. But I think this is a worthy read.

09/21/2018

For this week's BRITE ideas we want to highlight the issue of data quality on the Mechanical Turk (MTurk) platform.

MTurk is a common platform for behavioral research and has a lot of advantages, including being fast, fairly low-cost and providing a more diverse respondent panel than traditional university labs. But the big problem is that the data quality can potentially be low. There are a lot of mixed results on the general quality of MTurk respondents, including some signs that you often can't get about the same quality on MTurk.

My own (limited) MTurk experience has been quite positive and in particular I like using MTurk for pilot testing of Qualtrics experiments. I have found that my pilot tests on MTurk are reliable guides to what I see in the BRITE lab in most cases.

However, this summer there was a "mini-crises" among academics who use MTurk heavily because many were starting to see a rise in low-quality responses. There was initially a big fear that people were creating automated bots to take MTurk studies. However, it now appears that the issue is actually that there are a number of foreign "workers" who are using survey farms and proxy servers to take multiple studies and often to pass themselves off as being in the U.S. so that they can qualify for studies that set higher screens.

There was some really interesting research on this that I'd like to highlight here . There was also a good twitter thread this week on a technique for getting these respondents using server farms out of any study you might run.

09/28/2018

This week we want to highlight a new working paper by Chapman, Snowberg, Wang and Camerer, called "Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)" .

The main reason we want to highlight the paper is methodological. The authors are introducing a new way to measure risk preferences and loss aversion instead of the standard multiple-price list or lottery-choice tasks. Their approach, which builds on some other similar papers, is to give people a series of simple choices. Which choice a person sees, though, depends on the choices they've made previously. The choice sequences are designed using Bayesian approaches to maximize the information you get about a person.

They show in the paper that this approach is accurate, stable and fast at the individual level. In terms of results, the paper is also interesting. While we focus a lot on loss aversion, this paper's key novel finding is that: Although many are loss averse, there are also many people who appear "loss tolerant".

10/05/2018

This week we want to highlight a nice overview article on the topic of mental accounting by Yiwei Zhang and Abigail Sussman from University of Chicago.

Mental Accounting addresses the fact that how we frame and label different sources of money can have important consequences for behavior.

Justin sees this as an area of interest in behavioral economics right now and guesses that more careful examinations of systematic aspects of mental accounting in applied settings could be productive.

10/19/2018

This week we want to highlight a new paper in *Science* by Armin Falk and Johannes Hermle looking at the nature of gender differences in economic preferences (e.g., risk aversion, patience, altruism, etc...) across countries with different levels of economic development and gender inequality.

They find a very interesting (and to Justin initially counter-intuitive) pattern that gender differences are ***larger*** in more developed countries with more gender equality. Economic development and gender inequality tend to make the genders perform more similarly. Their interpretation of the result is that in places with more resources there is less focus on universal goals of basic subsistence and more scope for gender-specific preferences to reveal themselves. They also argue that gender equality may make it easier for the genders to express different preferences.

Justin is not sure he's convinced on this, but wanted to highlight it because a) it is generally interesting and b) it highlights the important point that the preferences we measure in the lab can be culturally specific, which may be an important point for some applied research.

11/02/2018

This week we want to highlight the paper Global Evidence on Economic Preferences by Falk et al., which just came out in the *Quarterly Journal of Economics.*

The paper reports on a very large cross-national study of economic preferences (e.g., risk, time, altruism). The data for this paper were also used in the Falk and Hermle study we highlighted in a recent BRITE Ideas post.

The key reason we want to highlight it here is that the data are available from the study and may provide some useful background information to BRITE studies.

11/09/2018

For this week in BRITE ideas we want to highlight a new chapter that will be coming out in the Handbook of Behavioral Economics by Dan Benjamin at USC that surveys the theory and evidence in behavioral economics on the topic of errors in probabilistic reasoning and judgement biases.

