Field Experiments in Organizations

Dov Eden

Coller School of Management, Tel Aviv University, Tel Aviv 6997801, Israel;
email: doveden@post.tau.ac.il

Annu. Rev. Organ. Psychol. Organ. Behav. 2017. 4:91–122

First published online as a Review in Advance on December 21, 2016

The *Annual Review of Organizational Psychology and Organizational Behavior* is online at orgpsych.annualreviews.org

https://doi.org/10.1146/annurev-orgpsych-041015-062400

Keywords

field experiment, experimental control, randomization, quasification, quasi-fields, little data

Abstract

Field experimentation, although rare, is the sterling-gold standard of organizational research methods. It yields the best internally valid and generalizable findings compared to more fallible methods. Reviewers in many psychology specialties, including organizational psychology, synthesize largely nonexperimental research, warn of causal ambiguity, and call for experimental replication. These calls go mostly unheeded. Practical application is a raison d'être for much organizational research. With the emergence of evidence-based management, field experiments enable us to deliver the most actionable tools to practitioners. This review explicates the role of experimental control and randomization and enumerates some of the factors that mitigate field experimentation. It describes, instantiates, and evaluates true field experiments, quasi-experiments, quasi-fields, combo designs, and triangulation. It also provides practical tips for overcoming deterrents to field experimentation. The review ends describing the merging of new technologies with classical experimental design and prophesying the bright future of organizational field experimentation.

INTRODUCTION

Field experimentation should be the method of choice for studying the many independent variables in organizational research for which field-experimental treatments can be devised. Organizational scholars too often overlook field experimentation as a design option, even when investigating treatable (aka, manipulable) independent variables. It is crucial that organizational research establish causality. To be applicable it must be both internally valid and generalizable. Reviewers' ubiquitous calls for experimental replication (Bass & Bass 2008, Kluger & DeNisi 1996, Stajkovic et al. 2009), specifically for field experiments (Scandura & Williams 2000), evidence unquenched thirst for organizational field experimentation.

Social psychologists prefer laboratory experiments over field experiments. For example, Deci et al. (1999) expressly excluded field experiments from their review of experimental research on the effects of reward on motivation because they are "noncomparable with the laboratory experiments" (p. 635). This preference is inappropriate for organizational researchers, with our commitment to creating actionable knowledge. Fortunately, the many field experiments reviewed here—and many others not cited here—are evidence that editors and reviewers of top journals do appreciate well-done field experiments.

This is not a review on method. Numerous publications explicate experimental method thoroughly (Cook & Campbell 1979, Lipsey & Cordray 2000, Shadish et al. 2002), including field experimentation in organizations (Aguinis et al. 2009, Dipboye 1990, Grant & Wall 2009, Highhouse 2009, King et al. 2012, Stone-Romero 2011). Boruch & Wothke (1985) discuss randomization in field experiments in detail. Distinctions and confusion regarding external validity, ecological validity, generalization, and the role of mundane realism (Aronson & Carlsmith 1968, Highhouse 2009) are not discussed here. The review shows what has been done—and how much more could be done—applying the field-experimental method in organizations to advance organizational research.

Neighboring fields are experiencing rapid growth of field experimentation. These include marketing (e.g., Ein-Gar & Steinhart 2011, Petersen & Kumar 2015) and behavioral economics (e.g., Azar et al. 2013, Bandiera et al. 2011, De Paola et al. 2014, Pelligra & Stanca 2013, Shavit et al. 2014, Stoop 2014). There is also expanding methodological literature on field experimentation in economics (e.g., Harrison & List 2004, Levitt & List 2009). Organizational psychology and organizational behavior (OPOB) may soon have to play catch-up.

If the review is successful, the reader (*a*) when planning new research will consider field experimentation as an option; (*b*) when reading research reports, will contemplate how field experimentation might better explore the same relationships; and (*c*) when writing discussion sections about nonexperimental research, will suggest field-experimental follow-up.

The sidebar Major Topics Covered lists the key themes discussed below. The prime importance of causality in organizational research, and the unique appropriateness of field experimentation to provide such causal knowledge, is a focus. The review discusses briefly experimental treatment, control, and randomization, then provides a description of some of the earliest attempts to conduct experiments in organizations. Distinctions among true experiments, quasi-experiments, field experiments, and quasi-fields are elaborated. Laboratory experimentation on organizational variables (see Weick 1967) is not reviewed. The review then discusses the factors that militate against field experimentation and dispels some misconceptions. These include presuming that field experiments are too difficult to do, that samples are too hard to obtain, and that randomization is impossible in organizational settings. The variety of published and unpublished organizational field experiments, including field experiments in organizations, laboratory-like experiments in the field, large- and small-scale field experiments, and in-house, problem-focused

MAJOR TOPICS COVERED

Centrality of causality
Rarity of field experimentation
Experimental treatment and control
Randomization: friend and foe
Pioneers: the earliest organizational field experiments
Quasi-experiments, quasification, quasi-fields, and combo designs
Lab-like experiments in the field
Generalization
Big and small field experiments
"Little data": unsung field experimenters
Factors mitigating field experimentation
Triangulation and mixed methods
Overcoming deterrents to field experimentation:
- utilize management indifference
- when randomization and fairness align
- invert the treatment
- piggyback on naturally occurring events
- transform delicate data
- use emerging technologies

Approaching the ideal design
A model field experiment
The mother-in-law test (MILT)
The future of organizational psychology and organizational behavior field experimentation

practitioner-conducted field experiments, is instantiated. Experience-based "do's" and "don'ts" are listed. The review closes with optimistic prophecy regarding the future of field experimentation in organizations.

CENTRALITY OF CAUSALITY

It is hard to refrain from causal thinking. Kahneman's (2011) System 1 thinking, which he dubs fast thinking, has us automatically inferring causality: "We are evidently ready from birth to have impressions of causality which do not depend on reasoning about patterns of causation" (p. 76). We are wired to assume that relationships evidence causality. Kahneman describes this as a major bias with potential for erroneous intuition. Organization scholars are not immune to fast thinking and fallacious causal inference. The experiment's relative causal certainty justifies the causal inference that we so covet and elevates it above other methods, including the novel and ever more sophisticated analyses designed to infer causality from observational data. Highhouse (2009) stated it most cogently: "Randomized experiments are the most potent research design for determining whether or not *x* causes *y*" (p. 554).

Unfortunately, most organizational research lacks causal persuasiveness. Aguinis & Vandenberg (2014) described it as a central shortcoming:

> Better decisions presumably stem from findings for which strong confidence exists as to the cause and effect of a phenomenon. The reality, though, is that such confidence does not exist in the majority of

> cases because of how the study was originally designed. . . . [T]he most frequently used research designs are passive observation studies in which we observe whether the rank order of values in one (or more) variable(s) is associated with the rank order of another variable(s). Causal inferences are improbable under those circumstances. (pp. 584–85)

Testing causal hypotheses rigorously is crucial for scientific purposes and indispensable for putting the very best, evidence-based tools into the hands of practitioners (Giluk & Rynes 2012, Rousseau 2012, Rynes & Bartunek 2017). Lacking evidence for causality leaves major unfinished business. Practical application of results without evidence of causality borders on malpractice. This review is based on the tenets that causal inference is paramount to understanding organizations and that experimentation best establishes causality. If conducted in organizational settings where application is contemplated, the results are most likely to be practicable.

RARITY OF FIELD EXPERIMENTATION

Shadish & Cook (2009) traced the rise, fall, and current renaissance of field experimentation in intervention evaluation research as new solutions have been devised for earlier problems. Podsakoff & Dalton (1987) classified only 2.15% of the studies published in five leading OPOB journals in 1985 as field experiments. Scandura & Williams (2000) noted a decline in field experimentation in leading OPOB journals from the 1980s to the 1990s. The decline was not statistically significant, but that is inconsequential; the percentage dropped from only 3.9% to 2.2%. Similarly, Austin et al. (2002) noted a negligible number of field experiments in the *Journal of Applied Psychology* throughout the twentieth century. A current count would be unlikely to detect an upsurge.

Reviewers across all areas of psychology bemoan the dearth of experimentation. When there are a few experiments, reviewers usually either exclude them or meta-analyze them together with much more numerous nonexperimental studies. For example, synthesizing feedback-intervention (FI) research, Kluger & DeNisi (1996) remarked that the literature "contains only a meager proportion of studies that reported a well-controlled FI experiment . . . . The lack of control groups . . . may bias our results to an unknown degree" (p. 276–77). Stajkovic et al. (2009) observed the same in their synthesis of the collective-efficacy literature. Even reviewing the huge literature on leadership—an eminently causal variable if ever there was one—Bass & Bass (2008) lamented that causal studies "still form a distinct minority" (Bass & Bass 2008, p. 1169). Reverse causality lurks: "Supervisors may be supportive because they have productive subordinates, or subordinates may be productive because they have supportive supervisors" (Bass & Bass 2008, p. 1169). It is preposterous that, with ever-expanding research on leadership and on self-efficacy, Eden et al. (2015) could find only one field experiment on the impact of leadership on self-efficacy (Dvir et al. 2002).

Field experimentation is not on the OPOB radar screen. Many scholars errantly contrast experimental studies—meaning laboratory research—with field studies—meaning nonexperimental research—and ignore experiments in the field. Dipboye (1990) used the phrase "laboratory versus field research" meaning laboratory experiments versus field studies. Stajkovic & Luthans (1998) spoke of "laboratory settings versus field settings," meaning experimental versus nonexperimental research. However, not every laboratory study is experimental (although most are; some are observational), and some studies in the field are experiments (although most are not). Similarly, Aguinis & Vandenberg (2014) wrote, "Conducting experimental research involves many practical constraints and also often results in decreased external validity. Consequently, given the nature of organizational science research foci, we are constrained for the most part with collecting data from the field" (pp. 584–85). Note the mutually exclusive choice between "experimental research" characterized by "decreased external validity," presumably because it is conducted in the laboratory, and "collecting data from the field," presumably in nonexperimental field studies. Aguinis & Vandenberg were likely thinking of laboratory experiments versus field studies. Again, field experiments are ignored; "laboratory" and "experiment" are used interchangeably, and research done in the field is presumed to be nonexperimental. Austin et al. (2002) did not even use the term field experiment in their report on the frequency of different designs in use. Juxtaposing experiment and field as mutually exclusive is as wrong as it is widespread; experimental research in the field, or field experimentation, is ubiquitously ignored.

Fortunately, for some organization scholars, field experimentation ranks high among their method priorities. Grant (2008) reported three constructively replicating field experiments in one article testing the same hypotheses using varied treatments and measures. An outlier is Eden et al.'s (2000) report on seven field-experimental replications on diverse samples to test several theory-based workshop designs.

FIELD EXPERIMENTATION

The experimental method is the universally recognized gold standard for establishing causality. As Pinker (2011) suggested, the best way to cope with ". . . the social science rat's nest of confounded variables . . ." is by ". . .using the gold standard of science: an experimental manipulation and a matched control group" (pp. 123–24).

Experimental Treatment and Control

The sine qua non for research to be experimental is that variance is systematically produced in an independent variable, usually (but not always) in the form of an experimental treatment, followed by observation of dependent variables. Experimental treatments in organizations are usually interventions intended to improve some aspect(s) of functioning or effectiveness. Experimental treatments are commonly called "manipulations." Although not a misnomer, the term manipulation can convey excess meaning. It is often confused with deception. Experimental treatments can be free of deception (e.g., training experiments), and deception is often used with other methods (e.g., disguising the true purposes of a survey). Furthermore, manipulation suggests guile or malice that gets participants to act contrary to their own interests to favor the experimenter. This is not what we do; therefore, experimental treatment is the preferred term.

Control is indispensable to the experimental method. Physical control over the laboratory enables experimenters to reduce the "rat's nest" of potential confounders by eliminating them; they create impoverished environments and let only the independent variable(s) vary systematically. But laboratory experimenters still must use randomization to control the infinite potential confounders—individual differences—that participants bring with them to the laboratory. In contrast, the aliveness of organizational settings limits field experimenters' physical control. However, it does not reduce their experimental control; they, too, randomize.

Experimental control is widely misinterpreted. For example, Greenberg & Tomlinson (2004) wrote, ". . . because it is difficult, if not impossible, to control the impact of variables in a field experiment, they tend to lack the same high degree of control found in most lab experiments" (p. 707). In fact, control in field experiments can be as rigorous as control in laboratory experiments. Experimental control does not require dominating, constraining, or otherwise "controlling" participants; limiting their behavior; or blocking the free flow of the myriad extraneous variables in perennial flux in organizations. Rather, "control" in the experimental context derives from the Latin *contra* (against) and *rotulus* (wheel). It originated with the medieval method of checking for production errors using a duplicate register. A mainstay of textile manufacturing during the industrial revolution, the control register checks for deviations in the main register. In the modern experiment, the experimental group is compared against (contra) an untreated "control" group. This is the essence of experimental control, in the field as well as in the laboratory, and it makes field experiments every bit as internally valid as laboratory experiments (Campbell & Stanley 1966). Experimenters manipulate variables, not people, and control potentially confounding variables, not people. Thus, true experimentation does not require a control freak, but it helps if one is a randomization freak.

