Despite increasing efforts to encourage the adoption of field experiments in marketing research (e.g., Campbell 1969; Cialdini 1980; Li et al. 2015), the majority of scholars continue to rely primarily on laboratory studies (Cialdini 2009). For example, of the 50 articles published in Journal of Marketing Research in 2013, only three (6%) were based on field experiments. The goal of this article is to motivate a methodological shift in marketing research and increase the proportion of empirical findings obtained using field experiments. The author begins by making a case for field experiments and offers a description of their defining features. She then demonstrates the unique value that field experiments can offer and concludes with a discussion of key considerations that researchers should be mindful of when designing, planning, and running field experiments.

**Keywords:** field experiments, lab experiments

### Field Experimentation in Marketing Research

**THE CASE FOR FIELD EXPERIMENTS**

Researchers have proposed several reasons why field experiments constitute only a small fraction of experiments conducted by social scientists (e.g., Cialdini 2009). First, field investigations tend to be highly time consuming: identifying potential partners and obtaining their consent, in addition to planning and collecting data, can take several years. In less successful instances, one’s efforts might prove futile for reasons beyond his or her control (e.g., a company’s legal or public relations considerations). These uncertainty and time intensiveness characteristics decrease the attractiveness of field experiments for any researcher, especially during the early stages of one’s academic career. Second, field experiments focus on observed behavior, thus limiting researchers’ ability to investigate the psychological processes underlying that behavior and, with it, the project’s theoretical contribution (e.g., Fong, Fang, and Luo 2015). Finally, because of the dynamics inherent in field (vs. lab) settings, field experiments often involve a high degree of noise and limited control over the experimental procedure, which in turn might introduce challenges for data analyses and interpretation. At this point, one might be asking why, given the challenges and potential downsides, researchers should even consider running field experiments.

To me, the answer is clear. It begins with the simple yet profound notion that behavioral research aims to understand how people behave, what influences their preferences and choices, how initial behaviors shape future behaviors, and so on. Although most published articles discuss how the results they report apply to real-world situations, they do so by extrapolating from research conducted in the lab, which often requires a considerable leap of faith. By contrast, results obtained using field experiments, which study behavior in the most relevant environment from an outcome perspective, can be taken at face value—they report the size of the effect, identify who is most affected and under what conditions, and the short- and long-term implications for, for example, sales, loyalty, profitability, and voting participation. I therefore propose that marketing research would benefit greatly from articles that offer converging evidence, using theoretical models, the laboratory, field experiments, and naturally generated empirical data.

**What Constitutes a Field Experiment?**

In field experiments, participants are unaware that they are taking part in a study—or, if they are aware, they are engaging in activities as they normally would, regardless of the experiment (Charness, Gneezy, and Kuhn 2013). Typically, field experiment participants are unaware that the researcher is manipulating factors in the environment and measuring behaviors and outcomes. Examples include changing prices for products (e.g., Anderson and Simester 2003), varying expiration dates of promotional offers (e.g., Shu and Gneezy 2010), and testing responses to different ads (e.g., Bertrand et al. 2010).

Field experiments can be instrumental in several ways. For example, an experiment conducted in the field can be used to (1) apply a previously established theory/framework and observe the extent to which it applies to real life (VanEpps, Downs, and Loewenstein 2016); (2) identify the effect size;
and (3) capture second-order effects, long-term effects, and so on (e.g., John et al. 2017; Putnam-Farr and Riis 2016). Alternatively, in some instances, it makes sense to start with a field experiment and, if necessary, follow up with laboratory experiments to understand the process underpinning the effect. Examples include research on voting behavior (e.g., Bryan et al. 2011), response to location-based promotional offers (e.g., Fong, Fang, and Luo 2015), and pay-what-you-want pricing (e.g., Gneezy et al. 2010, 2012). Finally, researchers can use field experiments to pit several theories against one another and observe the relative impact of each, thereby providing a more complete account of behavior within a given domain (e.g., Sudhir, Roy, and Cherian 2016).

**A Word on Field Experiments and External Validity**

Scholars have argued that research conducted in naturally occurring settings offers greater external validity (e.g., Calder, Phillips, and Tybout 1982; Winer 1999). Lynch (1999) rejects this proposition and claims that, similar to the lab, one real-world setting is unlike another, and thus any claim about the generalizability of any type of experiment is spurious. Instead, Lynch proposes that the key to understanding whether findings can be generalized is to “have a theory that specifies moderator variables and boundary conditions and that specifies what variables should not moderate the findings reported and to test for the asserted pattern of interactions” (p. 368).

