

Screening Experiments and the Use of Fractional Factorial Designs in Behavioral Intervention Research

Vijay Nair, PhD, Victor Strecher, PhD, Angela Fagerlin, PhD, Peter Ubel, MD, Kenneth Resnicow, PhD, Susan Murphy, PhD, Roderick Little, PhD, Bibhas Chakraborty, MA, and Aijun Zhang, MA

Health behavior intervention studies have focused primarily on comparing new programs and existing programs via randomized controlled trials. However, numbers of possible components (factors) are increasing dramatically as a result of developments in science and technology (e.g., Web-based surveys). These changes dictate the need for alternative methods that can screen and quickly identify a large set of potentially important treatment components.

We have developed and implemented a multiphase experimentation strategy for accomplishing this goal. We describe the screening phase of this strategy and the use of fractional factorial designs (FFDs) in studying several components economically. We then use 2 ongoing behavioral intervention projects to illustrate the usefulness of FFDs. FFDs should be supplemented with follow-up experiments in the refining phase so any critical assumptions about interactions can be verified. (*Am J Public Health*. 2008;98:1354–1359. doi:10.2105/AJPH.2007.127563)

THE LANDSCAPE IN HEALTH

behavior intervention studies is changing rapidly. Recent developments in science and technology have resulted in a dramatic increase in the available types and formulations of feasible interventions and in the ways in which interventions are delivered, messages are presented, data are collected, and so on. These advances, in turn, are leading to an explosion in the number of possible treatment components (or design factors) that can be studied.

Traditional behavioral intervention studies are typically large-scale randomized controlled trials (RCTs) in which the goal is to confirm the superiority of a new program over an existing one. For example, such a trial might assess whether prostate cancer patients who receive a decision aid (e.g., an extensive online presentation about the disease) are better informed about their treatment options and more involved in their health care decisions than are patients not receiving a decision aid.

Often in such trials, the new program consists of a combination of many interventions. Decision aids, for instance, contain many different components, each of which may influence the primary outcome variables. These confirmation trials do not provide direct information on which components are active and whether they have been set at optimal levels. Post hoc analyses based on non-randomized data are usually conducted to tease out this additional information.

When RCTs are used to obtain this information, they usually involve adding or subtracting components one at a time or, at most, in small groups (e.g., 2×2 factorial designs). These studies can assess the impact of only a limited number of treatment components. By the time these findings are disseminated, the population of interest may have changed or the technology may be different (e.g., new communications media are in place or the population of interest has become more sophisticated), and as a result the conclusions may no longer be valid. All of these considerations suggest the need for alternative methodologies in health behavior research.

Over the past 5 years, the Center for Health Communications Research, funded by the National Cancer Institute, has developed and implemented a multiphase experimentation strategy for systematically studying new interventions and confirming their superiority over existing ones. Adapted from a similar framework that has been successfully used in engineering applications for many years,¹ this “multiphase optimization strategy,”² as we have labeled it, consists of 3 phases—screening, refining, and confirming—involving separate randomized trials.

The goal in the first phase is to “screen” a large set of potentially important treatment components quickly and efficiently and identify components that are in fact important. This is done through a screening experiment in which the effects of all components are examined

simultaneously. Two-level fractional factorial designs (FFDs) are useful in accomplishing this goal economically. The Pareto principle—according to which only a small subset of the components and their interactions will be important—underlies the screening phase. Thus, many interactions can be excluded a priori, increasing the efficiency of the design.

The second phase is aimed at refining understanding of the effects of the important components identified in the first phase. Existing knowledge or working assumptions need to be further examined and verified in follow-up experiments, which can untangle important effects, determine optimal “dosage” levels (i.e., appropriate levels of quantitative factors) via experiments with 3 or more levels, and so on. An optimal treatment program can be formulated from the information gained from this phase.

The final phase consists of a confirmation trial designed to compare the new program with the gold standard and assess its advantages. Although this phase is similar to RCTs with 2 arms, the multiphase approach allows inclusion of only important components at their optimized levels.