Randomization: Friend and Foe

Kenny (1979) highlighted randomization as "the backbone of experimental inference" (p. 1). As with laboratory experimenters, field experimenters achieve experimental control by randomizing. Randomization is a blind balancing act that controls infinite potential confounders—including those unknown to the experimenter—by creating pre-experimental equivalence. This is what imbues the experiment with supreme internal validity and makes it the gold standard. When conducted in the field, the experiment also has high external or ecological validity. This enables field experimenters to have it both ways—to grab the fabled validity stick by both ends—making the field experiment sterling gold.

But equivalence is not equality; randomization "assures" equivalence only probabilistically. Randomization is determined by the procedure used to make the participant assignments, not by the outcome of that procedure. Even properly randomized groups differ in many ways. However, the probability that they differ is accounted for by inferential statistics (Rubin 1974). Confidence in equivalence increases as the number of participants—individuals or organizational units—increases. When randomized groups do differ significantly on one or more potential confounders, Kenny (1979) dubbed this "unhappy randomization" (p. 269). But that should not make you unhappy. Do not feel "persecuted by fate" and do not think "the experiment has been doomed and there is no way to achieve valid inference" (p. 269). You can control such potentially confounding pretest differences (*a*) by analyzing those variables as covariates (Kenny 1979, Rausch et al. 2003), (*b*) by blocking on them, (*c*) by using the propensity score method (Austin 2011, Rosenbaum & Rubin 1983), or (*d*) by using a rerandomization procedure to mitigate the imbalance (Morgan & Rubin 2012). These post hoc controls preserve the randomized experiment's internal validity.

Randomization's downside is that it is inimical to organization. Teams, departments, and divisions are constructed, and tasks and missions are assigned, in ways that are anything but random. Therefore, the experimenter must randomize individuals within units or randomize subunits within larger organizational aggregations. Just as randomization controls infinite individual differences, it controls differences among organizational units. Large organizations with multiple, geographically dispersed sites are especially suitable venues for group-level field experimentation: branch banks; military, police, and first-responder organizations; schools; chain stores; government agencies; health-delivery organizations; sports clubs or leagues.

In this review, vignette experiments are classified as laboratory experiments and are not reviewed. Vignette experiments test the effects of variations in vignette content on participants' verbal responses. They do not test the effects of independent variables on behavioral dependent variables. This does not disparage vignette research; the research on "the romance of leadership" (Meindl et al. 1985) instantiates how impactful an organizational vignette study can be. Aguinis & Bradley (2014) reviewed organizational vignette experiments and suggested best practices.

Also excluded is causal (cognitive) mapping. As its name implies, it can be used to reveal what people think about causal X → Y relations (e.g., Schraven et al. 2015). But it is not behavioral mapping; it does not reveal how people act. Applied to organizations, neither vignette experiments nor causal mapping can deliver what field experiments accomplish: illuminating how independent variables affect behavioral outcomes.

IN THE BEGINNING

The earliest behavioral studies in industry predated scientific OPOB (e.g., Münsterberg 1913, Ringelmann 1913). Münsterberg advocated for a new science in which "the psychological experiment is systematically to be placed at the service of commerce and industry" (p. 3). The best-known early "experiments" are the Hawthorne studies (Roethlisberger & Dickson 1939). Their methods were primeval, and whether there really was a Hawthorne effect in those studies has been controversial (e.g., Adair 1984). Nevertheless, we should not criticize those trailblazers for not using methods that were developed later; they pioneered using what was available at the time.

Kurt Lewin led and inspired researchers that began studying leadership and group dynamics experimentally, opening up the fields of experimental-social psychology and experimental-industrial psychology (Lewin & Lippitt 1938, Lewin et al. 1939). Experiments on worker participation in decision making followed (Coch & French 1948, French et al. 1960, Marrow 1969, Morse & Reimer 1956). However, as with the Hawthorne studies, these "experiments" were actually quasi-experiments that would not pass review under modern standards. Lewin et al. (1939) were aware of the limitations: "Analysis of causal relationships . . . is still far from complete" (p. 298). Martin et al. (2013) can be viewed as a true field-experimental replication of Lewin et al. with the addition of a moderator.

Researchers followed with quasi-experiments (mislabeled "field experiments") on organizational interventions (e.g., Lawler & Hackman 1969, Rosen 1970, Scheflen et al. 1971), sometimes even with no control group (Lawler et al. 1973). The true, randomized, organizational field experiment had not yet debuted. Only in the wake of Campbell & Stanley's (1966) landmark treatise on experimental design did true organizational field experiments begin appearing.

TRUE EXPERIMENTS, QUASI-EXPERIMENTS, QUASI-FIELDS, AND COMBO DESIGNS

Experiment → Field Experiment is not a dichotomy; it's a continuum. There are gradients of control and of "fieldness" that characterize research designs. Calibrating experimentation in terms of a control gradient is widely accepted. Stone-Romero (2011) stated it succinctly: "Control is high with randomized experimental designs, moderate with quasi-experimental designs, and very low with nonexperimental designs" (p. 38). Fieldness, or the extent to which an experiment is conducted in a true field, is a new design construct.

Quasi-Experiments

Sometimes experimenters cannot—or believe they cannot—assign participants randomly to conditions. Nonrandom assignment characterizes quasi-experiments. What makes quasi-experiments experimental is their treatment of the independent variable(s). What makes them "quasi" is their lack of randomly created, pre-experimental equivalence. This degrades internal validity to an inestimable extent. Nevertheless, quasi-experiments are worth doing; more can be learned about $X \rightarrow Y$ from quasi-experiments than from observational methods that lack experimental treatments. The internal validity of a quasi-experiment waxes with the number of potential confounders evidenced not to differ between the experimental and control conditions.

Unfortunately, establishing pre-experimental equivalence without randomization is severely limited. The quasi-experimenter must measure and compare participants in all conditions on each potential confounder and then attempt to control differences statistically. But this can be done for only observed variables. Using this cumbersome procedure, quasi-experimenters can control a handful of the infinite potential confounders (see Shadish et al. 2008). However, infinity minus a handful still leaves infinite uncontrolled potential confounders. Such is the price of not randomizing. Worse still, Carlson & Wu (2012) detailed why ex post facto statistical control may be illusory.

Cook & Campbell (1979) and Shadish et al. (2002) described quasi-experimental method in detail. Grant & Wall (2009) provided excellent discussion of the "neglected science and art" of quasi-experimentation in organizations. If the true experiment is the gold standard, the quasi-experiment is the silver. King et al. (2012) wisely advised researchers to "Go for the silver" when a true experiment is impossible. Quasi-experimentation has been improved by analytical advances such as propensity-score analysis and sensitivity analysis (D'Agostino & Rubin 2000, Rosenbaum & Rubin 1983). An incentive—albeit an ignoble one—for preferring quasi-experimentation over true experimentation is the evidence that less rigorous designs are more likely to confirm hypotheses (Terpstra 1981). Below, only true, randomized experiments are dubbed "experiments."

Quasification

A quasi-experiment is worthwhile when an experiment is underway but circumstances impose shifting individual participants or units around, undoing randomization. When an experiment retrogrades to a quasi-experiment, "quasification" has occurred. [In sociology, quasification means something else (Beardsworth & Bryman 1994).] It's like gold in your hands tarnishing into silver. Adopting King et al.'s (2012) advice, don't discard the silver!

Assignment drift occurred in Eden's (1986) team-development experiment among randomized combat company command teams. One company, then another and then still another, were withdrawn from the intervention for emergency redeployment to combat zones. They were replaced by control companies, but the design could no longer be considered randomized. The study was continued as a quasi-experiment and yielded useful results that were replicated in a follow-up field experiment that maintained its original design (Eden 1985). Similarly, "changes in program management" reduced Zielhorst et al.'s (2015, p. 103) experiment to a quasi-experiment.

Thus, risk of quasification lurks. Quasification can be disheartening, but it should not be a reason to shut down a study. Quasi-experiments are second in internal validity only to true experiments.

Greenberg's (1990) theft study is an enlightening example of quasi-experimentation. Greenberg compared three similar manufacturing plants. Due to lost contracts, two plants had to cut wages temporarily; the third did not. Employees in one plant got an "adequate" explanation designed to reduce feelings of inequity; the other plant got an "inadequate" explanation. Greenberg assigned the two pay-cut plants to the two equity treatments "at random." However, $2 - 1 = 1$ *df* randomization does not do much to create pre-experimental equivalence. Greenberg did show that the plants were comparable on several relevant variables measured before the treatment. Most remarkable was Greenberg's dependent variable: theft. Employees in the adequate-explanation plant stole less than did employees in the inadequate-explanation plant, confirming the equity hypothesis. Theft rate remained constant in the control plant.

Unethical behavior can be studied easily in laboratory experiments (e.g., Greenberg 1993), but readers may dismiss them as artificial and not replicable "in the real world." Studying theft quasi-experimentally in the field establishes generalizability. Greenberg's keen eye to spot the potential for field experimentation on a sensitive variable is scientific opportunism at its best.

Quasi-Fields

Fieldness is commonly conceived dichotomously: An experiment is either a laboratory experiment or a field experiment. Classifying a study as a laboratory experiment is straightforward.

However, classifying all other experiments as field experiments is problematic because of the huge range of circumstances in which experiments are conducted outside the laboratory. Experiments among members of an organization fulfilling their organizational roles are most obviously field experiments; the organizational "field" is well-defined. For example, organizational-development experiments require "an organization."

At the other extreme, randomized experiments are conducted among strangers scattered across the nation or around the world. Such participants have no common organizational connection. They simply happened to respond to an Internet research-recruitment appeal via, say, Mechanical Turk (Buhrmester et al. 2011) or jsPsych (de Leeuw 2015), or through word of mouth (Thau et al. 2015). Participants respond online to organizationally relevant stimulus materials, but there is no "organization." Measuring attitudes in response to media-mediated stimuli does not require that the respondents share organizational membership. However, studying a scattered convenience sample is quite different from studying individuals who share membership in an organization. This huge variety of circumstances and sites strains the meaning of "field." Such research can be categorized as true experiments in "quasi-fields." New information technologies are nudging scholars toward expanded definitions of "organization."

An example of a true experiment in a quasi-field is Di Stefano et al.'s (2014) study of chefs' willingness to share culinary knowledge with other chefs. The experimenters manipulated independent variables in scenarios sent to head chefs at Michelin-listed restaurants. Although the authors called this "a scenario-based field experiment" (p. 1660), their "field" was a scattered, unrepresentative sample of respondents from a well-defined population. No organization was defined, nor was there one to define, nor was there an identifiable "field." Classification of such experiments depends on the definition of "field." In this review, such "stranger" experiments are treated as quasi-field experiments because they are conducted neither in a laboratory nor in a definable field. This lends fieldness a very broad range. Such research implies an expanded definition of organizational research. It is the field that is quasi. True experiments can be done in quasi-fields, and quasi-experiments can be done in true fields. The most valid experiments are true experiments in true fields.

Combo Designs

Inasmuch as both experimental research and field research have mutually exclusive desirable properties, scholars have proposed middle-ground "combo" designs attempting to capture some advantages of both and shed some disadvantages of both. But these are not true field-experimental designs. Goodman (1970) proposed the natural controlled experiment, aka, the part-replica design: The researcher hires people to work, but the hiring "organization" is a part-replica of an organization made to look like a bona fide business. It is not, but the hiree thinks it is; hence its façade of realism. Independent variables (e.g., pay) can be manipulated and dependent variables (sales, satisfaction) can be measured.

Antonakis et al. (2015) brought Goodman's idea to fruition in an organization that they created. They hired real people recruited through an employment agency to perform real part-time work and paid real money under conditions that enabled them to randomize and manipulate independent variables. Experimental participants heard a charismatic motivation speech; control participants heard a standard speech. The former outperformed the latter, confirming the charismatic-communication hypothesis. The situation was real, obviating the need for deception, and the internal validity was supreme.

Although genuine, Antonakis et al.'s (2015) "field" was temporary. One may wonder about generalization. Based on their findings, should we expect charismatic communication to affect performance similarly among employees reporting to their supervisors for months or years? Replication will tell. Nevertheless, these experimenters' resourcefulness in creating the field-experimental site is worthy of emulation.

