Placebo and Belief Effects: Optimal Design for Randomized Trials

Scott Ogawa[1] & Ken Onishi[2]
Department of Economics
Northwestern University

Abstract

The mere possibility of receiving a placebo during a randomized trial has the potential to alter behavior because it alters subjects' beliefs. This is distinct from the traditional "placebo effect" and complicates the identification of relevant parameters. We assume that two factors determine human-subject experimental outcomes: *Treatment* and the subject's *belief about the probability of treatment*. Furthermore, we require that each subject must have correct beliefs about the probability of treatment. Given these two constraints, we investigate identification and optimal experimental design. Ultimately, we make three concrete optimal-design recommendations that maximize the statistical precision of various treatment effects. Most notably, we argue for a specific design in which the researcher introduces *variation in the probability* that each subject receives a placebo.

[1] sogawa@u.northwestern.edu
[2] onishi@u.northwestern.edu

I. Introduction

Since being famously considered and espoused by Henry Beecher (1955), the role of placebos in medical trials[3] has not only greatly impacted the interpretation of treatment effects but also the design of experiments themselves. Yet ironically the typical blind[4] experimental design introduces another problem: Since no subject believes with certainty that he or she is being treated during an experiment it is unclear that any observed treatment effect applies when subjects become fully aware that they are being treated.

Examples

As a leading example, consider an AIDS vaccine trial. It is possible that the treatment effect will be strong for subjects who are uncertain about treatment, but negligible for those who know they are being treated. Consider the following extreme example that illustrates the possibility. Suppose the vaccine works by making people immune from contracting AIDS from half of the infected population (maybe there are two strains). As it turns out, since the vaccine works against one of the two strains, it will be somewhat effective for people with one partner, but have almost no effect for those with many partners. If people who are unsure about treatment (the experimental population) choose to have one partner then there will be an experimental treatment effect. But if people who know they are being vaccinated choose to have many partners then there will be little treatment effect when the vaccine is implemented.

Notice that in the above example the behavioral response by the subject interacts with the treatment to determine the outcome. Furthermore, the response is not to the realized treatment, but to the belief about the probability of treatment. Besides AIDS vaccines, there are many other reasonable examples.

Cholesterol Medication: A subject who receives the medication might eat more bacon. Here it would not be unreasonable to think that medication will be "more effective" for people who eat lots of bacon. In other words, the difference in cholesterol levels between treated and non-treated subjects

[3] Within the medical liturature these experiments are typically referred to as Randomized Controlled (or Clinical) Trials. Henceforth we will refer to these as RCT's or simply "experiments." We believe the recommendations in this paper, however, apply to a wide range of human-subject experiments, including those in medicine and the social sciences.

[4] Note that this paper has nothing to do with the distinction between blind versus double-blind trials. Throughout we assume that the realized treatment is revealed only to the person analyzing the data.

will be largest among those who eat an unhealthy diet. “More effective”, however, is a debatable description in this case since the overall effect of the medication could be undesirable from the doctor’s perspective if it induces more bacon consumption.

Nicotine Supplements: A subject who wants to quit smoking might be more likely to do so if he believes the subsequent headaches will not be so bad. In this example, the behavioral response (i.e. did the person quit smoking) is what ultimately matters.

Adderall: If parents and students believe that the student is being treated with medication that will control behavior, the student may have a higher probability of being placed in particular environments. These environments alone may help the student focus.

In all of the examples above, there is an actual behavioral response to the belief of treatment. Thus the mere possibility of receiving a placebo has the potential to change the context in which the treatment is acting. It is this observation that is central to our paper.

Restrictions on Experimental Design

We assume that all RCT’s must adhere to one important restriction: Researchers must fully disclose to subjects the probability that they will receive the actual treatment. This can be thought of as either an ethical (or legal) constraint or simply a desirable property of an experiment since it requires no deception and thus increases the credibility of results.

Implications and Recommendations

This restriction, along with the aforementioned possibility behavioral responses, creates a significant problem for the researcher: It precludes the possibility of ever really knowing what the effect of a treatment in a non-experimental setting since it is impossible to observe the control outcome for individuals who think they are definitely being treated. In reality, we suppose most practitioners simply ignore this problem and assume the treatment effect measured using a classic RCT (in which some subjects are treated and some are not – and all are told there is a possibility of either) is the same as the treatment effect on individuals who definitely know they are receiving the treatment. In this context, we make the following three recommendations. The numbers presented are derived in

the appendix. Each maximizes the precision of a particular estimator that we detail in the next section.