We focus on screening experiments and the use of FFDs in public health intervention research. We discuss the role of screening experiments in this context and illustrate the usefulness of FFDs. Factorial designs and FFDs have a long history.^{3–6} They were originally developed in the context of agricultural applications and have

since found widespread use in engineering. Here we provide an overview of FFDs and use 2 projects from our center to demonstrate their usefulness (more information about FFDs is available from standard textbooks^{1,7,8}).

Successful use of FFDs relies on the principle of effect sparsity. There are 2 types of sparsity, one in which few factors are active and one in which higher order interactions are negligible. One can use existing knowledge (theory, experience, or empirical evidence) in formulating working assumptions about interactions. Results from the screening experiment will suggest which of these assumptions are critical, and suitable follow-up experiments must be conducted in the refining phase to determine which groups of interactions are “aliased” (as described later).

GUIDE TO DECIDE PROJECT

The first example we use to illustrate the value of FFDs is the Guide to Decide project, which focuses on the effectiveness of decision aids for women who are at high risk of breast cancer. Tamoxifen reduces the risk of a primary diagnosis of breast cancer by 50% but has significant side effects.⁹ The decision to take tamoxifen requires that women understand the benefits (reducing their risk of developing breast cancer) versus the risks (side effects) of the drug. Women must also know their baseline risk of breast cancer.

Our goal was to determine how decision aids influence women’s knowledge of complex statistical information, their risk perceptions, and their health behaviors. The benefits of decision aids are well established.^{10,11} However, only limited research has attempted to provide an understanding of why

decision aids are effective and which of the different components (factors) contribute to better decisionmaking.

The screening phase of the study consisted of an examination of the effectiveness of 5 communication factors, each with 2 levels, in a Web-based decision aid: information presented in text only or text in combination with a pictograph (“type of information display”; factor A), risk statistics presented in a denominator of 100 or 1000 (“presentation of statistics”; factor B), information on risks presented in an incremental format (incremental risk of tamoxifen side effects) or total risk format (“risk presentation”; factor C), order of presentation of risks and benefits (“order of presentation”; factor D), and information on other health risks provided or not provided (“health risk context”; factor E). We return to this example later in the article.

FULL FACTORIAL DESIGNS

For simplicity, we restrict attention to the first 4 factors—A (type of information display), B (presentation of statistics), C (risk presentation), and D (order of presentation)—assessed in the Guide to Decide project. Table 1 shows a full factorial design corresponding to the 4 factors and all of their interactions. Because there are 16 (2⁴) possible combinations of the 4 factors each at 2 levels (high and low), there are 16 groups (rows in Table 1). The minus and plus signs under the A through D columns in Table 1 indicate the 2 settings (i.e., low or high, respectively) of the 4 factors. For example, all of the participants assigned to group 1 (row 1) will receive the treatment combination with all 4 factors (A–D) set at their low level.

Participants were assigned to the 16 groups as follows: *N* participants were randomly assigned

to the 16 groups, with *K* participants in each group. Let *N* be the total number of participants in the study. It is most efficient, in a statistical sense, to assign an equal number of participants to all groups. Therefore, let *K*=*N*÷16 be the number of participants per group.

Note that this design leads to a single randomized trial rather than 16 different trials corresponding to the 16 groups. In particular, the main effect of a factor is obtained by combining the data from all 16 groups. To illustrate this process, let *Y*₁, *Y*₂, . . . , *Y*₁₆ be the average response in each group (row in Table 1); that is, *Y*₁ is the average of the responses from the *K* participants in group 1, and so on. Then the main effect of factor A is denoted by

$$(1) ((Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 + Y_7 + Y_8) - [Y_9 + Y_{10} + Y_{11} + Y_{12} + Y_{13} + Y_{14} + Y_{15} + Y_{16}]) \div 16,$$