Enumerating difficulties of randomizing in organizations, Lawler (1977) proposed adaptive experiments. They are not true experiments, either. They use nonequivalent control-group design (Campbell & Stanley 1966) with added features to improve measurement, strengthen relationships with the organization, and extend follow-up. Similarly, citing the dwindling frequency of organizational experiments and fearing their "extinction," Greenberg & Tomlinson (2004) proposed the situated experiment. They argued that it has the strengths of both laboratory and field experiments without the weaknesses of each. But the situated experiment actually is a laboratory experiment "in" an organization. These "adaptive" and "situated" combo designs all fall short of the strengths of true field experiments.

LAB-LIKE EXPERIMENTS IN THE FIELD

Although field experimentation may be level or declining in OPOB, it is thriving in behavioral economics and microeconomics, and it is "burgeoning" in other social and behavioral sciences (Ditlmann & Paluck 2015). Reviewing intrafirm and interfirm economics field experiments, Bandiera et al. (2011) wrote, "Field experiments are at the heart of a growing empirical literature that is expanding economists' understanding of firm behavior" (p. 78). Harrison & List (2004) developed a taxonomy of economics field experiments: conventional, artifactual, framed, and natural field experiments, thus defining an artificiality-realism continuum depending on the participants (students or "real people"), the nature of the task (tokens or real money), and whether the situation is contrived or real (natural in the lives of the participants). In this taxonomy a study is a "natural field experiment" only if the participants do not know they are in an experiment. EconPort (**http://www.econport.org/econport/request?page=web_home**) provides experimental economists with software for game-theory, market, and auction experiments. However, few economics experiments are of the natural variety: "[F]ew experimenters ever examine field behavior in a serious and large-sample way. It is relatively easy to say that the experiment could be applied to real people, but to actually do so entails some serious and often unattractive logistical problems" (Harrison & List 2004, p. 1016). Unfounded belief in the same deterrents is rife among organizational psychologists and other organizational-behavior researchers.

Early on, the natural character of field experiments was dubbed "mundane realism" (Aronson & Carlsmith 1968); it concerned ecological validity. In organizational psychology, as in economics, some field experiments have a distinctly unnatural, lab-like flavor, similar to Greenberg & Tomlinson's (2004) "situated experiment." An example is Bareket-Bojmel et al.'s (2014) study of the effects of short-term bonuses on motivation. They contrasted "actual incentives" to high-tech workers in a "real workplace." The bonuses included $25 in cash and a voucher for a $25 family pizza. Thinking of mundane realism, one wonders whether being rewarded for outstanding work with a pizza voucher instantiates "real life" in organizations.

Generalizing from such field experiments is debatable. But that does not restrain some scholars. Ariely's (2008) *New York Times* op-ed "What's the Value of a Big Bonus" discussed the effects of "multimillion-dollar compensation package(s)" for top executives at major banks and investment houses. Ariely based the article on experimental findings among students playing games with metal puzzles and tennis balls, performing arithmetic or simple mechanical tasks, and solving anagrams for performance-based bonuses of $0.50, $5, or $50, and in one experiment, $600. Ariely (2008) claimed that, "[H]igher bonuses may not only cost employers more but also discourage executives from working to the best of their ability" (p. A43). Generalizing across the vast void separating students solving puzzles for a few dollars to top executives getting multimillion-dollar salaries and bonuses for managing complex financial organizations is—at best—questionable.

GENERALIZATION

External validity concerns the generalizability—and applicability—of the research results. Scholars pursuing science for science's sake may disparage external validity (e.g., Mook 1983) but organizational researchers cannot. Our aim to produce actionable knowledge renders generalization indispensable.

Several reviewers have shown that the results of organizational laboratory and field experiments may be similar (Dipboye 1990, Locke 1986). Anderson et al. (1999) concluded the same in their review across a broad range of topics in psychology. Stoop (2014) reported convergent results in four tests of the same hypothesis in samples ranging from students in a laboratory to citizens at home unaware they were participating in an experiment. One might question expending the effort involved in field research when the laboratory will yield the same results. It is the field experiment's situs and mundane realism that imbue it with the aura of applicability.

BIG AND SMALL FIELD EXPERIMENTS

Field experiments vary in scope. They can be hard or easy to do. They can involve several investigators and many sites and participants, encompass many variables, be prolonged, and require heavy grant support, or they can be "light" in all these aspects. Shadish & Cook (2009) provided examples of how outlandishly complex and large-scale some early field experiments were.

To illustrate, consider two theory-based field experiments that tested the effects of training on reemployment. Caplan et al. (1989) conducted their experiment among nearly 1,000 unemployed individuals in Michigan; Eden & Aviram (1993) studied 66 unemployed individuals in Tel Aviv. The Michigan team was backed by the resources of the Institute for Social Research and supported by State and Federal grants; an unfunded doctoral student conducted the Tel Aviv experiment alone, with the support only of his dissertation supervisor. The Michigan training involved eight three-hour sessions over two weeks; the Tel Aviv training involved eight two-hour sessions over two-and-a-half weeks. The Michigan experimenters conducted workshops to inoculate their participants against setbacks and to preserve their motivation to seek reemployment by fortifying their job-seeking self-efficacy. The Tel Aviv experimenters focused on preventing decline in general self-efficacy due to job loss and boosting it using behavioral modeling for reemployment skills. The Michigan experimenters collected posttest data four weeks and four months after the training; the Tel Aviv experiment posttested on the final day of the workshop and two months later. Both experiments focused on self-efficacy enhancement, implemented workshop training of similar length, and emphasized job-search behavior and reemployment.

Both experiments confirmed the self-efficacy hypothesis. Both workshops boosted job search and reemployment among randomly assigned unemployed persons. Both were published in the same journal. The Tel Aviv experiment had a lower dropout rate, yielded effect sizes ten times larger than the Michigan experiment with one-tenth the sample size, and produced a higher reemployment rate. These experiments are mutually constructively replicating, as the findings converge despite differences in operationalization and population. Caplan et al. (1989, p. 768) had called for replication in "other times, settings, cultures, and subcultures"; the Tel Aviv experiment provided it.

Nevertheless, the differences are noteworthy. The Michigan experiment was rich in resources, scope, sample size, and number of variables measured. This funded long-term follow-ups that enriched knowledge of unemployment and reemployment on distal variables [e.g., wages, mental health (Vinokur et al. 1991, 1995)]. In contrast, the Tel Aviv experiment was poor in resources and more limited in scope. Nevertheless, both tested and confirmed the same basic hypothesis using field-experimental design.

Training to help unemployed persons gain reemployment continues to attract field experimenters (e.g., Noordzij et al. 2013), sometimes with even smaller samples and shorter training (e.g., Gray 1983, Rife & Belcher 1994). In a field-experimental test of self-efficacy training in a completely different situation with very different participants, Fan & Lai (2014) had a small sample and short training. They neither acknowledged a grant nor needed one to replicate field-experimentally Eden & Aviram's (1993) finding that general self-efficacy moderates training effects. Other small-sample reemployment training experiments were similarly simple yet effective (Jackson et al. 2009, Spera et al. 1994, Yanar et al. 2009).

Field experiments of this genre are not conducted "in" any particular organization. Rather, they use omnibus samples of unemployed workers who are not currently members of any work organization. They are quasi-field true experiments.

So how big must field experiments be? They can be large-scale, complex, and costly, but they don't have to be. The smallest size required based on statistical power analysis should determine the size. Ambition to maximize size is misplaced; maximize quality instead. The larger the scale of an experiment, the harder it is to control circumstances that can wreck it. Field experiments of manageable proportions should be easy to understand. KISS can mean, "Keep it small and simple."

"LITTLE DATA": UNSUNG FIELD EXPERIMENTERS

Practitioners conduct countless field experiments in organizations, but they seldom test theoretical hypotheses and get published. In-house practitioners respond to urgent demands for concrete and immediately actionable knowledge. In a 2014 SIOP (Society for Industrial and Organizational Psychology) session entitled, "Little Data: Conducting Focused Research Within Organizations," panelists from several corporations shared their experiences coping with the challenges of conducting true experiments in their companies. These scientist-practitioners must constantly contend with the oft-conflicting goals of valid research on one hand and immediate business needs on the other hand. They must do small-scale experiments fast. Hence the title "Little Data," juxtaposing their "focused" field experiments to the recent emergence of Big-Data research. ("Little data" means something different in information science and marketing research contexts; Borgman 2015).

The topics the panelists researched experimentally included training, communication, wellness-program adoption, benefits utilization, engagement, and performance management. Unfortunately, corporations covet their intellectual property rights and protect any competitive advantage that these experiments contribute. Merely to present at meetings experimenter-practitioners must obtain clearance from HR supervisors, PR departments, and corporate legal authorities. This makes for a trove of scattered, unpublished organizational field experiments of unknown size and quality that may enrich local practice but not science. Lacking cross-fertilization among experimenters (this SIOP panel was a rare exception), benefits remain parochial. The research is rarely written-up to publishable standards and shared with the academic community. Exceptions are semipublic discussions and presentations at practitioner consortia, which operate somewhat like conversational salons and serve as practitioner support groups.

Lack of open reporting widens the oft-bemoaned scientist-practitioner rift manifest in SIOP's lopsided overweighting of academe and underweighting of practice (Cascio & Aguinis 2008, Rynes

2007). Thus, any publication-based count of organizational field experiments underestimates their true volume. It would be impossible to mine the body of "Little Data" field experiments, exasperating the "file-drawer problem" (Rosenthal 1979) in attempts to synthesize organizational field experiments. We cannot judge the quantity or the quality of unpublished, in-house field experiments. Nevertheless, they epitomize the vibrant practicality of small, simple field experiments.

FACTORS MITIGATING FIELD EXPERIMENTATION

Cialdini (2009) noted three developments that have curtailed field experimentation in social psychology: the cognitive revolution, the demand of journals that manuscripts report several studies, and the requirement to include mediators. These developments disincentivize young scholars on their march to tenure from conducting field experiments that top journals may be unlikely to publish. One can do several laboratory experiments and test mediators in the time it takes to conduct one field experiment. Furthermore, true experimental mediation testing requires two experiments (Eden et al. 2015). If social psychology's predilections encroach on organizational psychology, organizational field experimentation will ebb further.

Experimenter self-efficacy is the field experimenter's key resource; its deficiency may be the major impediment to field experimentation. Current folklore has it that field experiments are hard to do: "Research in organizational settings is challenging owing to the difficulties of using experimental methods to study relevant phenomena" (Becker et al. 2015, p. 165). This misconception derives from, expresses, and perpetuates low field-experimenter self-efficacy. Becker et al. are saying that field experiments are too hard for them to do. As with other varieties of self-efficacy, enactive attainment is the major source of experimenter self-efficacy (Bandura 1997). Nothing builds field-experimenter self-efficacy more than conducting a couple field experiments successfully. One aim of this review is to strengthen readers' field-experimenter self-efficacy through verbal persuasion: If so many others can do it, so can you!

TRIANGULATION AND MIXED METHODS

To achieve greater justification for causal inference, researchers have used "multi-" and "mixed-methods" approaches (e.g., Creswell 2003, Scandura & Williams 2000, Teddlie & Tashakkori 2008). There is a journal dedicated to mixed-methods research (see Mertens 2015). Early on, Jick (1979) proposed triangulation (i.e., mixed methods) for organizational research, and it is being used more and more (e.g., Mathias & Smith 2016). Molina-Azorín (2012) showed that in strategic management mixed-methods research has more impact than single-method research.

Qualitative observation is often one of the triangulated methods (Lee 1999), as is experimentation. As Lipsey & Cordray (2000) wrote, "Qualitative methods. . .are a worthwhile adjunct to even the most comprehensive and rigorous experimental design" (p. 368). Qualitative findings constitute real-life examples of what quantitative results represent, nurturing insight regarding particular actions that produce quantitative data. This brings arid statistics to life, enriches understanding, and spices-up research reports.

When researchers do triangulate they most often use two rather than three methods and rarely call it triangulation. Typically, researchers test the same hypothesis in a pair of studies using different methods—usually a laboratory experiment and a field study (e.g., Allen & Rush 1998, Antonakis et al. 2011, Aquino et al. 2006, Grant & Rothbard 2013, Long & Christian 2015, Pastor et al. 2007, Oreg & Berson 2015, Thau et al. 2015); sometimes it is a laboratory experiment and two field studies (e.g., Mitchell & Ambrose 2012). An example of literal triangulation (i.e., three studies using different methods) is Giessner et al.'s (2009) tests of the same hypotheses in a laboratory experiment, a scenario experiment, and a field study; convergent results persuasively bolster their claim to internal and external validity. Lee et al. (2014) conducted four studies using different methods, including clever timing to vary their untreatable independent variable, the weather.