The handbook will not come out for a year or so, but the chapter is available now as this NBER working paper . The chapter gives a great overview of topics like gambler's fallacy, hot-hand biases, and a range of issues in biased inference. One of the interesting points he makes at the end in his discussion section is that as people try to incorporate these ideas more into models of economic and business decision-making it often becomes important to understand a couple of things about beliefs that we rarely collect data on. Specifically, he argues that we need to know more about the extent to which people group series of sequences together for information processing versus updating sequentially. He also notes that many models also rely on how people think other people will update their beliefs after seeing signals, but we have a lot more evidence about how people update than about how they think other people update. If you are working on these topics at all, I encourage you to read up on this (long) chapter.

11/16/2018

For this week's BRITE Ideas we want to highlight a new paper by Pustejovsky and Tipton on a method for doing clustering standard errors in situations where you have relatively small sample sizes and fixed effects models.

They argue that this is important in practice for almost all RCTs and experiments with clustering. So in the BRITE Lab context you would think about studies that have session-level effects because people are interacting with each other in some way. One of the authors had a nice twitter thread about the method and noted that they have written both an R package and a STATA command for implementing the method.

There is a lot of work on clustering appropriately, so Justin wants to be clear that he has not yet figured out for himself when exactly to use this technique versus other small-sample correction approaches. However, this seems like a useful potential resource for BRITE Lab users. If anyone ends up implementing this at some point or looks into it deeply and decides to explicitly use another technique, please email Justin about your experiences so we can share it as a group.

12/07/2018

For this week's BRITE ideas we want to highlight a new paper " Many Labs2: Investigating Variation in Replicability Across Sample and Setting ."

This is a second paper in the "many labs" project trying to replicate seminal or influential studies in psychology and behavioral science. Like the prior many labs, they find that some things replicate, but many do not. In this case they are finding about half the studies replicate and typically at smaller effect sizes than the original publication when they do.

More interestingly, in this version they tried a number of things to get to the questions of whether things don't replicate is because the finding is somehow "fragile". That is, either the phenomena does show up strongly, but only under certain conditions, or it is generally not real "effects". Here is what one of the leading authors, Brian Nosek wrote about this: " The main purpose of ML2 was to examine heterogeneity across sample & setting. Some heterogeneity was observed. It was mostly in large effects, not in weak effects. The notion that some “fragile” effects are highly sensitive to sample had no support here." That is a very interesting finding.

Another finding from the study was that within large/strong effects, while there was some heterogeneity across populations (e.g., WEIRD vs non-WEIRD populations), the "replicable effects" are usually similar across different populations. Brian Nosek had a nice "tweek storm" about the findings
that you can see here .

01/14/2019

For this week's BRITE Ideas we want to highlight a recent working paper by Armando Meier at Chicago Booth titled "Emotions, Risk Attitudes, and Patience."

His paper looks at whether measured preferences toward risk and patience are influenced by emotional states. He uses data from the German Socio-Economic Panel that allows him to track individual differences in recently experienced emotions, such as anger, happiness and fear and correlate those within person changes to how people answer a general risk question and how they perceive their general level of patience.

He's able to get at causal effects by looking at plausibly exogenous shocks involving the death of a parent or child. The key take-away is that there are sizable changes in measured risk attitudes and patience driven by changes in emotions.

02/02/2019

For this week's BRITE ideas we want to highlight a methodological issue that often comes up in studies. The issue is what to do if you observe an outcome at two points in time for the same person with a treatment in between -- should you a) use the change in outcomes (i.e., the difference) as your measure or should you b) run your regression with the second outcome and control for the initial outcome level? That is, should you difference or control for prior outcomes?

I saw a nice blog post discussing this issue and illustrating the fact that this involves bias-variance trade-offs. For observational studies, there is no obvious right answer. However, for most experimental studies where you have randomized treatments, the right answer is actually to control for prior outcomes rather than use differences. The reason is that using the difference is better for dealing with bias, but with randomization the bias should be small. In these cases, it is better to control because that improves the variance. I think this is a reasonably well known issue, but thought it was nice to have this blog post summarizing the points.

02/08/2019

For this week's BRITE Ideas we want to highlight that the 2nd volume in the new Handbook of Behavioral Economics -- Foundations and Applications is now out. You can download a copy of all the chapters here . There is also a link there where you can download the 1st volume as well.