1. *"The Free Lunch"*

   Any researcher who plans to run a classic RCT should instead break the subject pool into *two groups of equal size*: One will be told (truthfully) that there is a high probability (85%) of treatment while the other will be told (also truthfully) that there is a low probability (15%) of treatment. If properly designed, this modification comes at no cost to the researcher. If beliefs do not matter, then as before it will efficiently detect the treatment effect since half of subjects receive treatment and half do not. If, however, beliefs do matter, this setup will maximize the chances of detecting such an effect.

2. *Estimating the Realized Treatment Effect*

   We think it is quite likely that the researcher wants to know the treatment effect for subjects who know they are being treated. Yet if the researcher thinks that beliefs matter, then it is impossible to directly observe this effect (as discussed in the next section). We propose assuming that there is a linear relationship between beliefs and treatment effect. In this context, the optimal experimental design places 15% of subjects in one group, and 85% in another group. Those in the smaller group then receive the treatment with 15% probability, while those in the larger group receive the drug with 85% probability.

3. *Estimating the Placebo Effect*

   Finally, perhaps the researcher is specifically interested in estimating the effect of beliefs (the "placebo effect") irrespective of actual treatment. Optimal estimation is not straightforward (Kienle and Kiene, 1997) especially given the truth-telling constraint. If the researcher is willing to assume beliefs affect outcomes in a linear fashion then two thirds of subjects should receive the treatment with 75% probability, while the remaining third should not receive any treatment.

II. Framework

In order to analyze the problem faced by an experimentalist who is worried about the effect of beliefs, we use the following framework.

$t \in \{0,1\}$ – Whether or not the subject is treated.

$b \in [0,1]$ – The subject's belief about the probability of being treated.

$y(t,b)$ – Outcome as function of treatment and belief about probability of treatment.

In short, the outcome is a function of two variables. Now consider the following treatment effects, for which we have assigned suggestive names.

OTE: Overall Treatment Effect

This is what would happen if the researcher treated a person on the street and told him exactly what she was doing. It might be positive or negative depending on the effectiveness of the treatment and the behavioral response.

$$OTE = Ey(1,1) - Ey(0,0) \tag{1}$$

ETE: Experimental Treatment Effect

This is what a typical RCT study yields. Here $a \in (0,1)$ is the fraction of the subjects that are treated.

$$ETE = Ey(1,a) - Ey(0,a) \tag{2}$$

RTE: Realized Treatment Effect

This is the medical effectiveness of the drug for people who know they are taking the drug. We suppose that doctors are most interested in this effect since it isolates the medical effect of the vaccine given the beliefs (and thus behavior) of people when the treatment is implemented.

$$RTE = Ey(1,1) - Ey(0,1) \tag{3}$$

UTE: Unknowing Treatment Effect

This is the effect of the treatment on a population of people that do not know they are being treated. Realistically, this is unlikely to be of practical importance since it would be highly unethical to implement. On the flip side, however, this is probably the most interesting effect from a purely

scientific standpoint given it completely isolates the effects of the treatment (as opposed to the behavioral response).

$$UTE = Ey(1,0) - Ey(0,0) \tag{4}$$

PPE: Pure Placebo Effect

This is the effect on subjects of being told that they are being treated even though in reality no treatment is taking place.

$$PPE = Ey(0,1) - Ey(0,0) \tag{5}$$

Observations

To reiterate points we made in the previous section, it worth making the following observations. First, we can never observe $Ey(1,0)$ or $Ey(0,1)$. Thus without making further assumptions about $y(\cdot,\cdot)$ we can never know RTE, UTE, or PPE. Second, in general it is <u>not</u> true that UTE = ETE = RTE. In other words, the treatment effect found in a typical randomized controlled trial is not necessarily predict what will happen when the treatment is introduced to individuals who know that they are being treated.

Design Objective

In the preceding framework, the researcher has control over what each subject is told regarding the probability of treatment. Once this happens, the actual treatment for all subjects is determined by a randomization process. In essence, the researcher chooses a distribution function over beliefs. This will usually be a discrete mass function with two points of support. We will thus use lower case *p* to denote how subjects should be divided. In reality, this will be a deterministic process by the researcher and thus using *p* is sloppy notation since it does not represent a probability in the typical sense.

Functional Form Assumptions

The following table outlines several different functional form assumptions we can make about the nature of the $Ey(t,b)$. Informally, the upper-left corner of the table represents strong assumptions, while the lower right represents very weak assumptions.