Sometimes organizational scholars test methodological issues using two experiments—one laboratory experiment to ensure internal validity and one field experiment to ensure external validity (e.g., Harrison & McLaughlin 1993, Podsakoff et al. 2015). Sometimes researchers test methodological issues using just one field experiment to achieve both internal and external validity (e.g., Aguinis et al. 2009, Alderfer 1968, Rosen 1970, Tucker & Rowe 1977, Tziner & Kopelman 1988). Some report the laboratory experiment first and the follow-up field study as a replication; others report the field study first and present the laboratory experiment as a replication. Aguinis & Molina-Azorín (2015) discussed the factors that scholars may weigh in prioritizing their mixed methods.

Combining a laboratory experiment and a field study might yield a similar degree of overall validity as a single field experiment. It can take two or three studies with various methods to obtain valid results that one field experiment can yield. Ultimately, consilience regarding any organizational topic will be reached when findings from the widest variety of methods converge.

COMMON MISCONCEPTIONS

Antiexperimental Bias

Some believe field experiments are inappropriate—not just too hard—for our science. Criticizing current research on such variables as communication and coordination, Schein (2015) opined, "Studying these phenomena experimentally has proven to be difficult because of the ethical implications of asking subjects to do things that etiquette and cultural rules prohibit.... [E]xperimentation is basically impossible with human subjects..." (p. 15). Similarly, Bernerth & Aguinis (2016) wrote, "Unfortunately, implementing experimental and quasi-experimental designs is practically difficult in organizational research due to logistical and ethical issues" (p. 230). To the extent that organization scholars share this sentiment, a glass ceiling is suppressing organizational field experimentation. This bias persists despite the many field experiments cited here, and many more not cited here.

Treatability

Some lines of organizational research involve independent variables for which treatments can be administered easily in the field. One is goals. Laboratory and field experimentation on goal-setting quickly mushroomed (Locke & Latham 1990), and goal-setting field experimentation continues (Moulton et al. 2015). Others include job-preview and new-employee orientation (Ganzach et al. 2002, Wanous & Reichers 2000), as well as pay (e.g., Bellé 2015). In contrast, a major impediment to field experimentation is that many variables cannot be altered by experimental treatments. However, when it makes sense theoretically, such variables can be studied experimentally as they naturally occur. One way is to manipulate moderators (Vancouver & Carlson 2015). Chen et al. (2009) studied the effects of stress on strain by enhancing workers' coping resources. The independent variable, stress, could not be increased ethically. Chen et al. studied its effects by altering a moderator and showed that stress was associated with strain for some participants but not others, depending on the moderator.

Another feature that limits treatability is sensitive, value-laden variables that deter many experimenters. King et al. (2012) suggested ways of field-experimenting on such topics. Examples are field experiments by Hebl et al. (2007) on hiring discrimination against pregnant women and by Bertrand et al. (2007) on institutionalized corruption. Survey research would be far less valid for such topics.

Random Assignment versus Random Sampling

Sometimes scholars confuse random assignment, which creates internal validity, with random sampling, which supports generalization. Randomization means random assignment. Randomization contributes nothing to external validity and random sampling contributes nothing to internal validity. The external validity of any experiment—laboratory or field—is limited. Although field experimenters lay claim to external validity, it hinges on the representativeness of the sample, of the treatment, and of the measures.

Many researchers treat one section of a college course (or organizational department) as the experimental group and another as the control group. However, individuals are rarely assigned to course sections or departments at random, rendering them quasi-experiments. Sometimes authors fail to state how they assigned participants to conditions (Earley & Lind 1987, Gordijn & Stapel 2008, Scandura & Graen 1984, Thibaut et al. 1974). When random assignment is not specified, assume the research was quasi-experimental.

Not Every "Experiment" is an Experiment

"Experiment" is a mislabel that too often slips by reviewers and editors. Calling nonexperimental or quasi-experimental research experimental or field-experimental is not uncommon (e.g., Azar et al. 2013, Beer & Cannon 2004, Liang et al. 2014, Mitchell et al. 1997, Morgan et al. 2013). Some of these articles reported case studies, which Bass et al. (1976) called "a succession of tryouts—what 'experiments' used to be in the original sense of the word" (p. 355). It is astonishing that authors can still conduct a quasi-experiment but nevertheless title it "a field experiment" in a top journal (e.g., Noordzij et al. 2013). The lesson is clear: Not every randomization imbues a study with internal validity and not every study tagged "experiment" is an experiment.

Labels can vary and be confusing. Aguinis et al. (2009) variously called their study "an experimental field study" and "a true field experiment." Oddly enough, Becker (1978) conducted a field experiment but labeled it "field study." That's akin to having gold but calling it silver. Beware of labels and do not rely on authors, reviewers, and editors to judge whether "experiment" properly designates research; judge for yourself. Caveat lector!

Claiming Too Many Degrees of Freedom

A common malpractice that too often gets past reviewers and editors is randomly assigning groups to conditions and then basing *df* on the much larger number of individuals in those groups. If groups were randomized, *df* should be based on the number of groups. Examples abound (Barling et al. 1996, Graen et al. 1986, Lam & Schaubroeck 2000, Meglino et al. 1988). The lesson is clear: Not every randomization imbues a study with the internal validity of an experiment. Many of these experiments preceded the advent of multilevel modeling. Its use can relieve some threats to internal validity. Reynolds & Bennett (2015) suggest proper multilevel analysis when units have been randomized but individuals within units have not. Beware how you (and others) count *df*!

Samples Are (Not) Hard to Get

A common misconception is that sites for field experimentation are hard to find. Most adherents to this misbelief have never sought sites. One ubiquitous type of setting that is especially accessible and conducive to organizational field experimenters is training. Training is a natural situation in which many potential confounders are not active, management sometimes wants evaluation, randomization is easy, dependent variables are measurable, and posttraining follow-up is often practicable. This is not limited to training evaluation. Myriad organizational topics can be studied experimentally in training settings (e.g., Aguinis et al. 2009, Chen et al. 2009, Dvir et al. 2002, Light et al. 1990, Martocchio 1994, Oz & Eden 1994, Waung 1995).

Some field experiments are conducted using convenience samples. For academics, this is often student samples. The context may be a course, a course assignment, or a game, and the school is the organization (e.g., Coman & Hirst 2015, Cruz-Cunha 2012, De Paola et al. 2014, Dewar et al. 2013, Fan & Lai 2014, Yeow & Martin 2013). In business games, multiple units are assigned to conditions randomly (Brown & Latham 2006, Thibaut et al. 1974). Sometimes the circumstances are so experimenter-friendly that the game is altered to suit the experimenters' requirements (Hägg & Johanson 1975, Seijts & Latham 2011).

Although business games simulate organizational reality and the studies can be true experiments, some deny they are field experiments because the game is contrived and the sample is not "organizational." Many of these experiments do exude a strong laboratory flavor, and generalization is a concern. Some journals even have an explicit editorial policy against publishing research on student samples (*Journal of International Business Studies*, *Journal of Management Studies*, *Journal of Occupational and Organizational Psychology*). Landers & Behrend (2015) have argued persuasively against such blanket rejection. Others argue that schools and universities are organizations.

To overcome such misgivings, some field experimenters establish generalization by testing a hypothesis among students and replicating among nonstudents, or establish internal validity by replicating a field study experimentally among students in the laboratory. Studying overconfidence in judgment regarding financial estimates, Glaser et al. (2013) conducted one experiment among students and another among traders and bankers. Earley & Lind (1987) studied procedural justice in the laboratory with students and in a field experiment in a mail-order company. Chen et al. (2003) conducted their Study 1 among MBA students and replicated it in Study 2 among employees in a financial services company. These experimenters varied not only type of participant but also tasks and measures thereby achieving a great degree of constructive replication. Stoop (2014) conducted the same experiment four times in settings ranging from students in a laboratory to ordinary citizens in their "natural field." Jones et al. (2014) conducted a laboratory experiment among students to test the effects of organizational corporate social performance and replicated it among visitors to a job fair. Friedman & Ronen (2015) tested training implementation among business students and replicated it among retail supervisors. Marr & Thau (2014) replicated a field study in two laboratory experiments. Caza et al. (2015) replicated a field study among university employees with an experiment among students. However, they did not state how they assigned students to reward-choice conditions; presumably this was a quasi-experiment.

These examples of two or more studies in sundry situations exemplify mixed-methods research. Some scholars regard their laboratory experiment as replicating their field study and others posit the field study as replicating their laboratory experiment. From the viewpoint of science, it matters not; such studies are constructively replicating (Lykken 1968).

Students may come to your aid in another way. Wheeler et al. (2014) suggested guidelines for having students recruit samples for organizational research.

OVERCOMING DETERRENTS TO FIELD EXPERIMENTATION

Typically, managers are untutored in social science and lack knowledge of experimental design and inferential statistics. They are turned off by our experimental-methodological requirements and perceive them as academic mumbo jumbo. The following advice should help overcome such resistance.

Refrain From Jargon

Do not use the word experiment in the field. Many lay managers associate experiment with other disciplines and will suspect you want to use their employees as guinea pigs. Call your experiment a study, research, or project.

Explain Randomization to Lay Managers

Most managers think they know what an experiment is. They do not, nor do they distinguish random from haphazard. To them true randomization sounds like hocus pocus. They may not say that is why they are not approving your experiment, but it deters many. Therefore, it is crucial to prepare and rehearse a 2-to-3-minute, jargon-free, lay explanation of what randomization accomplishes, how easy it is to do, and what price is paid by not randomizing. In the spirit of full disclosure, you can conduct the randomization in managers' presence to show that no magic is involved.

Capitalize on Management Indifference

Sometimes managers do not care who gets which treatment and when. Describing circumstances that are particularly conducive to field experiments, Cook & Campbell (1979) included those in which some innovation is intended for all individuals or units but cannot be delivered to all at once: "This makes it possible deliberately to plan an *experimentally staged introduction*, with chance determining in what order the innovation is received.... [T]he units changed first and last are probabilistically equivalent and can provide an experimental and control comparison" (p. 375; italics in original). This is also known as randomized-rollout or waiting-list design (Ditlmann & Paluck 2015).

An example is Eden & Moriah's (1996) internal-auditing experiment. All branches of a bank in designated regions were scheduled for auditing during the upcoming year. Having no preference for the order in which they got audited so long as they all got audited that year, management permitted us to randomize the order, creating a true field experiment. An added advantage of randomized-rollout design is built-in replication; after the posttest the treatment is administered to the control group. If it is followed by a second posttest, it is replication. At the end of six months, half had been audited, half had not, and financial-performance data were collected. Then the control branches got audited, enabling a replication test of the impact of auditing on performance in a second posttest. This design also enables detection of fade-out and sleeper effects in the original experimental group.

Taverniers et al. (2010) applied the waiting-list design. They randomly assigned special-forces candidates to high-stress and control conditions. The high-stress group was subjected to an extremely stressful exercise, and the controls were not (yet). Control candidates underwent the same high-stress treatment after the experimental candidates—and after the posttest. Management did not care who underwent the exercise first or last. All these men would have endured the same high-stress exercise had there been no experiment. The stressed candidates responded with elevated indicators of stress, validating the exercise as highly stressful.

A military organization with 54 logistics units wanted team development for some of its command teams and had no preference for how they were selected, the number to be trained depending on budget constraints. Using a randomized-block design, Eden (1985) matched the units and then randomly assigned one of each pair (i.e., block) to team-development workshops and the other to the control condition. Command's indifference regarding which units got team development made a randomized field experiment possible.

Tziner & Eden (1985) studied the effects of the combination of crew members' qualifications on crew performance. Assigning inductees to three-man tank crews, armor command had no formula for composing crews; assignments were haphazard. Command agreed to randomization. We randomly created crews with all combinations of high and low ability and motivation and tested the effects of all combinations on crew performance. The results showed that ability and motivation interact in affecting performance, with implications for optimizing team composition. Implementing the experimental design required no resources or disruption of routine. The organization invested nothing beyond goodwill, but it—and science—gained.

Using situations in which management is indifferent to aspects of changes in structure, process, equipment, or staffing for field experimentation requires only mindfulness and opportunism. These experiments require little effort on the part of experimenters and impose little or no cost on the organization. It was effortless and cost-free to schedule auditing randomly, to delay the highly stressful exercise for the randomized control group, and to render tank-crew assignments in a manner that randomly operationalized combinations of qualifications. Management would have audited all the branches, subjected all the special-forces candidates to high stress, administered team development to the same number of units, and assigned all the tank crews had there been no experiments.

Experimenters need not execute the experimental treatment by "their own hand" (manipulate from *manus*, which is Latin for hand). So long as the treatment is rendered randomly, it is an experiment. Natural occurrences in the ebb and flow of organizational events are often opportunities to take advantage of management indifference and conduct field experiments. These abundant opportunities include training, introduction of new equipment/technology, refurbishment of furniture or decor (e.g., Greenberg 1988; see below as well), managerial transitions, reorganizations, and organizational-development interventions. Scholars who are alert to these naturally occurring events and to management indifference in these situations can step up and conduct true field experiments.