TABLE 1—Full Factorial Design Corresponding to 4 Factors (A–D) and Their Interactions Assessed in the Guide to Decide Project

Group	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
1	–	–	–	–	+	+	+	+	+	+	–	–	–	–	+
2	–	–	–	+	+	+	–	+	–	–	–	+	+	+	–
3	–	–	+	–	+	–	+	–	+	–	+	–	+	+	–
4	–	–	+	+	+	–	–	–	–	+	+	+	–	–	+
5	–	+	–	–	–	+	+	–	–	+	+	+	–	+	–
6	–	+	–	+	–	+	–	–	+	–	+	–	+	–	+
7	–	+	+	–	–	–	+	+	–	–	–	+	+	–	+
8	–	+	+	+	–	–	–	+	+	+	–	–	–	+	–
9	+	–	–	–	–	–	–	+	+	+	+	+	+	–	–
10	+	–	–	+	–	–	+	+	–	–	+	–	–	+	+
11	+	–	+	–	–	+	–	–	+	–	–	+	–	+	+
12	+	–	+	+	–	+	+	–	–	+	–	–	+	–	–
13	+	+	–	–	+	–	–	–	–	+	–	–	+	+	+
14	+	+	–	+	+	–	+	–	+	–	–	+	–	–	–
15	+	+	+	–	+	+	–	+	–	–	+	–	–	–	–
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Note. The columns A, B, C, and D refer to the settings (low [minus signs] or high [plus signs]) of the 4 components: A=type of information display (low = prose + pictograph, high = prose only), B = presentation of statistics (low = denominator of 100, high = denominator of 1000), C = information on risks (low = incremental risk format, high = total risk format), D = order of presentation (low = risks first, high = benefits first). The remaining columns (e.g., AB, AC) refer to the corresponding levels of the interaction effects. The first 4 elements of each row indicate the combinations of the 4 treatment components. For example, in the first combination, all 4 components are set at the low levels. The total number of possible combinations of 4 components at 2 levels each is 16, so there are 16 rows (groups or treatment combinations).

that is, multiplying the Ys by the minus and plus signs in the A column in Table 1, summing them, and then dividing by 16. Note that the main effect estimate is based on the data from all 16 groups, so the factorial design combines information across all of the groups (rows).

The columns AB (type of information display—presentation of statistics), AC (type of information display—risk presentation), and so forth in Table 1 correspond to 2-, 3-, and 4-way interaction effects. In 2-level designs, these interaction columns can be obtained through simply multiplying the corresponding main effect columns. For example, AB is obtained by multiplying columns A and B and treating the minus and plus signs as -1 and 1, respectively. The interaction effects are estimated in a manner similar to that for the main effects. For example, the AB interaction effect is denoted by

$$(2) \frac{[(Y_1 + Y_2 + Y_3 + Y_4 + Y_{13} + Y_{14} + Y_{15} + Y_{16}) - (Y_5 + Y_6 + Y_7 + Y_8 + Y_9 + Y_{10} + Y_{11} + Y_{12})]}{16},$$

that is, multiplying the Ys by the minus and plus signs in the AB interaction column, summing them, and dividing by 16.

The design in Table 1 is balanced in a number of different ways. For example, each factor occurs at low and high levels an equal number of times, and each combination of factors occurs an equal number of times (e.g., the 4 different combinations of the AB pair [minus–minus, minus–plus, plus–minus, plus–plus] all occur 4 times). This balance leads to statistical efficiency with respect to estimating main effects and interactions. Furthermore, the columns in the design matrix (Table 1) are orthogonal to each other, resulting in uncorrelated estimates.