Treatment Effect \ Belief Effect	None	Linear	Unrestricted
None	β_0	$\beta_0+\beta_1 b$	$g(b)$
Constant	$\beta_0+\alpha t$	$\beta_0+\beta_1 b+\alpha t$	$g(b)+\alpha_0 t$
Linear in b		$\beta_0+\beta_1 b+(\alpha_0+\alpha_1 b)t$	$g(b)+(\alpha_0+\alpha_1 b)t$
Unrestricted			$g_t(b)$

Table 1

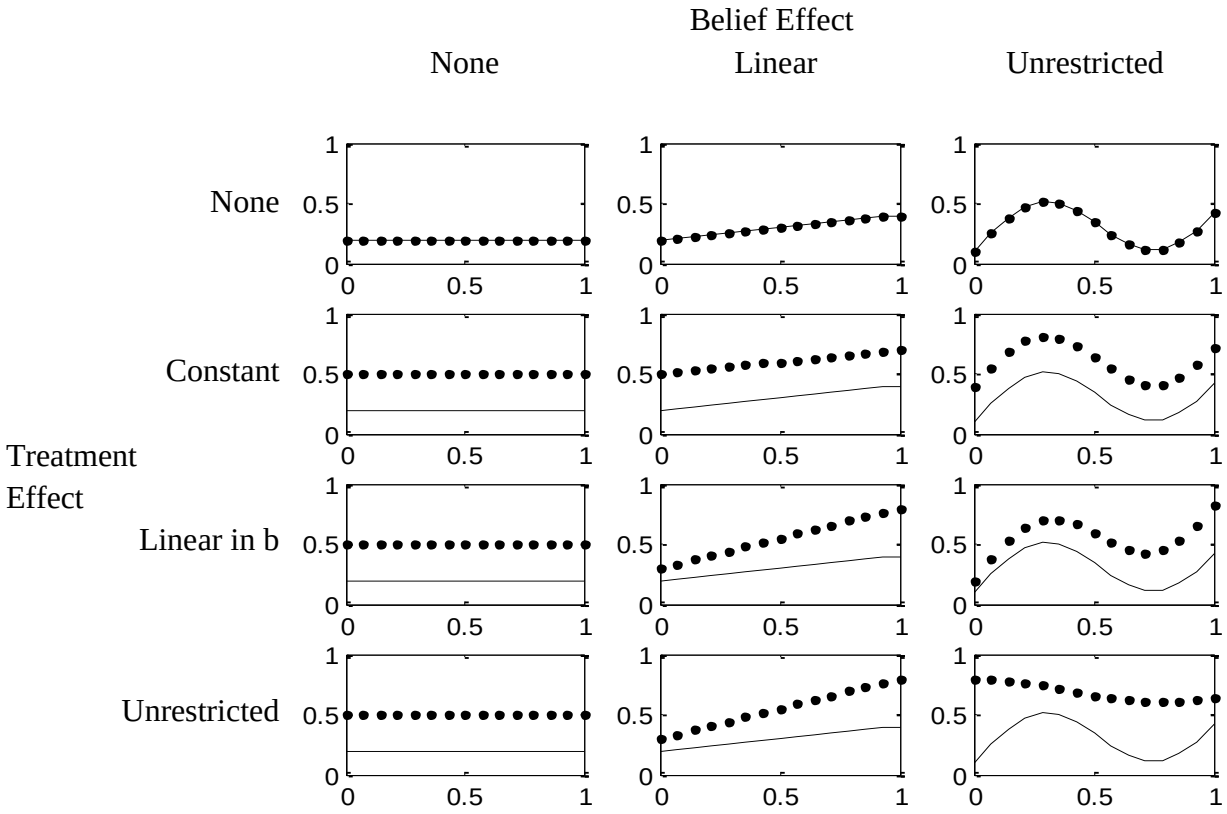

Figure 1: In the figure above, the horizontal axis has beliefs *b*, while the vertical axis is expected outcome *y*. The dotted line is an example of the outcome for the treated population while the solid line is the untreated population.

Identification and Estimation of Effects

We now outline identification and estimation of the previously discussed effects (OTE, ETE, RTE, UTE, PPE) under the various functional form assumptions. In most cases identification and optimal experimental design will be trivial. In some, however, this optimal design will not be so obvious. In all cases we assume there are a fixed number of subjects available to the researcher. The optimal design will be defined the one that minimizes the variance of the estimated parameters of interest.

OTE: Overall Treatment Effect

Under all specifications this effect is identified. Furthermore, the optimal design is to treat half the subjects and not treat the other half, and to reveal all information to all subjects.

$$\begin{aligned} p(b=0) &= 0.5 \\ p(b=1) &= 0.5 \end{aligned} \tag{6}$$

This is a simple yet important idea. It is unclear to us why OTE is not the most important treatment effect in many contexts. Consider again the cholesterol medication example: If people respond by eating more bacon then the overall effect will be weakened. If policy makers care only about cholesterol level in the population, then OTE is the only treatment effect that should matter.