When Randomization and Fairness Align

Randomization is often the fairest way to discriminate among individuals or units and it squares with our sense of justice. This is especially true when something desirable or odious is to be distributed or endured. The Vietnam War draft lottery was based on this principle, as are all lotteries.

Greenberg's (1988) equity experiment is an example of fairness-lottery randomization. Employees with varying levels of office desirability (size, carpeting, windows, etc.) were to be reassigned to offices of varying desirability in a different building for two weeks, after which they were to return to their newly refurbished offices. The temporary offices could be of higher, lower, or similar status (i.e., reward value) as their permanent offices. Management agreed to random temporary assignments. Randomization was simple, averted ill-feelings of favoritism, afforded an effective operationalization of equity, and enabled Greenberg to set up a true experiment. Greenberg scored employees as overpaid, underpaid, or equitably paid, depending on whether their temporary offices were of higher, lower, or similar status to their permanent offices. The results confirmed the equity hypothesis; overpayment enhanced performance and underpayment impaired performance.

Such opportunities are ubiquitous. All it takes to conduct a field experiment is an enterprising researcher to see and avail the opportunity.

Invert the Treatment

Sometimes the effects of a deleterious independent variable can be studied experimentally by reducing, eliminating, or inverting it. One example is stress. Institutional review boards (IRBs) disapprove proposals to subject workers to stress. However, experiments on stress abatement, buffering, or coping interventions get approved (e.g., Bruning & Frew 1987, Fan & Lai 2014, Fan & Wanous 2008, Schaubroeck et al. 1993, Zielhorst et al. 2015). The harmful effects of stress are evident in the untreated control condition; the treatment exposes these effects by preventing or alleviating them in the experimental group. That passes IRB scrutiny and tests the stress hypothesis.

Another harmful effect that cannot be studied experimentally by producing it is the Golem effect (Babad et al. 1982), that is, the debilitating effect of lowering leader expectations on follower performance. Lowering expectations would impair performance and harm participants and the organization. Inverting the treatment to vitiate low expectations, field experimenters (Davidson & Eden 2000, Oz & Eden 1994) have studied—with IRB approval—Golem effects by averting them. The experimental treatment prevented or mitigated formation of low expectations among experimental supervisors. The Golem hypothesis was confirmed by the naturally occurring low performance of the control groups in contrast (contra) to the higher-performing experimental groups. As in stress abatement, the harmful phenomenon of interest occurs in the untreated control condition and is exposed by averting it—ethically—in the experimental condition. Nature produces harmful effects ubiquitously, and field experimenters can study them by applying interventions that mitigate them.

Piggybacking on Naturally Occurring Events

Managers incessantly change things. They manipulate independent variables intending to effect improvements in dependent variables, but without experimental controls and measurement. Similarly, organizations have standard operating procedures that involve altering variables of theoretical interest. The major requisite ingredient in turning such events into field experiments is the researcher's alertness to spot such opportunities. The branch-bank auditing experiment, the tank-crew-formation experiment, and the temporary-office-assignments experiment instantiate cost-free piggybacking.

Transform Delicate Data

Field experimentation can require analysis of proprietary data that corporations understandably go to great lengths to keep cached. An example is banks, who may be unwilling to divulge performance data that could aid competitors. The way to overcome this is to ask the bank to make any linear transformation of the data that its statisticians believe will render it useless to competitors. This puts bank personnel at ease without distorting our statistical analyses. Resistance to hand over data in the branch-bank auditing experiment melted away once we (Eden & Moriah 1996) convinced management that we really had no need for the raw data. They turned their real data into what they considered garbage; for us it was gold. Similar data-transformation arrangements can be made with absence rates, units produced, police reports, intelligence information, or any other kind of sensitive data.

USING EMERGING TECHNOLOGIES TO APPROACH IDEAL DESIGN

The ideal design would yield supreme internal validity and supreme external and ecological validity: field experimentation on a representative sample. Most research falls short. Even if a sample represents all members of an organization or if all members of an organization are studied, that one organization does not represent any larger population. Macro-OB and organizational-ecology researchers sometimes study representative samples of organizations or entire populations of organizations (Carroll & Hannan 2000) and can claim that their results generalize, but micro-organizational experimenters have not done such sampling.

The advent of virtual teams, virtual organizations, and e-leadership (Avolio et al. 2014, Davenport & Daellenbach 2011, Thatcher et al. 2012) opens up new opportunities for field experimenters to maximize internal and external validity. Beal & Weiss (2003) and Uy et al. (2010) proposed using advanced technology to collect ecological momentary assessment data and for experience sampling. These are efficient means for studying within-person variance over time, which could greatly enrich field experimentation. Email, Internet, Skype, and cell phones enable efficient communication, sampling of large populations, random administration of field-experimental treatments, and cheap data collection (e.g., Agerström & Rooth 2011). An example is audit-experiment methodology (aka, audit study) in employment discrimination research (Pager 2007, Quillian 2006).

Using the similar correspondence-test approach, Milkman et al. (2012, 2015) field-experimented among a random sample of professors who responded to randomly determined information about "students" requesting office appointments via email. They found greater responsiveness among professors to white men than to women and minorities. The result cannot be attributed to uncontrolled confounders or to quirks in the sample. High internal validity and high external validity render Milkman et al.'s findings extraordinarily convincing—and disturbing: discrimination is rampant and multitudes are being treated unfairly. Gilliland et al. (2001, study 2) employed a similar method using conventional postal service.

Similar opportunities abound. Amazon's Mechanical Turk offers a range of Internet sampling and randomization opportunities (Chandler & Kapelner 2013, Lee et al. 2014, Flores et al. 2014), as does Twitter (Xu & Wu 2015). Affiliates of MIT's Abdul Latif Jameel Poverty Action Lab (aka, J-PAL) conduct randomized field experiments on a global scale, beyond the dimensions of any finite organization. Finally, Aguinis & Lawal (2012) have detailed how Internet freelancing, or e-Lancing, can be used to conduct efficient and effective field experiments on organizational issues.

Distance and online designs can be used to experiment in a well-defined organization (e.g., Aguinis et al. 2009) or in quasi-field experiments, imbuing them with a strong laboratory-like flavor. A new generation of eager scholars that master emerging technologies will increasingly merge cheap, efficient electronic tools and novel platforms with classical experimental methods (e.g., representative sampling, random assignment, delivery of experimental treatments) to overcome some of the difficulties that have been stymieing organizational field experimentation. These technologies will drastically enhance cost-effectiveness, facilitating more field experiments at less cost.

A MODEL FIELD EXPERIMENT

Glaub et al. (2014) conducted an exemplary field experiment that instantiates the attributes to strive for in any field experiment. Using a simple, randomized, pretest-posttest control group design, they tested the effectiveness of training designed to increase the personal initiative of 100 small-scale entrepreneurs in Uganda and to improve their business success. They anchored their experiment in theory that merges the evidence-based management approach with entrepreneurship (Frese et al. 2012), theoretical linkage between cognition and action to address the action-doing gap, and a theory of the role of personal initiative in entrepreneurial success (Frese 2009). The training included staged exposure to principles, practice, and trial application activities in the participants' own businesses between training sessions. As training progressed, participants practiced doing what they were being trained to do. The wait-listed control participants got the same training after the posttest. The controls knew they would get the training. Not being relegated to a no-treatment condition solely for the experimenters' benefit prevented resentful demoralization among the control participants, a design defect that can make the treatment look effective. Post-randomization experimental-control comparisons revealed no important pretest differences. Hard measures (sales, size, failure rate) gauged business success. The results confirmed the hypothesis that personal initiative, which the training had augmented, fully mediated the beneficial effect of the training on business success.

This field experiment had everything including hypothesis-testing and bringing scientific methods to the developing world. The authors livened up their report with qualitative descriptions of some participants' business successes and failures. The experimenters in no way compromised methodological rigor. They didn't have to, despite conducting their research far from home. They could even claim that work like theirs can help reduce world poverty. The case for field experimentation is valid universally, across all cultures. It is a matter of finding or making the opportunity and doing what you know how to do.

TIPS

The sidebar Tips lists what to do and what not to do, on the basis of much experience-based learning. Many of them are rarely found in methodology textbooks. Heeding them will minimize the grief you will experience when things do not go as planned; they rarely do.

The following are some do's, in more detail:

- Micro-plan your experiment. Your academic research proposal does not include nitty-gritty action steps. Before you start "selling" your experiment to an organization, make a step-by-step plan listing everything you want to do and everything you want people in the organization to do. Later they will resist doing things you did not request at the outset.
- File copies of signed, written IRB approval in several safe places. This can protect you from trouble that you cannot anticipate or even imagine until it is upon you.
- Use the power formulae to calculate needed sample size(s). Base it on number of groups if groups are to be randomized. If there are insufficient units, choose a bigger organization.
- Rehearse a three-minute lay explanation of why randomization is necessary.
- Do your own randomization. Do not rely on managers or graduate students who assure you that you can leave it to them. Provide your students a learning opportunity: Randomize with them.
- In the tradition of action research, invite clients to join as research partners. Most will be passive, but some might contribute useful ideas. Managers know much more about their organization than you ever will.
- Invite client-managers to peruse questionnaires. Involving them generates useful ideas and enhances their support for your experiment.
- Offer extra benefits [e.g., piggybacking on your questionnaire with questions clients have wanted to ask but have not had the opportunity; getting your (free) advice—if they ask for it—regarding issues troubling them that are not related to your experiment].
- Measure several potential confounders and check whether randomization created pre-experimental equivalence.

TIPS

Do's:
- Micro-plan your experiment.
- File copies of signed, written Institutional Review Board approval in several safe places.
- Use the power formulae to calculate needed sample size.
- Rehearse a three-minute lay explanation of randomization.
- Invite client-managers to join as research partners.
- Invite client-managers to peruse questionnaires.
- Offer extra benefits.
- Do your own randomization.
- Measure potential confounders to check preexperimental equivalence.
- Include a manipulation check.
- Use short forms of measures.
- Keep in mind the results tables that will test your hypotheses.
- KISS: Keep it small and simple.

Do not's:
- Do not start with a large-scale field experiment; start small and build on your success.
- Do not use the words "subject," "experiment," "random," or "manipulate" in the field.
- Do not overimpose; request only the bare-bones essentials.
- Do not promise more than you can deliver.
- Do not base *df* on the number of individuals if you randomized groups.
- Do not despair if after randomization the experimental and control groups differ.
- Do not invest time and effort writing proposals.
- Do not give up when you encounter resistance.
- Do not abandon field experimentation if your first attempt fails.

- Include a manipulation check.
- Use short forms of measures. The most costly aspect of your experiment for the organization may be the downtime when participants complete questionnaires. Minimize the organization's costs. However, use single-item measures only in the rarest of circumstances, lest you arouse "visceral reaction" in well-trained readers (including reviewers; see, e.g., Fisher et al. 2016, p. 4).
- From the proposal stage through data analysis, keep in mind the results tables that will test your hypotheses. They should guide you every step of the way, lest you end up not being able to test a key hypothesis because you left out an important measure or collected data that you cannot analyze.
- Avoid unwieldy field experiments; i.e., KISS!

The following are some do not's, in more detail:

- Do not embark on a large-scale, complicated project for your first field experiment. Start small and build on your success.
- Do not use the words subject, experiment, random, or manipulate in the field; laypersons misunderstand them.
- Do not overimpose. Request only the bare-bones essentials for conducting your experiment.
- Do not promise anything you are not absolutely certain you can deliver.
- Do not base *df* on the number of individuals if you randomized groups.

- Do not despair if after randomization (but before the treatment) the experimental and control groups differ significantly. Recall Kenny's (1979) "unhappy randomization" and take corrective action.
- Do not invest much time and effort writing proposals. Field experiments do not necessarily require budgets. Instead, invest your effort in planning and conducting simple experiments.
- Do not give up when you encounter resistance, unanticipated problems, or withdrawal of support from a backer in the organization. Any action arouses some resistance and unexpected problems always arise. Be prepared to deal with difficulties. You are not alone. Similar stressors beset colleagues conducting quasi-experiments, surveys, and field research.
- Do not abandon field experimentation if your first attempt fails. "Failures" can end up in leading journals (e.g., Eden 1986) and conferences (e.g., Hezkiau-Ludwig & Eden 2011), and they are useful pilots for your next field experiment.