TABLE 2—Numbers of Groups in a 2-Level Full Factorial Design as Numbers of Factors Increase

No. of Factors	No. of Groups
2	4
3	8
4	16
5	32
6	64
8	256
10	1 024
15	32 768

The problem with full factorial designs is that the number of groups increases rapidly with the number of factors and their levels. Table 2 shows the situation for 2-level factors. The problem is worse for factors with more levels; even for 3 factors at 5 levels, there are 125 (5×5×5) groups. Full factorial designs are geared toward estimating main effects and higher-order interactions. However, in many experiments, it is likely that only a small proportion of the factors are active. Also, most of the higher-order interactions will be negligible and are not of primary interest in the screening stage. As noted by Box et al., “there tends to be a redundancy in [full factorial designs]—redundancy in terms of an excess number of interactions that can be estimated and sometimes in an excess number of [components] that are studied.”^{14(p375)} FFDs exploit this redundancy, allowing the effects of additional factors to be examined economically.

HALF-FRACTION FRACTIONAL FACTORIAL DESIGNS

Suppose one wants to use an FFD with 16 groups to study all 5 Guide to Decide project factors. If the fourth-order ABCD interaction

is negligible, one can vary the fifth factor (E) according to the ABCD column in Table 1. This results in the 2 effects being “aliased”; that is, the effect of E cannot be separated from that of ABCD (E=ABCD). (If U and V are 2 effects, it can be stated that U=V if U and V are aliased.) If our assumption about the ABCD interaction is valid, then any significant effect associated with the ABCD column should be attributed to the main effect of factor E.

There are additional consequences associated with aliasing. The relationship E=ABCD implies that A=BCDE, B=ACDE, C=ABDE, and D=ABCE; in other words, each main effect is aliased with a fourth-order interaction. In addition, all 2-factor interactions are aliased with 3-factor interactions: AB=CDE, AC=BDE, AD=BCE, AE=BCD, BC=ADE, BD=ACE, BE=ACD, CD=ABE, CE=ABD, and DE=ABC.

We can estimate 2-factor interactions only if we know that the 3-factor interactions are negligible. This is reasonable in many situations. If so, we can use the 16-group design to study 5 factors simultaneously. This FFD is a half fraction of a 2⁵ full factorial. It is attractive in that all main effects are aliased with fourth-order or higher order interactions and all 2-way interactions are aliased with third-order or higher-order interactions. Thus, we can estimate all main effects and second-order interactions provided all third-order and higher-order interactions are negligible.

GUIDE TO DECIDE REVISITED

The usefulness of a half fraction for studying 5 factors in 16

groups can be illustrated with the Guide to Decide example. (A 16-group design can also be obtained as a one quarter fraction of a 2⁶ full factorial design. Later we describe how this design can be used to study 6 factors in 16 groups.) Table 3 shows the 16-group FFD, which we obtained by setting E=ABCD, used in the screening phase of the study. This design reduced the number of groups by half but allowed us to estimate all of the main effects and 2-factor interactions assuming that third-order and higher-order interactions were absent.

The screening phase of the study involved 632 women who were at high risk of having a first breast cancer diagnosis in the subsequent 5 years. Three primary outcome measures were assessed: (1) participants’ knowledge of the risks and benefits of tamoxifen, (2) their perceptions of these risks and benefits, and (3) their intentions to take additional action or seek more information.

Table 4 shows the results of our analysis for one outcome measure: knowledge of risks and benefits. Only significant main effects and interactions are shown (with the exception of the pictograph measure, for which there was a small main effect but significant interactions). The main effect of incremental risk format was significant; the negative coefficient indicated that the low level (incremental risk format) was more effective than the high level (total risk format). The significant positive interaction with pictograph showed that when the incremental risk format was used, knowledge scores were lower among women who received risk information in a text-only format but not among women who received risk information in a pictograph format.