ETE: Experimental Treatment Effect

As with the OTE, this is identified in all cases. If the choice of the parameter *a* is not important to the researcher then she should pick $a = 0.5$ and use the following design.

$$p(b=0.5) = 1 \tag{7}$$

This is how most RCT's are structured. The classic RCT design optimally estimates ETE, and then extends this estimation with the assumption that ETE = RTE. In doing so the researcher implicitly assumes that behavioral responses to beliefs do not change the treatment effect. The only way a

researcher can justify the above experimental design is if she is confident that beliefs do not alter the treatment effect yet do alter the outcome. This is the model $Ey(t,b) = \beta_0 + \beta_1 b + \alpha t$

We find it plausible, however, that if the researcher assumes the above model then she is also assuming that beliefs do not matter at all: $Ey(t,b) = \beta_0 + \alpha t$. Under this assumption, we see no reason to not use the following design:

$$\begin{aligned} p(b = 0.15) &= 0.5 \\ p(b = 0.85) &= 0.5 \end{aligned} \tag{8}$$

In other words, if there are 1000 subjects, 500 will be told that there is a 15% chance of treatment, and thus 75 people in this group will be treated (in expectation). The other 500 are told there is a 85% chance of treatment, thus 425 will be treated in expectation. Thus 500 of the 1000 subjects are treated in expectation. Therefore this design delivers the same level of statistical precision as equation (6) if we hypothesize that beliefs do not matter (the researcher simply ignores beliefs and conducts analysis as she would have previously).

At the same time, however, this design optimally estimates β_1 (the effect of beliefs) in the $\beta_0 + \beta_1 b + \alpha t$ model. Amazingly, it also is best for estimating α_1 (the effect of beliefs on the treatment effect) in the $g(b) + (\alpha_0 + \alpha_1 b)t$ model. In this sense, we believe it is a bit of a "free lunch" and should be used in most RCT's.

RTE: Realized Treatment Effect

Among all models considered in Table 1, only the most general ($Ey(t,b) = g_t(b)$) leaves RTE unidentified. In all other cases RTE is identified. Furthermore, if the researcher is willing to assume that the treatment effect is constant then she should simply design the experiment to optimally estimate the ETE. However, for the more general model of $Ey(t,b) = g(b) + (\alpha_0 + \alpha_1 b)t$ we suggest the following design in order to efficiently estimate RTE:

$$p(b=0.15)=0.15$$
$$p(b=0.85)=0.85 \tag{9}$$

Notice most subjects are led to believe with high probability that they will receive the treatment while a few are told there is a low probability of treatment. We believe this is well illustrated by the way people use two hand to play pool: They first "anchor" the cue stick with the hand closer to the cue ball – this is analogous to using most subjects to get a very good estimate for the effect on subjects with a belief similar to what the researcher is ultimately interested in estimating. The other hand is then used to guide the cue stick – generally it is good to place this hand as far away from the anchor point in order to minimize the effect of trembling hands.

UTE: Unknowing Treatment Effect

This is quite similar to RTE case and thus yields a similar optimal design:

$$p(b=0.15)=0.85$$
$$p(b=0.85)=0.15 \tag{10}$$

PPE: Pure Placebo Effect

The "pure placebo effect" may be resulting from a behavioral response by subjects (the focus of this paper) or simply some other mechanism that broadly falls under the category of "placebo effect." Either way it is impossible to identify this effect without making an assumption about the effects of beliefs since we cannot observe $Ey(0,1)$. The weakest possible assumption among those offered in table 1 is to assume that beliefs affect the outcome linearly (and the treatment can vary with beliefs): $Ey(t,b)=\beta_0+\beta_1 b+(\alpha_0+\alpha_1 b)t$. To estimate PPE we thus must estimate β_1. The best way to do this is to tell a third of the subjects that they will not receive the treatment, and then to tell the other two thirds that there is a 75% chance of treatment:

$$p(b=0)=0.33$$
$$p(b=0.75)=0.67 \tag{11}$$

Appendix: Optimization Procedures

This section outlines the optimization procedures use to derive the recommendations made in the previous two sections. The general strategy is to first assume only two mass points for b[5]. We also assume that the researcher will regress the outcome (or difference in outcomes between two groups, depending on the context) on b. They will use generalized least squares where the weight matrix is a known function of *b*.