I concur with Lynch’s assertion that external validity depends on theory, not method. At the same time, it is important to remember that experiments conducted in natural settings can often uniquely explain where and when certain phenomena are likely to occur; help refine existing theories; and, importantly, highlight their significance and likely impact (Campbell 1969; Cialdini 1980, 2009; Lewin 1997; McGuire 1969).

**FIELD EXPERIMENTS EXAMPLES**

One of the most important decisions firms face is how to price their products or services. Although the classic assumption in economic modeling is that firms indeed set profit-maximizing prices, empirical evidence has shown that this assumption is often violated (e.g., Simon 1955). Ostensibly, a firm can identify the best price when it estimates the demand function—observing sales at different price levels—and subsequently chooses the profit-maximizing price. Alternatively, demand elasticity can be estimated by employing econometric techniques using historical price and sales data (Nevo 2001; Rosse 1970), which can then be used to derive a demand function. Although elegant and simple, this approach hinges on the critical assumption that the firm has been setting profit-maximizing prices, an assumption that—even if true—restricts price variation, thereby handicapping the econometric approach. Finally, historical demand is often nondiagnostic of future demand. For example, the demand for VCR players was likely stable at some point but is surely quite different now.

A cleaner, more straightforward, and more accurate approach for deriving optimal prices is to run a field experiment in which researchers vary prices and observe demand. Notably, this approach does not rely on assumptions and is flexible enough to detect and correct for fluctuations in demand over time.

**Example 1: Price, Quality, and Demand**

Gneezy, Gneezy, and Lauga (2014) study the price–quality heuristic in an experiment in which they varied the wine’s quality (high vs. low) and price ($10, $20, $40). Participants were winery visitors who came for a wine tasting: they received a printed page with the names and prices of nine wines available for tasting and chose six wines to taste. This design enabled the authors to examine the effect of price and quality on liking, demand, revenue, and profits. The main finding is that for the higher-quality wine, demand went up when the price increased from $10 to $20 and went slightly down when the price was $40. By contrast, demand for the lower-quality wine consistently decreased with price. The most profitable combination was $20 for the higher-quality wine. Using these findings, the winery priced its better Cabernet Sauvignon at $20, resulting in an 11% profit increase. Note that if the same experiment were conducted in the lab, any conclusion regarding the implications for demand, sales, and profitability would require some leap of faith and take the form of informed speculation at best.

**Example 2: Price and Perceived Fairness**

On July 12, 2011, Netflix announced its decision to replace its membership plan, which offered both unlimited streaming and DVDs, with two distinct plans: one with unlimited streaming and one with unlimited DVDs, priced at $7.99 each. This 60% price increase caused the company’s market value to drop by over two-thirds—$11 billion—in three months. The outrage of Netflix’s customers was rooted in a deep sense of unfairness and disappointment, which could not be captured by standard pricing strategies. A field experiment, however, could have uncovered the effects of different prices on fairness perceptions, identified the pros and cons associated with each, and enabled Netflix’s leadership to make a well-informed decision.

Anderson and Simester (2008) use a field experiment to investigate the impact of prices on perceived fairness and sales. Collaborating with a women’s plus-size clothing catalog, the authors mailed four versions of the catalog to potential customers and varied the magnitude of the price premium for the larger sizes. The results show that demand decreased as the premium increased, arguably because of perceptions of unfairness. Using actual sales data, the authors estimated the cost of this sense of unfairness to the firm to be a 6%–8% decrease in gross profits.

**KEY CONSIDERATIONS**

This section offers a brief discussion of important considerations and decisions associated with the design and implementation of field experiments. Note that this is not a complete account of the steps and considerations, but rather a discussion of factors that are unique to field experiments and yet are often overlooked or unacknowledged by researchers.

**Context/Setting and Experiment Type**

In some cases, the nature of one’s research question necessarily dictates the experimental context. For example, research examining the effect of Facebook “liking” on subsequent attitudes and behavior needs to be done on Facebook (John et al. 2016). Sometimes, however, researchers have more flexibility in choosing the setting for their experiment, enabling them to conduct independent field experiments.

In general, field experiments can be conducted with (collaborative) or without (independent) a nonacademic partner (e.g., organization/firm/seller). The decision is entirely the
researchers’, as long as it allows them to test predictions, proposed theories, and so on. Note, however, that independent experiments are often limited to a one-time interaction, meaning that they do not lend themselves easily to the measurement of second-order, spillover, and long-term effects.

Independent experiments—designing an ad hoc field experiment. Sometimes researchers can design and conduct an experiment that has all the necessary features of a field experiment. For example, to investigate the factors influencing behavior under what-you-want pricing, researchers offered donuts (Saccardo et al. 2016) and coffee (Jung et al. 2014) to passersby. Another example is the lost-letter technique (Müllermann, and Harter 1965), used to assess community orientation toward political groups, human cooperation (Yezer, Goldfarb, and Poppen 1996), and other types of social preferences. Needless to say, this freedom comes at a cost, so it is critical to successfully identify those real-world behaviors and situations that best fit the research question.