TABLE 3—Fractional Factorial Design for the Guide to Decide Project

Group	Factor A: Information Display	Factor B: Presentation of Statistics	Factor C: Risk Presentation	Factor D: Order of Presentation	Factor E: Health Risk Context
1	Pictograph	100	Incremental	Benefits first	Present
2	Pictograph	100	Incremental	Risks first	Absent
3	Pictograph	100	Total	Benefits first	Absent
4	Pictograph	100	Total	Risks first	Present
5	Pictograph	1000	Incremental	Benefits first	Absent
6	Pictograph	1000	Incremental	Risks first	Present
7	Pictograph	1000	Total	Benefits first	Present
8	Pictograph	1000	Total	Risks first	Absent
9	Prose only	100	Incremental	Benefits first	Absent
10	Prose only	100	Incremental	Risks first	Present
11	Prose only	100	Total	Benefits first	Present
12	Prose only	100	Total	Risks first	Absent
13	Prose only	1000	Incremental	Benefits first	Present
14	Prose only	1000	Incremental	Risks first	Absent
15	Prose only	1000	Total	Benefits first	Absent
16	Prose only	1000	Total	Risks first	Present

Note. The 4 columns for factors A–D correspond to those in Table 1 (see Table 1 for fractional factorial design used here). The final column (factor E; information on other risk factors provided or not provided) corresponds to the ABCD column in Table 1. This fractional factorial design aliases the main effect of E with the ABCD interaction.

TABLE 4—Results of Analyses of Guide to Decide Participants’ Knowledge Scores

	b	p
Pictograph (vs text)	0.001	.996
Incremental risk (vs total)	-0.674	<.001
Pictograph × Incremental Risk	0.791	<.001
100 risk denominator (vs 1000)	0.493	.003
Pictograph × Risk Denominator	-0.364	.114

Note. Analyses controlled for participants’ numeracy scores. Although the main effect of pictograph was not important, it is included because of its significant interactions with other factors.

Risk denominator was also significant; use of the 1000-person denominator increased knowledge relative to use of the 100-person denominator. The Pictograph × Risk Denominator interaction suggested that pictographs could partially mediate knowledge deficits resulting from the use of 100-person denominators; however, this interaction was only marginally important.

The results from the screening phase provided clear guidelines for

the phase 2 refining experiment, which is currently under design. In phase 2, all numerical information will be presented in pictographs, and the incremental risk format will be used (because of the strong main effect of incremental risk and its interaction with pictograph). Contextual information will not be included. Phase 2 will examine 4 new components of the risk–benefit decision aid. In addition, both tamoxifen and raloxifene (a recently approved drug) will be assessed.

you smoke a tobacco cigarette in the past 7 days?”).

Six components (factors) were assessed in the project, each with 2 levels. Factor A was type of exposure (a single, large set of materials or multiple correspondences over several weeks). Factor B was outcome expectation depth (the high-depth group received individualized feedback and advice for quitting, and the low-depth group received general feedback related to motives for quitting). Factor C was success story depth (members of the high-depth group read a story tailored to their specific sociodemographic backgrounds; members of the low-depth group read a story tailored to their gender).

Factor D was efficacy expectation (the high-depth group received feedback and advice on the most significant barriers to quitting, and the low-depth group received advice on a broader range of barriers). Factor E was source personalization level (the high-depth group received personalized materials, including a photograph of and supportive text from the study HMO’s individual smoking cessation team; the low-depth group version included only a photograph). Finally, factor F was type of framing (gain framing [positive aspects of quitting] or loss framing [negative effects of continued smoking]).

A full factorial experiment with 6 factors requires 64 groups, which is not practical. With respect to the Project Quit application, previous experience suggested that factor B (outcome expectation depth) could interact with factors C (success story depth), D (efficacy expectation), and F (type of framing). In addition, we believed a priori that the DF interaction (Efficacy Expectation × Framing) might be active and that all other interactions

PROJECT QUIT

Project Quit, another Center for Health Communications Research project funded by the National Cancer Institute, focuses on smoking cessation. We use it here to demonstrate a one fourth fraction of a 64-group design. (For illustrative purposes, we present a slightly modified version of the actual experiment.) Computer-tailored programs have been used in several smoking cessation studies.^{12–15}

However, most of the research thus far has focused on whether the new treatments are effective and not on why or how they are effective. Our goal in this project was to use the multiphase framework to identify the treatment components important for smoking cessation. Phase 1 (the screening phase) was a 6-month Web-based study involving 1848 participants. The primary outcome measure was abstinence over a 7-day period (assessed via the question, “Did

would be small. Thus, we wanted to be able to estimate the BC, BD, BF, and DF interactions assuming that all other interactions were small. We could then use a 16-group design with the aliasing relationships $E=ABC$ and $F=ACD$ (when products are taken, this implies the additional aliasing $EF=BD$), representing a one fourth fraction of a 64-group FFD.