This is an amazing compilation of articles by many of the leaders in the field of behavioral economics. It's a great resource for anyone wanting to get familiar with or catch up on the forefront of behavioral economics. If people are interested, Justin proposes that we could create a reading group over the summer to go over the chapters and think about what new research ideas might be at the forefront of these topics. If that interests you, shoot Justin a short email and we'll see if we have a quorum.

02/27/2019

For BRITE Ideas this week we want to highlight a chapter put together by Susan Athey and Guido Imbens (aka the demigods of applied econometrics) about how to do statistical analysis for randomized experiments. This is somewhat old now (first came out in 2016), but Justin recently saw it again and thought it was a great resource for BRITE researchers.

The biggest point running through it is that they argue that people should use **randomization inference** techniques when calculating standard errors and confidence intervals for experimental data instead of classical "sampling based" techniques.

They also talk about a number of other issues, such as stratified random sampling vs. ex-post correction for covariate imbalance. There they argue that you should do stratified random sampling but also make a helpful note that you don't want to go too extreme on stratification. If you have a paper where you have used randomization inference, send it along to Justin so he can see what people are doing.

03/29/2019

For this edition of BRITE Ideas, we want to highlight a **write-up** of a recent workshop that was held at University of California, Berkeley. The workshop was on the future of forecasting and it highlighted exciting work on forecasting the results of social science research.

The key theme was that there is increasing interest in embedding tasks that get experts (and sometimes professionals or lay audiences) to predict the outcome of research projects before they are known. As the article highlights, this can be useful for a few reasons: It can help the research community overcome hindsight bias because it clarifies what was really known and expected prior to an experimental finding; It also may help researchers decide which treatments to explore when you have a number of possibilities but limited funding.

I think this is an interesting idea and I'd encourage everyone to take a look at this write-up. It has some **useful links** to slides and a number of very interesting studies by highly influential scholars in behavioral/experimental economics and psychology.

04/05/2019

For BRITE Ideas this week we want to highlight an exciting new resource for conducting studies. There is now an open-source program developed as an alternative to Qualtrics called formr . The new platform is based on the R programming language and is designed to let you do some pretty interesting things with data collection, such as easily displaying within a survey experiment information about the responses of others and displaying complex contingent graphics.

Justin has not fully read through this or tried it yet, but seems potentially quite powerful and interesting. This paper here describes the new platform with examples from three studies that have used it. The study examples are pretty interesting. Please let the group know if you use this at some point, as we would love a group tutorial or any thoughts on how it is working compared to Qualtrics for running web-based experiments.

* There is not a package on the CRAN yet. But if you are interested in the programming details of "formr", you could visit their GitHub space .

09/07/2019

For the first BRITE Ideas post of the fall, I (Justin) wanted to highlight one of my own studies that was run both in the BRITE Lab and online with Mechanical Turk subjects.

Many studies attempt to estimate risk attitudes by asking people to make choices over lotteries. These choices can then be fit to structural models to estimate parameters like classic (utility curvature) risk aversion, probability weighting and loss aversion. My coauthors and I tested whether these models can predict decisions people make over insurance choices. In theory, a good model of risk preferences should be informative to how people make insurance choices. So we conducted an experiment where people chose over lotteries designed to allow us to calibrate a big number of common utility models (e.g., expected utility, prospect theory) and then also had to decide how much of a potential loss they wanted to insurance at different prices and probabilities of loss.

The punchline finding is that none of the structural models do a good job at predicting insurance choices. In fact, they typically perform worse that just random choice in predicting what people choose in the insurance task. There are a couple of key reasons for this, maybe most generally interesting that the models estimated from lottery choices predict that people should respond a lot more than they do to the price of insurance. Ultimately we think this study suggests a) that there are some deeper decision processes (e.g., heuristics) at play when people decide about insurance that are not captured by models of risk attitudes and b) the exercise of estimating models of risk preferences from experimental data may have extremely limited external validity. Of side interest for our group, we find more noise in M-Turk samples than BRITE Lab, but otherwise the patterns of choices and preference parameters look very similar.