THE MOTHER-IN-LAW TEST

Summarizing a century of development in industrial and organizational psychology methods, Austin et al. (2002) wrote, "The complexity of research methods has changed the capability of traditional audiences to understand the level of discourse, and it seems that the audience now consists of other I-O psychologists. The peril of this approach is the gradual lessening of the relevance of the field to previous audiences" (p. 19). Similar concerns regard growing complexity of our theoretical models (e.g., Kraut 2013). This complexity continues to accelerate. It is the analyses that are becoming ever-more complex, not the basic methods. This is unfortunate and unnecessary. As Cialdini (2009) wrote, growing complexity is causing "reduction in the clarity with which nonacademic audiences. . .can see the relevance of academic psychology to their lives . . . . When my colleagues and I have studied which messages most spur citizens to reduce household energy usage, the results don't have to be decoded or interpreted or extrapolated. The pertinence is plain" (p. 5). The adage, "Simplicity is the ultimate sophistication," is widely attributed to Leonardo. Well-done field experiments, like Cialdini's, can be simple enough to pass the mother-in-law test (MILT). That is, you should be able to explain your research to your mother-in-law or other laypersons.

THE FUTURE OF FIELD EXPERIMENTATION IN ORGANIZATIONS

The nature of experimental treatments is liable to become more complex. Field experimental data analyses are becoming more complex, even for simple field experiments (Shadish 2002). Moreover, experimenters are coming to recognize that one treatment does not fit all. Inspired by evidence-based medicine and clinical psychology, adaptive interventions will be coming into use (Nahum-Shani et al. 2012). Interventions will be individuated and adapted over time to the changing needs of particular individuals or units. Nahum-Shani et al. proposed the sequential-multiple-assignment-randomized-trial (SMART) design and procedures for analyzing SMART data. Throughout this review, I have tried to make field experimentation sound simple and easy. It can be. But the encroaching complexity of customized interventions will make experiments harder and more complicated. However, this will not displace well-conceived, simple, small-scale, field-experimental tests of specific hypotheses.

We have a chicken-and-egg quandary. Is the growing availability of sophisticated analyses for nonexperimental data causing decline in experimentation, or is waning experimentation causing intensification of efforts to develop such analyses? Stone-Romero et al. (1995) foresaw decline in organizational experimentation as software for "causal" analysis of nonexperimental data becomes widely available. They noted that it already had stoked nonexperimental studies as "some organizational researchers may assume, quite incorrectly, that it is appropriate to derive causal inferences from studies that use nonexperimental designs if the data from such studies are analyzed with CSA-based procedures (e.g., EQS, LISREL)" (Stone-Romero et al. 1995, p. 143). Conversely, observing that organizational field experimentation is declining, Edwards (2008) concluded that this "underscores the importance of determining the degree to which nonexperimental studies justify causal inferences" (p. 473). Edwards has a point, but ebbing field experimentation first and foremost underscores the importance of doing more field experiments. It seems reasonable to conclude that development of sophisticated nonexperimental causal analyses and de-emphasis of field experimentation are mutually reinforcing. We would be mortified by any move to restrict development of analytic tools to improve causal interpretation, and we should look no less askance at the marginalization of field experimentation.

Summarizing the field's progress in the first volume of the *Annual Review of Organizational Psychology and Organizational Behavior*, Porter & Schneider (2014) wrote, "It seems obvious, at least to us, that there has been substantial progress in recent years on at least one significant dimension of scholarship in our fields: namely, research design and methodology. In contrast to several decades ago, there is now greater attention to aspects of research design involving such features as the development of more sophisticated means to rule out the effects of confounding variables" (p. 15). They cited increased use of multivariate statistics. They spoke to a crucial design issue—ruling out the effects of confounders—without mentioning experimentation. This instantiates experimentation's low profile among organizational scholars.

Field experimentation seems like a seductive beauty that is hard to get. Schein (2015) noted that "[A]bstract, quantitative, and statistical methodologies are driving the research process more and more" (p. 14). As talented methodologists accelerate their efforts to develop analyses that justify causal conclusions from nonexperimental data, more effort should be invested in developing experimental methodology. As Ho & Rubin (2011) aphorized, "Research design trumps methods of analysis" (p. 17). Shadish (2010, p. 5) emphasized "the primacy of design over analysis" and cited Light et al.'s (1990, p. viii) pithy statement that "you can't fix by analysis what you bungled by design." And, of course, *the* design is the field experiment. The future is a choice: The more organizational scholars invest in developing analyses, the less field experimentation we will see. If more of the effort invested in this relentless search were invested in field experimentation, organizational research would flourish and enter into a new renaissance.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Adair JG. 1984. The Hawthorne effect: a reconsideration of the methodological artifact. *J. Appl. Psychol.* 69:334–45

Agerström J, Rooth D. 2011. The role of automatic obesity stereotypes in real hiring discrimination. *J. Appl. Psychol.* 96:790–805

Aguinis H, Bradley KJ. 2014. Best-practice recommendations for designing and implementing experimental vignette methodology studies. *Organ. Res. Methods* 17:351–71

Aguinis H, Lawal SO. 2012. Conducting field experiments using eLancing's natural environment. *J. Bus. Venturing* 27:493–505

Aguinis H, Mazurkiewicz MD, Heggestad ED. 2009. Using web-based frame-of-reference training to decrease biases in personality-based job analysis: an experimental field study. *Pers. Psychol.* 62:405–38

Aguinis H, Molina-Azorín JF. 2015. Using multilevel modeling and mixed methods to make theoretical progress in microfoundations for strategy research. *Strateg. Org.* 13:353–64

Aguinis H, Pierce CA, Bosco FA, Muslin IS. 2009. First decade of Organizational Research Methods: trends in design, measurement, and data-analysis topics. *Organ. Res. Methods* 12:69–112

Aguinis H, Vandenberg RJ. 2014. An ounce of prevention is worth a pound of cure: improving research quality before data collection. *Annu. Rev. Organ. Psychol. Organ. Behav.* 1:569–95

Alderfer CP. 1968. Comparison of questionnaire responses with and without preceding interviews. *J. Appl. Psychol.* 52:335–40

Allen TD, Rush MC. 1998. The effects of organizational citizenship—field study and a laboratory experiment. *J. Appl. Psychol.* 83:247–60

Anderson CA, Lindsay JJ, Bushman BJ. 1999. Research in the psychological laboratory: truth or triviality? *Curr. Dir. Psychol. Sci.* 8:3–9

Antonakis J, d'Adda G, Weber R, Zehnder C. 2015. *"Just words? Just speeches?" On the economic value of charismatic leadership*. Work. Pap., Dep. Organ. Behavior, Univ. Lausanne

Antonakis J, Fenley M, Liechti S. 2011. Can charisma be taught? Tests of two interventions. *Acad. Manag. Learn. Educ.* 10:374–96

Aquino K, Tripp TM, Bies RJ. 2006. Getting even or moving on? Power, procedural justice, and types of offense as predictors of revenge, forgiveness, reconciliation, and avoidance in organizations. *J. Appl. Psychol.* 91:653–68

Ariely D. 2008. What's the value of a big bonus? *New York Times*, Nov. 19, p. A43. **http://www.nytimes.com/2008/11/20/opinion/20ariely.html?_r=0**

Aronson E, Carlsmith JM. 1968. Experimentation and social psychology. In *The Handbook of Social Psychology*, Vol. 2, pp. 1–79, ed. G Lindzey, E Aronson. Reading, MA: Addison-Wesley. 2nd ed.

Austin J, Scherbaum C, Mahlman. 2002. History of research methods in industrial organizational psychology: measurement, design, analysis. In *Handbook of Research Methods in Industrial and Organizational Psychology*, ed. SG Rogelberg, pp. 77–98. Malden, MA: Blackwell

Austin PC. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav. Res.* 46:399–424. doi:10.1080/00273171.2011.568786

Avolio BJ, Sosik JJ, Kahai S, Baker B. 2014. E-leadership: Re-examining transformations in leadership source and transmission. *Leadersh. Q.* 25:105–31

Azar OH, Yosef S, Bar-Eli M. 2013. Do customers return excessive change in a restaurant? A field experiment on dishonesty. *J. Econ. Behav. Organ.* 93:219–26

Babad EY, Inbar J, Rosenthal R. 1982. Pygmalion, Galatea, and the Golem: investigations of biased and unbiased teachers. *J. Educ. Psychol.* 74:459–74

Bandiera O, Barankay I, Rasul I. 2011. Field experiments with firms. *J. Econ. Perspect.* 25(3):63–82

Bandura A. 1997. *Self-Efficacy: The Exercise of Control*. New York: Freeman

Bareket-Bojmel L, Hochman G, Ariely D. 2014. It's (not) all about the Jacksons: testing different types of short-term bonuses in the field. *J. Manag.* **https://doi.org/10.1177/0149206314535441**

Barling J, Kelloway EK, Weber T. 1996. Effects of transformational leadership training on attitudinal and financial outcomes: a field experiment. *J. Appl. Psychol.* 81:827–32

Bass BM, Bass R. 2008. *The Bass Handbook of Leadership: Theory, Research, and Managerial Applications*. New York: The Free Press

Bass BM, Cascio WF, McPherson JW, Tragash HJ. 1976. PROSPER—training and research for increasing management awareness of affirmative action in race relations. *Acad. Manag. J.* 19:353–69

Beal DJ, Weiss HM. 2003. Methods of ecological momentary assessment in organizational research. *Organ. Res. Methods* 6:440–64

Beardsworth A, Bryman A. 1994. Late modernity and the dynamics of quasification: the case of the themed restaurant. *Sociol. Rev.* 228–57

Becker LJ. 1978. Joint effect of feedback and goal setting on performance: a field study of residential energy conservation. *J. Appl. Psychol.* 63:428–33

Becker TE, Atinc G, Breaugh JA, Carlson KD, Edwards JR, Spector PE. 2015. Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *J. Organ. Behav.* 37:157–67
Beer M, Cannon MD. 2004. Promise and peril in implementing pay-for-performance. *Hum. Resour. Manag.* 43:3–20
Bellé N. 2015. Performance-related pay and the crowding out of motivation in the public sector: a randomized field experiment. *Public Adm. Rev.* 75:230–41
Bernerth JB, Aguinis H. 2016. A critical review and best-practice recommendations for control variable usage. *Pers. Psychol.* 69:229–83
Bertrand M, Djankov S, Hanna R, Mullainathan S. 2007. Obtaining a driver's license in India: an experimental approach to studying corruption. *Q. J. Econ.* 122:1639–76
Borgman CL. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT Press.
Boruch RF, Wothke W. 1985. Seven kinds of randomization plans for designing field experiments. In *Randomization and Field Experimentation*, ed. RF Boruch, W Wothke, pp. 95–118. San Francisco: Jossey-Bass
Brown TC, Latham GP. 2006. The effect of training in verbal self-guidance on performance effectiveness in a MBA program. *Can. J. Behav. Sci.* 38(1):1–11
Bruning NS, Frew DR. 1987. Effects of exercise, relaxation, and management skills training on physiological stress indicators: a field experiment. *J. Appl. Psychol.* 72:515–21
Buhrmester M, Kwang T, Gosling SD. 2011. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect. Psychol. Sci.* 6:3–5
Campbell DT, Stanley JC. 1966. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally
Caplan RD, Vinokur AD, Price RH, van Ryn M. 1989. Job seeking, reemployment, and mental health: a randomized field experiment in coping with job loss. *J. Appl. Psychol.* 74:759–69
Carlson KD, Wu J. 2012. The illusion of statistical control: control variable practice in management research. *Organ. Res. Methods* 15:413–35
Carroll GR, Hannan MT. 2000. *The Demography of Corporations and Industries*. Princeton, NJ: Princeton Univ. Press
Cascio WF, Aguinis H. 2008. Research in industrial and organizational psychology from 1963 to 2007. *J. Appl. Psychol.* 93:1062–81
Caza A, McCarter MW, Northcraft GB. 2015. Performance benefits of reward choice: a procedural justice perspective. *Hum. Resour. Manag. J.* 25:184–99
Chandler D, Kapelner A. 2013. Breaking monotony with meaning: motivation in crowdsourcing markets. *J. Econ. Behav. Organ.* 90:123–33
Chen S, Westman M, Eden D. 2009. Impact of enhanced resources on anticipatory stress and adjustment to new information technology: a field-experimental test of Conservation of Resources Theory. *J. Occup. Health Psychol.* 14:219–348
Chen Y, Brockner J, Greenberg J. 2003. When is it "A pleasure to do business with you"? The effects of relative status, outcome favorability, and procedural fairness. O*rgan. Behav. Hum. Decis. Processes* 92:1–21
Cialdini RB. 2009. We have to break up. *Perspect. Psychol. Sci.* 4:5–6
Coch L, French JRP Jr. 1948. Overcoming resistance to change. *Hum. Relat.* 1:512–32
Coman A, Hirst W. 2015. Social identity and socially shared retrieval-induced forgetting: the effects of group membership. *J. Exp. Psychol.: Gen.* 144:717–22
Cook TD, Campbell DT. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally
Creswell JW. 2003. *Research Design: Quantitative, Qualitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage. 2nd ed.
Cruz-Cunha MM. 2012. *Handbook of Research on Serious Games as Educational, Business and Research Tools*. Hershey, PA: IGI Global
D'Agostino RB, Rubin DB. 2000. Estimating and using propensity scores with partially missing data. *J. Am. Stat. Assoc.* 95:749–59
Davenport S, Daellenbach U. 2011. "Belonging" to a virtual research centre: exploring the influence of social capital formation processes on member identification in a virtual organization. *Br. J. Manag.* 22:54–76

Davidson OB, Eden D. 2000. Remedial self-fulfilling prophecy: two field experiments to prevent Golem effects among disadvantaged women. *J. Appl. Psychol.* 85:386–98

de Leeuw JR. 2015. jsPsych: a JavaScript library for creating behavioral experiments in a Web browser. *Behav. Res. Methods* 47:1–12

De Paola M, Gioia F, Scoppa V. 2014. Overconfidence, omens and gender heterogeneity: results from a field experiment. *J. Econ. Psychol.* 45:237–52

Deci EL, Koestner R, Ryan RM. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol. Bull.* 125:627–68

Dewar AJ, Kavussanu M, Ring C. 2013. The effects of achievement goals on emotions and performance in a competitive agility task. *Sport Exercise Perform. Psychol.* 2(4):250–64

Di Stefano G, King AA, Verona G. 2014. Kitchen confidential? Norms for the use of transferred knowledge in gourmet cuisine. *Strateg. Manag. J.* 35:1645–70

Dipboye RL. 1990. Laboratory versus field research in industrial-organizational psychology. In *International Review of Industrial and Organizational Psychology*, Vol. 5, ed. CL Cooper, IT Robertson, pp. 1–34. New York: Wiley

Ditlmann R, Paluck EL. 2015. Field experiments. In *International Encyclopedia of the Social and Behavioral Sciences*, Vol. 9, ed. JD Wright, pp. 128–34. Oxford: Elsevier. 2nd ed.