In this example, unlike the decision aid example, some of the 2-factor interactions were aliased with other 2-factor interactions. Consider the 2-factor interactions we identified a priori as possibly being active, that is, those on the left-hand sides of the following equations: $BC=AE$, $BD=EF$, $BF=DE$, and $DF=BE$. Fortunately, these interactions were aliased with other second-order interactions that were considered negligible. This was not by accident. The aliasing structure ($E=ABC$ and $F=ACD$) was selected judiciously to accomplish this goal.

Of course, our assumption about the interactions could have been wrong. Many intervention researchers view such a possibility as a severe limitation of FFDs.^{16,17} However, in our framework, FFDs are embedded within a multiphase strategy, so one can conduct follow-up experiments in the refining phase to verify any critical working assumptions (this concept is well known within the design literature; see Collins et al.² and Wu and Hamada⁸ for examples of follow-up experiments intended to resolve ambiguities in aliased interaction effects).

In the analysis, we fit the binary outcome variable (abstinence or nonabstinence) to the intervention components and several baseline sociodemographic characteristics. Motivation was the only significant sociodemographic variable.

Among the components, the effects of source personalization ($P=.027$) and success story ($P=.046$) were significant; high depth levels of both of these factors increased quit rates. Outcome expectations and efficacy expectations were marginally significant ($P=.195$ for both). The other main effects were not significant. In addition, none of the 2-factor interactions were significant (including BC, BD, BF, and DF, which we had believed might produce significant results).

These findings demonstrate the effect-sparsity principle: only 2 of the 6 components were important, and there were no significant interactions. We were able to examine several components economically using only a 16-arm trial. The refinement phase of Project Quit is under way; in this phase, the effects of added source personalization and variations in story design on smoking cessation will be examined in another randomized trial.

SUMMARY

Advances in technology (e.g., Web-based surveys) are leading to a vast increase in the number of possible treatment components in behavior intervention research. There is a need for alternative methods that can screen a large set of potential components and identify important ones that can be subsequently used in intervention programs. There is also a need for an understanding of optimal levels of program components. However, full factorial experiments require a large number of groups and thus are not practical. For example, in our projects, such an experiment would have required the design and implementation of 32 or 64 combinations of tailoring programs. FFDs are a promising alternative.

These designs are used extensively in engineering, and as shown in this article, they can also be useful in health behavior studies.

Fractional factorial experiments have been conducted in the behavior intervention literature.^{16,17} However, some of these experiments have involved fractional designs only in the loose sense of the term, without the elegant structure and statistical properties associated with the designs discussed here. The latter are sometimes referred to as regular fractional factorial designs.⁸ They involve a very simple aliasing structure in which 2 effects are completely aliased with each other. In other types of FFDs, such as Plackett–Burman designs, the aliasing structure is more complex.^{1,7,8} The interaction effects of these FFDs are not as easy to untangle in the refining phase of a multiphase experiment, so we do not recommend their use in the screening phase.

It is important to emphasize the usefulness of FFDs in health behavior studies within a multiphase approach. The iterative strategy in the multiphase approach allows researchers to further investigate the important assumptions implicit in FFDs during the refining phase. Such a strategy is critical if FFDs are to be used successfully in behavioral intervention research. ■

About the Authors

Vijay Nair, Susan Murphy, Bibhas Chakraborty, and Aijun Zhang are with the Department of Statistics, University of Michigan, Ann Arbor. Victor Strecher and Kenneth Resnicow are with the Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor. Angela Fagerlin and Peter Ubel are with the Center for Behavioral and Decision Sciences in Medicine and the Division of General Internal Medicine, University of Michigan, Ann Arbor, and the VA Health Services Research and Development Center for Practice Management and Outcomes Research, VA Ann Arbor Healthcare System, Ann Arbor.