09/21/2019

***Sharing from Jihae Shin**:*

I recently conducted an experiment at the Brite lab for my paper on boredom and creativity. Recent research found that boredom can increase creativity (Gasper & Middlewood, 2014; Mann & Cadman, 2014). In my paper, I examine 'especially when does boredom increase creativity?' I hypothesized that boredom in one task increases creativity in another task especially when one engages in multitasking.

I conducted an experiment to examine this hypothesis. There were three conditions: a multitasking condition, a single-tasking condition, and a control condition. In the multitasking condition, participants worked on a boring task and a creativity task at the same time, and in the single-tasking condition, participants worked on the boring task and the creativity task separately and sequentially. In the control condition, participants only worked on the creativity task.

As hypothesized, I found that participants in the multitasking condition showed highest creativity. I propose that multitasking is a valuable strategy one can use to capture the power of boredom. We are now working on analyzing the rest of the data and hope to have a working paper soon.

09/30/2019

***Sharing from Paola Mallucci**:*

Paola Mallucci, Diana Wu and Tony Cui (forthcoming) " Impact of Social Motives on Bilateral Negotiations: How Power Changes Perceptions of Fairness ". *Journal of Economic Behavior and Organizations* .

We ran a modified ultimatum game that varies the balance of power between players through a probability that creates a convex combination of the dictator and ultimatum game.

We found that the power change affect what is perceived as a fair outcome of bilateral bargaining and that subjects with more power demand more, as long as there is common knowledge about power distribution. This means that while people think they deserve more when they have more power, they do not ask for more unless the person they are bargaining with knows they have more power. In contrast, we find providing feedback on past bargaining results has no effect on bargaining outcome.

10/10/2019

***Sharing from Kevin Chung**:*

A third-party reviewer can be a person, company, or government agency and can be found in a plethora of sectors and industries. Recently, there has been mounting evidence that calls into question just how unbiased some third party review systems are.

Two colleagues and I ran a study at the BRITE Lab to examine factors affecting the decisions of third-party raters and their impact on producers/service providers. We manipulated two factors (1) whether the rater and the service provider know each other's identity, and (2) whether the rater and the service provider interact repeatedly.

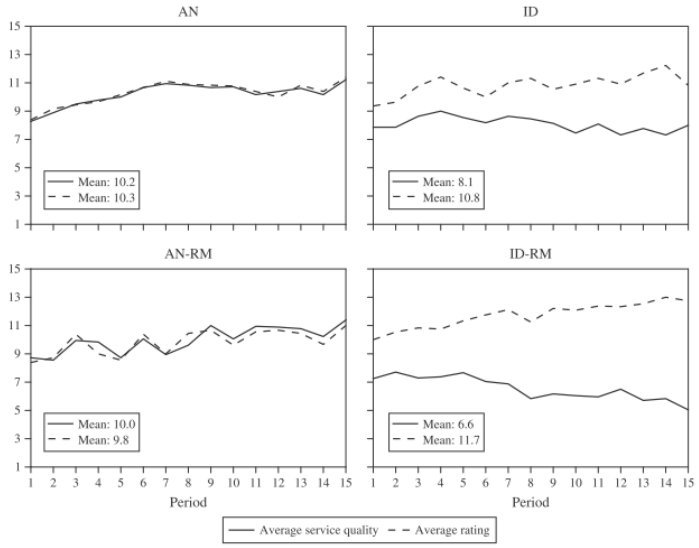
We find that decisions of both the rater and the service provider are very sensitive to the relational factors that govern their interaction. When the rater and the service provider know each other's identities, we observe a high proportion of overrating even though raters earn less monetary rewards for doing so, and the propensity to overrate is even stronger with repeated interactions (see below). Furthermore, the service provider chooses low quality levels.

This paper is now published at Management Science .