Dvir T, Eden D, Avolio B, Shamir B. 2002. Impact of transformational leadership on follower development and performance: a field experiment. *Acad. Manag. J.* 45:735–44

Earley PC, Lind EA. 1987. Procedural justice and participation in task selection: the role of control in mediating justice judgments. *J. Pers. Soc. Psychol.* 52:1148–60

Eden D. 1985. Team development: a true field experiment employing three levels of rigor. *J. Appl. Psychol.* 70:94–100

Eden D. 1986. Team development: quasi-experimental confirmation among combat companies. *Group Organ. Stud.* 11:133–46

Eden D, Aviram A. 1993. Self-efficacy training to speed reemployment: helping people to help themselves. *J. Appl. Psychol.* 78:352–60

Eden D, Geller D, Gewirtz A, Gordon-Terner R, Inbar I, et al. 2000. Implanting Pygmalion leadership style through workshop training: seven field experiments. *Leadersh. Q.* 11:171–210

Eden D, Moriah L. 1996. Impact of internal auditing on branch bank performance: a field experiment. *Organ. Behav. Hum. Decis. Process.* 68:262–71

Eden D, Stone-Romero EF, Rothstein H. 2015. Synthesizing the results of multiple randomized experiments to establish causality in mediation testing. *Hum. Resour. Manag. Rev.*

Edwards JR. 2008. To prosper, organizational psychology should. . .overcome methodological barriers to progress. *J. Organ. Behav.* 29:469–91

Ein-Gar D, Steinhart Y. 2011. The “Sprinter effect”: when self-control and involvement stand in the way of sequential performance. *J. Consum. Psychol.* 21:240–55

Fan J, Lai L. 2014. Pre-training perceived social self-efficacy accentuates the effects of a cross-cultural coping orientation program: evidence from a longitudinal field experiment. *J. Organ. Behav.* 35:831–50

Fan J, Wanous JP. 2008. Organizational and cultural entry: a new type of orientation program for multiple boundary crossings. *J. Appl. Psychol.* 93:1390–400

Fisher GG, Matthews RA, Gibbons AM. 2016. Developing and investigating the use of single-item measures in organizational research. *J. Occup. Health Psychol.* 21:3–23

Flores C, Bui D, Lilienthal L, Myerson J, Hale S. 2014. Age-related differences in simple and complex spans in online samples. Presented at Annu. Meet. Psychonom. Soc. 55th, Long Beach, CA, Nov. 20–23

French JRP Jr, Israel J, Ås D. 1960. An experiment on participation in a Norwegian factory. *Hum. Relat.* 13:3–20

Frese M. 2009. Towards a psychology of entrepreneurship: an action theory perspective. *Found. Trends Entrep.* 5:435–94

Frese M, Bausch A, Schmidt P, Rauch A, Kabst R. 2012. Evidence-based entrepreneurship (EBE): cumulative science, action principles, and bridging the gap between science and practice. *Found. Trends Entrep.* 8:1–62

Friedman S, Ronen S. 2015. The effect of implementation intentions on transfer of training. *Eur. J. Soc. Psychol.* 45:409–16

Ganzach Y, Pazy A, Ohayun Y, Brainin E. 2002. Social exchange and organizational commitment: decision-making training for job choice as an alternative to the realistic job preview. *Pers. Psychol.* 55:613–37

Giessner SR, van Knippenberg D, Sleebos E. 2009. License to fail? How leader group prototypicality moderates the effects of leader performance on perceptions of leadership effectiveness. *Leadersh. Q.* 20:434–51

Gilliland SW, Groth M, Baker RC, Dew AE, Polly LM, Langdon JC. 2001. Improving applicants' reactions to rejection letters: an application of fairness theory. *Pers. Psychol.* 54:669–703

Glaser M, Langer T, Weber M. 2013. True overconfidence in interval estimates: evidence based on a new measure of miscalibration. *J. Behav. Decis. Making* 26:405–17

Glaub M, Frese M, Fischer S, Hoppe M. 2014. Increasing personal initiative in small business managers or owners leads to entrepreneurial success: a theory-based controlled randomized field intervention for evidence-based management. *Manag. Learn. Educ.* 13:354–79

Gordijn EH, Stapel DA. 2008. When controversial leaders with charisma are effective: the influence of terror on the need for vision and impact of mixed attitudinal messages. *Eur. J. Soc. Psychol.* 38:389–411

Graen GB, Scandura TA, Graen MR. 1986. A field experimental test of the moderating effects of growth need strength on productivity. *J. Appl. Psychol.* 71:484–91

Grant AM. 2008. The significance of task significance: job performance effects, relational mechanisms, and boundary conditions. *J. Appl. Psychol.* 93:108–24

Grant AM, Rothbard NP. 2013. When in doubt, seize the day? Security values, prosocial values, and proactivity under ambiguity. *J. Appl. Psychol.* 98:810–19

Grant AM, Wall TD. 2009. The neglected science and art of quasi-experimentation: why-to, when-to, and how-to advice for organizational researchers. *Organ. Res. Methods* 12:653–86

Gray D. 1983. A job club for older job seekers: an experimental evaluation. *J. Gerontol.* 38:363–68

Greenberg J. 1988. Equity and workplace status: a field experiment. *J. Appl. Psychol.* 73:606–13

Greenberg J. 1990. Employee theft as a reaction to underpayment inequity: the hidden cost of pay cuts. *J. Appl. Psychol.* 75:561–68

Greenberg J. 1993. Stealing in the name of justice: informational and interpersonal moderators of theft reactions to underpayment inequity. *Organ. Behav. Hum. Decis. Process.* 54:81–103

Greenberg J, Tomlinson EC. 2004. Situated experiments in organizations: transplanting the lab to the field. *J. Manag.* 30:703–24

Giluk TL, Rynes SL. 2012. Research findings practitioners resist: lessons for management academics from evidence-based medicine. See Rousseau 2012, pp. 130–64

Goodman PS. 1970. The natural controlled experiment in organizational research. *Hum. Organ.* 29:197–203

Hägg I, Johanson J. 1975. The business game as a research tool: experiments on the effects of received information on the decision process: II. *Swedish J. Econ.* 77:351–58

Harrison DA, McLaughlin ME. 1993. Cognitive processes in self-report responses: tests of item context effects in work attitude measures. *J. Appl. Psychol.* 78:129–40

Harrison GW, List JA. 2004. Field experiments. *J. Econ. Lit.* 42:1009–55

Hebl MR, King EB, Glick P, Singletary SL, Kazama S. 2007. Hostile and benevolent reactions toward pregnant women: complementary interpersonal punishments and rewards that maintain traditional roles. *J. Appl. Psychol.* 92:1499–511

Hezkiau-Ludwig R, Eden D. 2011. *What to do when a field experiment goes awry*? Presented at Annu. Conf. Soc. Ind. Org. Psychology, 26th, Chicago

Highhouse S. 2009. Designing experiments that generalize. *Organ. Res. Methods* 12:554–66

Ho DE, Rubin DB. 2011. Credible causal inference for empirical legal studies. *Annu. Rev. Law Soc. Sci.* 7:17–40

Jackson SE, Hall NC, Rowe PM, Daniels LM. 2009. Getting the job: attributional retraining and the employment interview. *J. Appl. Soc. Psychol.* 39:973–98

Jick TD. 1979. Mixing qualitative and quantitative methods: triangulation in action. *Adm. Sci. Q.* 24:602–11

Jones DA, Willness CR, Madey S. 2014. Why are job seekers attracted by corporate social performance? Experimental and field tests of three signal-based mechanisms. *Acad. Manag. J.* 57:383–404

Kahneman D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus & Giroux

Kenny DA. 1979. *Correlation and Causality*. New York: Wiley-Intersci. Revis. ed.

King EB, Hebl MR, Morgan WB, Ahmad AS. 2012. Field experiments on sensitive organizational topics. *Organ. Res. Methods* 16:501–21

Kluger AN, DeNisi A. 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.* 119:254–84

Kraut AI. 2013. Theory, schmeory. Let's keep our models simple and strong. *Ind. Organ. Psychol.* 6:188–90

Lam SSK, Schaubroeck J. 2000. A field experiment testing frontline opinion leaders as change agents. *J. Appl. Psychol.* 85:987–95

Landers RN, Behrend TS. 2015. An inconvenient truth: arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Ind. Organ. Psychol.: Perspect. Sci. Practice* 8:142–64

Lawler EE III. 1977. Adaptive experiments: an approach to organizational behavior research. *Acad. Manag. Rev.* 2:576–85

Lawler EE III, Hackman JR. 1969. Impact of employee participation in the development of pay incentive plans: a field experiment. *J. Appl. Psychol.* 53:467–71

Lawler EE, Hackman JR, Kaufman S. 1973. Effects of job redesign: a field experiment. *J. Appl. Soc. Psychol.* 3:49–62

Lee JJ, Gino F, Staats BR. 2014. Rainmakers: why bad weather means good productivity. *J. Appl. Psychol.* 99:504–13

Lee TW. 1999. *Using Qualitative Methods in Organizational Research*. Thousand Oaks, CA: Sage

Levitt SD, List JA. 2009. Field experiments in economics: the past, the present, and the future. *Eur. Econ. Rev.* 53:1–18

Lewin K, Lippitt R. 1938. An experimental approach to the study of autocracy and democracy: a preliminary note. *Sociometry* 1:292–300

Lewin K, Lippitt R, White RK. 1939. Patterns of aggressive behavior in experimentally created "social climates." *J. Soc. Psychol.* 10:271–99

Liang LH, Adair WL, Hideg I. 2014. When should we disagree? The effect of relationship conflict on team identity in East Asian and North American teams. *Negotiation Confl. Manag. Res.* 7:a282–89

Light RJ, Singer JD, Willett JB. 1990. *By Design: Planning Research in Higher Education*. Cambridge, MA: Harvard Univ. Press

Lipsey MW, Cordray DS. 2000. Evaluation methods for social intervention. *Annu. Rev. Psychol.* 51:345–75

Locke EA. 1986. Generalizing from laboratory to field settings: research findings from industrial-organizational psychology, organizational behavior, and human resource management. Lexington, MA: Lexington Books

Locke EA, Latham GP. 1990. *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice-Hall

Long EC, Christian MS. 2015. Mindfulness buffers retaliatory responses to injustice: a regulatory approach. *J. Appl. Psychol.* 100:1409–22

Lykken DT. 1968. Statistical significance in psychological research. *Psychol. Bull.* 70:151–59

Marr JC, Thau S. 2014. Falling from great (and not-so-great) heights: how initial status position influences performance after status loss. *Acad. Manag. J.* 57:223–48

Marrow AJ. 1969. *The Practical Theorist: The Life and Work of Kurt Lewin*. New York: Basic Books

Martin S, Liao H, Campbell EM. 2013. Directive versus empowering leadership: a field experiment comparing impacts on task proficiency and proactivity. *Acad. Manag. J.* 56:1372–95

Mathias BD, Smith AD. 2016. Autobiographies in organizational research: using leaders' life stories in a triangulated research design. *Organ. Res. Methods* 19:204–30