Estimating the Realized Treatment Effect (recommendation 2)

As previously discussed, estimating RTE with relatively week assumptions is possible. The model we consider in makes no assumptions about the effect of beliefs on outcomes, but does assume the treatment effect depends linearly on beliefs: $Ey(t,b) = g(b) + (\alpha_0 + \alpha_1 b)t$. In this framework, the RTE can be written

$$RTE = \alpha_0 + \alpha_1 \tag{12}$$

Our goal will be to minimize the variance of the estimated values.

The covariate matrix and the weight matrix can be written down as follows if we assume only two mass points.

$$X = \begin{bmatrix} 1 & b_L \\ 1 & b_H \end{bmatrix} \tag{13}$$

$$\Omega_{i,j} = \begin{cases} \dfrac{y(0,b_i)\left[1-y(0,b_i)\right]}{n_i(1-b_i)} + \dfrac{y(1,b_i)\left[1-y(1,b_i)\right]}{n_i b_i} & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

[5] Computer simulation suggests that this is always optimal. Formal proof, however, is an area that we are currently working on.

Since the weight matrix above is 2x2, $i=1$ corresponds to the "Low" value (denoted by subscript L) and the $i=2$ corresponding to the "High" value (denoted by H). Assuming a fixed number of subjects *N*, we naturally imposed the following restriction:

$$n_L + n_H = N \tag{15}$$

The problem now becomes a constrained optimization over three variables: b_L, b_H, n_L. The goal is to minimize the following value.

$$\text{var}\left(\hat{\alpha}_0 + \hat{\alpha}_1\right) = \text{var}\left(\hat{\alpha}_0\right) + \text{var}\left(\hat{\alpha}_1\right) + 2\,\text{cov}\left(\hat{\alpha}_0, \hat{\alpha}_1\right) \tag{16}$$

Where the variance-covariance matrix is given by:

$$\begin{bmatrix} \text{var}\left(\hat{\alpha}_0\right) & \text{cov}\left(\hat{\alpha}_0, \hat{\alpha}_1\right) \\ \text{cov}\left(\hat{\alpha}_0, \hat{\alpha}_1\right) & \text{var}\left(\hat{\alpha}_1\right) \end{bmatrix} = \left(X'\Omega^{-1}X\right)^{-1} \tag{17}$$

We have worked out this problem numerically (ie minimized its value in Matlab) assuming the actual probabilities $y(t,b)=0.5$ regardless of the treatment or belief.

Here is the expression that is minimized over three variables (b_L, b_H, n_L) for the UTE problem, which is to minimize $\text{var}\left(\hat{\alpha}_0\right)$.

$$\min_{b_L, b_H, n} \frac{n b_L^3 \left(1-b_L\right) + \left(1-n\right) b_H^3 \left(1-b_H\right)}{n\left(1-n\right) b_L \left(1-b_L\right) b_H \left(1-b_H\right) \left(b_H - b_L\right)^2} \tag{18}$$

As it turns out, these probabilities do not significantly affect the optimal experimental design. To the extent that they do, the optimal design can be adjusted accordingly based on educated guesses for the function *y*.

Free Lunch for most RCT's (recommendation 1)

We impose two additional restrictions to the optimization in the previous section.

$$\begin{aligned} n_L &= n_H \\ b_L &= 1 - b_H \end{aligned} \tag{19}$$

These two restrictions ensure that half of all subjects are treated and half are not. Optimizing under this constraint yields equation (9).

Estimating the Placebo Effect (recommendation 3)

In this case we assume the model $Ey(t,b) = \beta_0 + \beta_1 b + (\alpha_0 + \alpha_1 b)t$, though ultimately are only interested in estimating β_1 with greatest efficiency. This is the effect of beliefs on subjects who are not being treated.

We again assume two mass points for subjects and thus have the following weight matrix.

$$\Omega = \begin{bmatrix} \dfrac{y(0,b_L)\left[1 - y(0,b_L)\right]}{n_L(1-b_L)} & 0 \\ 0 & \dfrac{y(0,b_H)\left[1 - y(0,b_H)\right]}{n_H(1-b_H)} \end{bmatrix} \tag{20}$$

Otherwise the optimization is similar to before.

References

Beecher, Henry K. “The Powerful Placebo” *Journal of the American Medical Association*. 1955; 159(17):1602-1606.

Kienle, Gunvar, and Helmut Kiene. “The Powerful Placebo: Fact or Fiction?” *Journal of Clinical Epidemiology,* Vol 50, No 12, pp 1322 – 1318, 1997.

Kirsch I, Rosadini MJ. “Do double-blind studies with informed consent yield externally valid results? *Psychopharmacology* 1993: 110: 437 – 442.