Table 1. Description of Treatments

No.	Treatment label	Social setting	Repeated matching
1	AN	Anonymous	one round
2	ID	Identity disclosure	one round
3	AN-RM	Anonymous	three rounds
4	ID-RM	Identity disclosure	three rounds

Figure 3. Exp 1: Average Service Quality Level ($\bar{e}$) and Average Rating ($\bar{s}$) Over 15 Rounds



10/18/2019

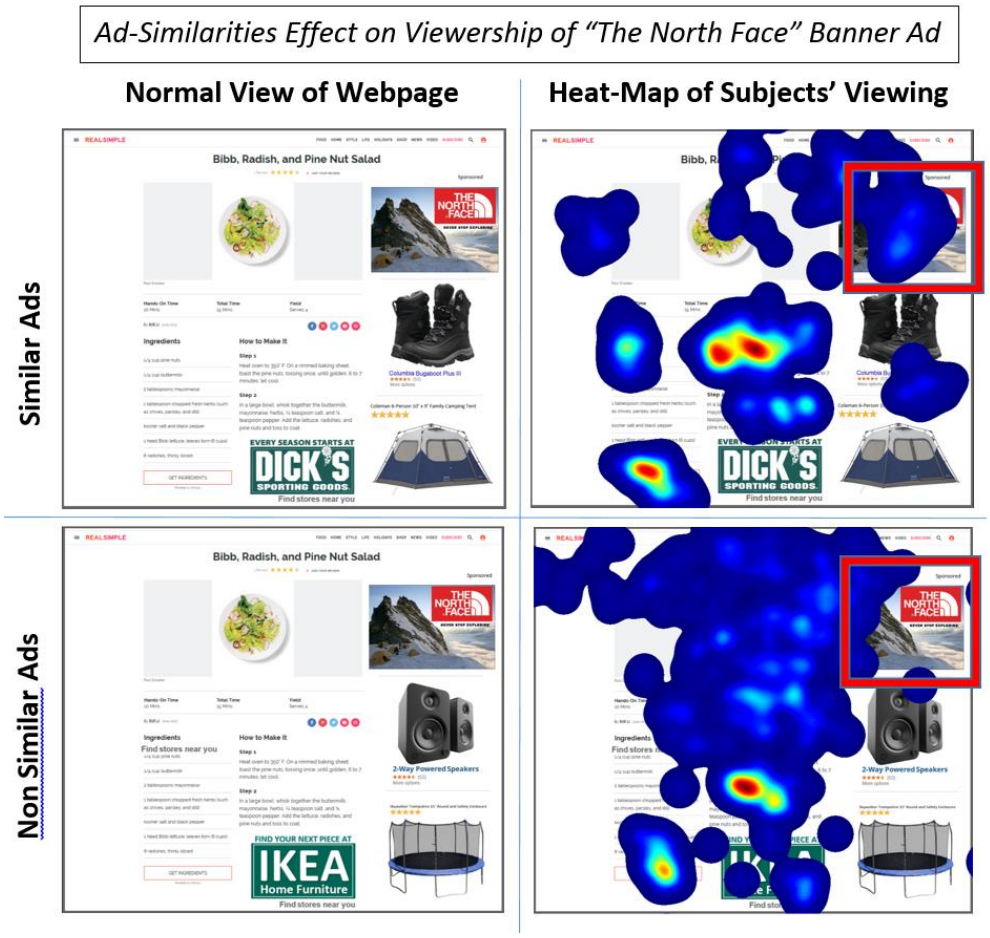
***Sharing from Paul Hoban**:*

Paul Hoban and Sean Melessa used a $4,500 BRITE Lab grant to create a persistent eye tracking lab at the Wisconsin School of Business. They have been using the lab to study how peoples' viewing habits of banner advertisements are impacted by the content of the ads to which they are exposed. They explore how peoples' viewing of a specific ad is impacted by the content of surrounding advertisements. In today's world of increasingly competitive targeting algorithms, it is becoming more common to see similar products being displayed concurrently on the same webpage. This is concerning for some advertisers because the traditional belief has been that having your ads placed next to ads by competitors would draw away the limited attention of consumers who are potentially interested in your product or product category.