Roderick Little is with the Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor.

Requests for reprints should be sent to Vijay Nair, PhD, Department of Statistics, University of Michigan, 439 West Hall, Ann Arbor, MI 48109 (e-mail: vnn@umich.edu).

This article was accepted January 15, 2008.

Contributors

The article was written by V. Nair with input from the other authors. The multiphase experimentation strategy and the use of fractional factorial designs were developed by V. Nair, S. Murphy, V. Strecher, and R. Little. V. Strecher, A. Fagerlin, P. Ubel and K. Resnicow were the directors of the Center for Health Communications Research projects funded by the National Cancer Institute. B. Chakraborty and A. Zhang were involved in the data analysis of the projects.

Acknowledgments

This work was supported by National Cancer Institute (grant P50 CA 101451).

We thank the Project Quit and Guide to Decide research teams at the University of Michigan, the Group Health Cooperative, and the Henry Ford Health System.

Human Participant Protection

This study was approved by the institutional review boards of the collaborating institutions and the institutional review board of the University of Michigan.

References

- Box GEP, Hunter WG, Hunter JS. *Statistics for Experimenters*. New York, NY: John Wiley & Sons Inc; 1978.
- Collins LM, Murphy SA, Nair V, Strecher V. A strategy for optimizing and evaluating behavioral interventions. *Ann Behav Med*. 2005;30:65–73.
- Yates F. *The Design and Analysis of Factorial Experiments*. Harpenden, England: Imperial Bureau of Soil Sciences; 1937.
- Fisher RA. *The Design of Experiments*. 3rd ed. Edinburgh, Scotland: Oliver & Boyd; 1942.
- Finney DJ. The fractional replication of factorial arrangements. *Ann Eugen*. 1945;12:291–301.
- Box GEP, Hunter JS. The 2^{kp} fractional factorial designs: parts I and II. *Technometrics*. 1961;3:311–351, 449–458.
- Montgomery DC. *Design and Analysis of Experiments*. 6th ed. New York, NY: John Wiley & Sons Inc; 2005.
- Wu CFJ, Hamada M. *Experiments: Planning, Analysis, and Parameter Design*

Optimization. New York, NY: John Wiley & Sons Inc; 2000.

9. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*. 1998;90:1371–1388.
10. O'Connor AM, Fiset V, Rostom A, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev*. 2001;3:CD001431.
11. O'Connor AM, Rostom A, Fiset V, et al. Decision aids for patients facing health treatment or screening decisions: systematic review. *BMJ*. 1999;319:731–734.
12. Shiffman S, Gitchell J. World's best practice in tobacco control: increasing quitting by increasing access to treatment medications: USA. *Tob Control*. 2000;9:228–236.
13. Shiffman S, Paty JA, Rohay JM, DiMarino ME, Gitchell JG. The efficacy of computer tailored smoking cessation material as a supplement to nicotine patch therapy. *Drug Alcohol Depend*. 2001;64:35–46.
14. Etter JF. Using new information technology to treat tobacco dependence. *Respiration*. 2002;69:111–114.
15. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. *Patient Educ Couns*. 1999;36:107–117.
16. West SG, Aiken LS, Todd M. Probing the effects of individual components in multiple component prevention programs. *Am J Community Psychol*. 1993; 21:571–605.
17. West SG, Aiken LS. Toward understanding individual effects in multicomponent prevention programs: design and analysis strategies. In: Bryant KJ, Windle M, West SG, eds. *The Science of Prevention: Methodological Advances From Alcohol and Substance Use Research*. Washington, DC: American Psychological Association; 1997.