Collaborative experiments—working with a nonacademic partner. A research partner could take the form of any marketplace actor interacting with consumers: a coffee cart manager (e.g., Jung, Perfecto, and Nelson 2016), a university’s human resource department (Grinstein and Kronrod 2016), hotel management (e.g., Baca-Motes et al. 2013; Goldstein, Cialdini, and Griskevicius 2008), government/public agencies (e.g., Yeager et al. 2014), and so on. The general rule of thumb is that the smaller the organization, the more quickly an understanding can be reached.

I recommend that researchers familiarize themselves with, and develop a good understanding of, potential nonacademic partners before approaching them. Researchers should know how such potential partners operate; what their challenges and needs might be; and, importantly, the potential costs and benefits of the proposed experiment for them. Although it will of course be important to highlight potential gains the study can bring, I also recommend proactively presenting potential caveats and discussing ways to address them.

Potential Benefits and Costs

From the perspective of potential partners, the potential benefits associated with running a field experiment with an organization that knows what it is doing might include increased sales and profits and improved effectiveness of promotions (Fong, Fang, and Luo 2015) and communications (Putnam-Farr and Riis 2016), for example. Most importantly, a well-planned and well-executed experiment can reveal what works (vs. not, or not as well), under what conditions, and the associated ramifications (Bone et al. 2017). Note that, in my experience, pitching the importance of enhancing scientific knowledge is usually ineffective in convincing the potential partner and might even backfire, driving the company/organization away.

Understanding possible concerns and addressing them before approaching a research partner is important. Common concerns include costs (e.g., time, money), reputation and legalities (see the criticism of Facebook’s emotional-manipulation experiment by Kramer, Guillory, and Hancock 2014), and interference with operations.

Design Tightness and Experimental Noise

As mentioned previously, field experiments often involve a significant loss of control and an increased number of variables that must be considered in the design phase, as well as during data analyses. Consider the case of time effects: people are likely to feel that they are “richer” earlier in the month, right after receiving their paycheck, which suggests differential preferences with respect to, for example, spending (Fong, Fang, and Luo 2015) and saving (Grinstein and Kronrod 2016). Similarly, predicting the specific time of day (e.g., before or after lunch) could substantially—and possibly systematically—influence participants’ preferences and behavior and should therefore be considered as well.

More generally, researchers should capture any variation in the experimental setting, small or large (preferably in real time), determine its potential to undermine the experiment, and take the steps necessary to address such concerns. Consider a study in which researchers offer chocolate cupcakes using pay-what-you-want pricing to visitors of a local farmers’ market over four consecutive Sundays. Imagine that in week 3, outside temperatures hit 100 degrees Fahrenheit, as opposed to 72 degrees Fahrenheit observed in prior weeks, significantly decreasing the attractiveness of the goods and, thus, consumers’ willingness to pay. The researcher could take several steps to address this problem, such as pausing data collection and resuming the following week, or continuing with data collection and controlling for temperature in the analyses. The main takeaway is simple: researchers should develop and maintain a system that enables them to capture and address (unexpected) situational factors in real time. One could deal with some of these issues, such as time of day, by developing a detailed protocol that defines randomization, data recording, documentation of irregularities, and so on. The more prepared a researcher is, the better. However, even the most meticulous protocol is unlikely to account for all possible problems, which is why the presence of the researcher on the ground is imperative, particularly during the first few sessions.

Finally, when planning a field experiment, researchers sometimes need to “give in”—adjust and modify aspects that are otherwise considered the golden standard. For example, (Putnam-Farr and Riis 2016) provide an account of compromises dictated by the circumstances surrounding their Experiment 2: lack of complete randomization within each client’s pool of employees and a slight variation in the rewards offered (and the text describing them) by each client. Although largely undesired, such hiccups are inherent to the nature of the real world and should be acknowledged and addressed to the extent possible.

A FINAL NOTE

The goal of this article is to demonstrate the importance of field experiments and make the case for their adoption in marketing research. The merits of field experiments are substantial, but they often come with some costs. My recommendation is clear and simple: researchers interested in understanding a behavioral phenomenon—its antecedents, consequences, and implications—should make field experiments an integral component of their everyday research methods toolkit. Finally, note that Journal of Marketing Research’s commitment to promoting the adoption of behavioral field experimentation is aligned with what seems to be a general shift in our field, as evidenced by editorials published in Journal of Consumer Research (Peracchio, Luce, and McGill 2014) and Marketing Science (Sudhir 2016).
REFERENCES