However, Paul and Sean's research stream on banner advertising and targeting had led them to suspect that this congruence in content from targeting may actually be helping advertisers. They hypothesized that users are self-trained to avoid looking at banner ads, and that it takes substantial stimuli to draw their attention towards ads. Having multiple ads for similar products and brands may actually create a stronger attention stimuli. Thus, it can actually help to advertise near competitors because while you may get less relative attention, together there is more "attention pie" to split.

To test this hypothesis, they ran a series of randomized surveys on M-Turk and followed them up with an eye-tracking study using the newly acquired lab provided by the BRITE Lab Grant. Their experimental results show significant positive synergies between simultaneously displayed ads from competing products and brands. The image below provides an illustration of the experiments and analysis they are conducting and the mechanism they believe is at play. In this instance, 120 subjects were asked to read and learn about a website. Half of the subjects saw a website containing unrelated advertisements, while the other half saw the same websites but containing similar ads. While the students read the website, their viewing paths on the computer screen were measured and recorded with eye tracking, and afterwards were tested for differences in viewing behavior between the two groups. The image below is a heat map illustrating a snap shot of ten subjects' aggregate viewing behaviors at a fixed point in time. Note how the viewing pattern of the webpage varies between the two conditions.

They are still finalizing their analysis for this project, but ultimately they believe that their work is providing new results counter to many currently help beliefs and will help contribute to a better understanding of digital advertising.



Description of images above:

The webpage on top shows ads that are similar in content or product category (outdoor); while the webpage on bottom shows ads without a consistent theme. To the right of each webpage is a heat-map of ten students' viewing behavior at a fixed point in time for those websites. The red box shows the region that was used to test for differences in viewing behavior – The North Face advertisement. In this instance, The North Face ad received more viewing attention in the ad similarity condition, than it did in the non-similarity condition.

11/22/2019

***Sharing from Paola Mallucci**:*

We run an experiment to test the implication of moral licensing for consumer behavior. We gave subjects a small amount of money and offered them an option to bet on horses. After betting is done and winnings (or losses) are realized, subjects can elect to donate part of their compensation to a charity helping horses in retirement.

We find people are more likely to bet on horses if they are told about the option to donate before they bet. We also find that subjects that have been pre-alerted to the donation option are slightly less likely to donate, although they donate more when they do. This suggests that letting subjects know they can do a good deed in the future, gives them an excuse to engage in a morally dubious activity, even though they might never actually do the good deed. We are now collaborating with a racetrack to test these results in the field.

01/31/2020

***Sharing from Willie Choi**:*

We conduct an experiment to examine whether the effects of performance measure noise on employees' learning depends on whether employees engage in experiential learning (learning from one's own experiences) or vicarious learning (learning by watching others' experiences). Participants in our experiment engage in an abstract decision-making task where they try to maximize expected payoffs over many rounds. The experiential learners do the task many times and receive feedback, while the vicarious learners observe an experiential learner going through this process. A critical feature of the experiment is that there is a random element that adds performance-measurement noise, which is meant to reflect the potential for uncontrollable factors that affect performance. Greater noise makes it harder to use the performance measure to learn the best way to perform the task.

We find experiential learners exhibit less learning as performance measure noise increases, but an increase in performance measure noise does not affect learning when it occurs vicariously. The differences are strong enough that in our setting while experiential learners learn more when noise is low, vicarious learners actually learn more when noise is high. Why do vicarious learners do better in the noisy environment? We argue that they have a comparative advantage in processing information because they adopt a more holistic information processing approach than experiential learners. The experiential learners use a narrower set of criteria and that more myopic approach becomes problematic under performance-measure noise.

One promising application of our results relates to strategy development and implementation to control and improve financial performance in firms. In this setting, employees implement strategy through strategic initiatives and other similar actions, and learn about the effectiveness of these actions using the firm's strategic performance measurement system. Performance measure noise is especially problematic in this setting because it inhibits learning of causal relationships between employees' actions and outcomes of interest. However, the results suggest firms can overcome this by encouraging more vicarious, rather than experiential, learning. Thus, one benefit of the recent trend towards more open workplace environments is a greater propensity for employees to engage in vicarious learning.