Martocchio JJ. 1994. Effects of conceptions of ability on anxiety, self-efficacy, and learning in training. *J. Appl. Psychol.* 79:819–25

Meglino BM, DeNisi AS, Youngblood SA, Williams KJ. 1988. Effects of realistic job previews: a comparison using an "enhancement" and a "reduction" preview. *J. Appl. Psychol.* 73:259–66

Meindl JR, Ehrlich SB, Dukerich JM. 1985. The romance of leadership. *Adm. Sci. Q.* 30:78–102

Mertens DM. 2015. Mixed methods and wicked problems. *J. Mixed Methods Res.* 9(1):3–6

Milkman KL, Akinola M, Chugh D. 2012. Temporal distance and discrimination: an audit study in academia. *Psychol. Sci.* 23:710–17

Milkman KL, Akinola M, Chugh D. 2015. What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *J. Appl. Psychol.* 100:1678–1712

Mitchell MS, Ambrose ML. 2012. Employees' behavioral reactions to supervisor aggression: an examination of individual and situational factors. *J. Appl. Psychol.* 97(6):1148–70

Mitchell TR, Thompson L, Peterson E, Cronk R. 1997. Temporal adjustments in the evaluation of events: the "rosy view." *J. Exp. Soc. Psychol.* 33:421–48

Molina-Azorín JF. 2012. Mixed methods research in strategic management: impact and applications. *Organ. Res. Methods* 15:33–56

Mook DG. 1983. In defense of external invalidity. *Am. Psychol.* 38:379–87

Morgan KL, Rubin DB. 2012. Rerandomization to improve covariate balance in experiments. *Ann. Stat.* 40:1263–82

Morgan WB, Walker SS, Hebl MR, King EB. 2013. A field experiment: reducing interpersonal discrimination toward pregnant job applicants. *J. Appl. Psychol.* 98:799–809

Morse NC, Reimer E. 1956. The experimental change of a major organizational variable. *J. Abnorm. Soc. Psychol.* 52:120–29

Moulton S, Collins JM, Loibl C, Samek A. 2015. Effects of monitoring on mortgage delinquency: evidence from a randomized field study. *J. Policy Anal. Manag.* 34:184–207

Münsterberg H. 1913. *Psychology and Industrial Efficiency*. Boston, New York: Houghton Mifflin

Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, et al. 2012. Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychol. Methods* 17(4):457–77

Noordzij G, van Hooft EAJ, van Mierlo H, van Dam A, Born MPh A. 2013. The effects of a learning-goal orientation training on self-regulation: a field experiment among unemployed job seekers. *Pers. Psychol.* 66:723–55

Oreg S, Berson Y. 2015. Personality and charismatic leadership in context: the moderating role of situational stress. *Pers. Psychol.* 68:49–77

Oz S, Eden D. 1994. Restraining the golem: boosting performance by changing the interpretation of low scores. *J. Appl. Psychol.* 79:744–54

Pager D. 2007. The use of field experiments for studies of employment discrimination: contributions, critiques, and directions for the future. *Ann. Am. Acad. Pol. Soc. Sci.* 609:104–33

Pastor JC, Mayo M, Shamir B. 2007. Adding fuel to fire: the impact of followers' arousal on ratings of charisma. *J. Appl. Psychol.* 92:1584–96

Pelligra V, Stanca L. 2013. To give or not to give? Equity, efficiency and altruistic behavior in an artefactual field experiment. *J. Soc. Econ.* 46:1–9

Petersen JA, Kumar V. 2015. Perceived risk, product returns, and optimal resource allocation: evidence from a field experiment. *J. Mark. Res.* 52:268–85

Pinker S. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Penguin Books

Podsakoff NP, Maynes TD, Whiting SW, Podsakoff PM. 2015. One (rating) from many (observations): factors affecting the individual assessment of voice behavior in groups. *J. Appl. Psychol.* 100:1189–202

Podsakoff PM, Dalton DR. 1987. Research methodology in organizational studies. *J. Manag.* 13:419–41

Porter LW, Schneider B. 2014. What was, what is, and what may be in OP/OB. *Annu. Rev. Organ. Psychol. Organ. Behav.* 1:1–21

Quillian L. 2006. New approaches to understanding racial prejudice and discrimination. *Annu. Rev. Sociol.* 32:299–328

Rausch JR, Maxwell SE, Kelley K. 2003. Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *J. Clin. Child Adolesc. Psychol.* 32:467–86

Reynolds GS, Bennett JB. 2015. A cluster randomized trial of alcohol prevention in small businesses: a cascade model of help seeking and risk reduction. *Am. J. Health Promot.* 29(3):182–91

Rife JC, Belcher JR. 1994. Assisting unemployed older workers to become reemployed: an experimental evaluation. *Res. Soc. Work Practice* 4:3–13

Ringelmann M. 1913. Recherches sur les moteurs animés: travail de l'homme. *Ann. Inst. Nat. Agron.*12(2):1–40. **http://gallica.bnf.fr/ark:/12148/bpt6k54409695.image.f14.langEN**

Roethlisberger FJ, Dickson WJ. 1939. *Management and the Worker*. Cambridge, MA: Harvard Univ. Press

Rosen NA. 1970. Demand characteristics in a field experiment. *J. Appl. Psychol.* 54:163–68

Rosenbaum PR, Rubin D. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55

Rosenthal R. 1979. An introduction to the file drawer problem. *Psychol. Bull.* 86:638–41

Rubin DB. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66:688–701

Rousseau DM, ed. 2012. *Oxford Handbook of Evidence-Based Management*. Oxford, UK: Oxford Univ. Press

Rynes SL. 2007. Let's create a tipping point: what academics and practitioners can do, alone and together. *Acad. Manag. J.* 50:1046–54

Rynes SL, Bartunek JM. 2017. Evidence-based management: foundations, development, controversies, and future. *Annu. Rev. Organ. Psychol. Organ. Behav*. 4:235–61

Scandura TA, Graen GB. 1984. Moderating effects of initial leader–member exchange status on the effects of a leadership intervention. *J. Appl. Psychol.* 69:428–36

Scandura TA, Williams EA. 2000. Research methodology in management: current practices, trends, and implications for future research. *Acad. Manag. J.* 43:1248–64

Schaubroeck J, Ganster DC, Sime WE, Ditman D. 1993. A field experiment testing supervisory role clarification. *Pers. Psychol.* 46:1–25

Scheflen KC, Lawler EE III, Hackman JR. 1971. Long-term impact of employee participation in the development of pay incentive plans: a field experiment revisited. *J. Appl. Psychol.* 55:182–86

Schein EH. 2015. Organizational psychology then and now: some observations. *Annu. Rev. Organ. Psychol. Organ. Behav.* 2:1–19

Schraven DFJ, Hartmann A, Dewulf GPMR. 2015. Resuming an unfinished tale: applying causal maps to analyze the dominant logics within an organization. *Organ. Res. Methods* 18:326–49

Seijts GH, Latham GP. 2011. The effect of commitment to a learning goal, self-efficacy, and the interaction between learning goal difficulty and commitment on performance in a business simulation. *Hum. Perform.* 24(3):189–204

Shadish WR. 2002. Revisiting field experimentation: field notes for the future. *Psychol. Methods* 7:3–18

Shadish WR. 2010. Campbell and Rubin: a primer and comparison of their approaches to causal inference in field settings. *Psychol. Methods* 15:3–17

Shadish WR, Clark MH, Steiner PM. 2008. Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random to nonrandom assignment. *J. Am. Stat. Assoc.* 103:1334–43

Shadish WR, Cook TD. 2009. The renaissance of field experimentation in evaluating interventions. *Annu. Rev. Psychol.* 60:607–29

Shadish WR, Cook TD, Campbell DT. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton-Mifflin

Shavit T, Rosenboim M, Shani Y. 2014. Time preference before and after a risky activity. *J. Econ. Psychol.* 43:30–36

Spera SP, Buhrfeind ED, Pennebaker JW. 1994. Expressive writing and coping with job loss. *Acad. Manag. J.* 37:722–33

Stajkovic AD, Lee D, Nyberg AJ. 2009. Collective efficacy, group potency, and group performance: meta-analyses of their relationships, and test of a mediation model. *J. Appl. Psychol.* 94(3):814–28

Stajkovic AD, Luthans F. 1998. Self-efficacy and work-related performance: a meta-analysis. *Psychol. Bull.* 124(2):240–61

Stone-Romero EF. 2011. Research strategies in industrial and organizational psychology: nonexperimental, quasi-experimental, and randomized experimental research in special purpose and nonspecial purpose settings. In *APA Handbook of Industrial and Organizational Psychology*, Vol. 1, *Building and Developing the Organization*, ed. S Zedeck, pp. 37–72. Washington, DC: Am. Psychol. Assoc.

Stone-Romero EF, Weaver AE, Glenar JL. 1995. Trends in research design and data analytic strategies in organizational research. *J. Manag.* 21:141–57

Stoop J. 2014. From the lab to the field: envelopes, dictators and manners. *Exp. Econ.* 17:304–13

Taverniers J, Van Ruysseveldt J, Smeets J, von Grumbkow J. 2010. High-intensity stress elicits robust cortisol increases, and impairs working memory and visuo-spatial declarative memory in Special Forces candidates: a field experiment. *Stress* 13:324–34

Teddlie C, Tashakkori A. 2008. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. New York: Sage

Terpstra DE. 1981. Relationship between methodological rigor and reported outcomes in organization development evaluation research. *J. Appl. Psychol.* 66:541–43

Thatcher SMB, Brown SA, Jenkins JL. 2012. E-collaboration media use and diversity perceptions: an evolutionary perspective of virtual organizations. *Int. J. e-Collaboration* 8(2):28–46

Thau S, Derfler-Rozin R, Pitesa M, Mitchell MS, Pillutla MM. 2015. Unethical for the sake of the group: risk of social exclusion and pro-group unethical behavior. *J. Appl. Psychol.* 100:98–113

Thibaut J, Friedland N, Walker L. 1974. Compliance with rules: some social determinants. *J. Pers. Soc. Psychol.* 30:792–801

Tucker DH, Rowe PM. 1977. Consulting the application form prior to the interview: an essential step in the selection process. *J. Appl. Psychol.* 62:283–87

Tziner A, Eden D. 1985. Effects of crew composition on crew performance: Does the whole equal the sum of its parts? *J. Appl. Psychol.* 70:85–93

Tziner A, Kopelman R. 1988. Effects of rating format on goal-setting dimensions: a field experiment. *J. Appl. Psychol.* 73:323–26

Uy MA, Foo MD, Aguinis H. 2010. Using experience sampling methodology to advance entrepreneurship theory and research. *Org. Res. Methods* 13:31–54

Vancouver JB, Carlson BW. 2015. All things in moderation, including tests of mediation (at least some of the time). *Organ. Res. Methods* 18(1):70–91

Vinokur AD, Price RH, Caplan RD. 1991. From field experiments to program implementation: assessing the potential outcomes of an experimental intervention program for unemployed persons. *Am. J. Community Psychol.* 19:543–62

Vinokur AD, Price RH, Schul Y. 1995. Impact of the JOBS intervention on unemployed workers varying in risk for depression. *Am. J. Community Psychol.* 23:39–74

Wanous JP, Reichers AE. 2000. New employee orientation programs. *Hum. Resour. Manag. Rev.* 10:435–51

Waung M. 1995. The effects of self-regulatory coping orientation on newcomer adjustment and job survival. *Pers. Psychol.* 48:633–50

Weick KE. 1967. Organizations in the laboratory. In *Methods of Organizational Research*, ed. VH Vroom, pp. 1–56. Pittsburgh: Univ. Pittsb. Press

Wheeler AR, Shanine KK, Leon MR, Whitman MV. 2014. Student-recruited samples in organizational research: a review, analysis, and guidelines for future research. *J. Occup. Organ. Psychol.* 87:1–26

Xu J, Wu Y. 2015. Using Twitter in crisis management for organizations bearing different country-of-origin perceptions. *J. Commun. Manag.* 19:239–53

Yanar B, Budworth M, Latham GP. 2009. The effect of verbal self-guidance training for overcoming employment barriers: a study of Turkish women. *Appl. Psychol.* 58:586–601

Yeow J, Martin R. 2013. The role of self-regulation in developing leaders: a longitudinal field experiment. *Leadersh. Q.* 24:625–37

Zielhorst T, van den Brule D, Visch V, Melles M, van Tienhoven S, et al. 2015. Using a digital game for training desirable behavior in cognitive–behavioral therapy of burnout syndrome: a controlled study. *Cyberpsychol. Behav. Soc. Netw.* 18(2):101–11

Annual Review of Organizational Psychology and Organizational Behavior

Volume 4, 2017

Contents

Errata

An online log of corrections to *Annual Review of Organizational Psychology and Organizational Behavior* articles may be found at http://www.annualreviews.org/errata/orgpsych