

# Fractional designs: A simulation study of usefulness in the social sciences

JOHANNES A. LANDSHEER  
*Utrecht University, Utrecht, The Netherlands*

and

GODFRIED VAN DEN WITTENBOER  
*University of Amsterdam, Amsterdam, The Netherlands*

Fractional designs can be extremely useful in social science research, especially when a large number of factors is involved. Reluctance for the use of fractional designs seems to be warranted for two reasons: (1) In the social sciences, the amount of measurement error is often large, which may decrease the power, and (2) higher order interactions are assumed to be nonsignificant, which is difficult to guarantee without sufficient research. This simulation study shows the effects of measurement error and assumption violations under various conditions. It is concluded that fractional designs handle measurement error gracefully and that they are as powerful as a full design when equal degrees of freedom are available. Significant interaction effects can cause serious problems, especially in situations with low or intermediate measurement error, and can lead to erroneous conclusions. Only when estimated confounded effects are clearly not significant, the chance of a wrong decision is reasonably small. Therefore, fractional designs are especially warranted for the exclusion of irrelevant factors. However, we note pitfalls in the use of Version 1.0 of the program Trail Run from SPSS, Inc., to implement the procedures.

In early stages of research, many attributes are often assumed to be relevant. Although common sense and clear thinking may reduce the number of relevant aspects, guesswork is frequently involved. The researcher, however, wants more secure techniques. Social scientists are beginning to realize that fractional design theory, which has been developed in the technical sciences over several decades, may offer a resolution to this problem. Fractional designs are cost-effective designs, which use a reduced number of cases. They are especially useful for experiments with a large number of factors and when the relevance of these factors is uncertain and cannot be excluded a priori.

However, because fractional designs are relatively new to social scientists, the specific features of these designs are not generally known. We have used simulation studies to investigate the effect of large measurement error on fractional designs and the distorting effects of significant higher order interactions. The latter is a known weakness of fractional designs. Simulation studies can show the circumstances in which fractional designs are applicable. The results of these simulation studies clearly illustrate the advantages and disadvantages of fractional design. The methodology of the simulation studies is explained and illustrated in this paper.

Fractional designs use a reduced number of factor combinations and, therefore, cases (or runs). They do this in a systematic way, so that it remains possible to estimate main effects and sometimes second-order interaction effects; higher interaction effects cannot be estimated. Fractional designs are therefore cost-efficient factorial designs for research of a relatively large number of factors with a minimal number of cases. They are important for a first screening, whenever a relatively small number of relevant effects is anticipated among a large number of insignificant factors. The designs involve a tradeoff between power and estimability of interaction effects on the one hand and efficiency and costs on the other.

Recently, fractional designs have become more popular in the social sciences. The fractional design is often used in conjoint analysis, especially in marketing research (Carroll & Green, 1995; Louviere, 1988; Vriens, 1995). Several studies in this area deal with consumer choices (Cestre & Darmon, 1998; Kaul & Rao, 1995), but the fractional design is also used in direct mail research (Berger & Magliozzi, 1993).

The oldest and best-known instance of a fractional design is the Latin square; however, nowadays, a large variety of fractional designs is available. There are also many books devoted to the subject of fractional designs (Box, Hunter, & Hunter, 1978; Dodge, Fedorov, & Wynn, 1988; Edwards, 1993; Kempthorne, 1952/1973; Louviere, 1988; McLean & Anderson, 1984; Upperman, 1993; Winer, Brown, & Michels, 1991), some of them specifically written for use in the social sciences. Fractional designs originated in the field of industrial quality

---

We thank Joop Hox for bringing us together. Correspondence should be addressed to J. A. Landsheer, Utrecht University, Faculty of Social Sciences, Department of Methodology and Statistics, P. O. Box 80140, NL3508TC Utrecht, The Netherlands (e-mail: j.a.landsheer@fss.uu.nl).

control (Davies, 1956; Hahn, 1977), and many publications can be found in technical journals, such as *Technometrics*. For instance, they are used for research of the production quality of electronic devices (Mesenbrink & Lu, 1994). In industrial quality control, the most important rationale for using fractional designs is the separation of the vital few factors from more trivial ones (Nair, 1992).

A review of various computer programs for the construction of these fractional designs can be found in Nachtheim (1987). Recently, SPSS issued a special program (*Trial Run*) to generate these and other designs, for using efficiently a reduced number of experimental conditions at the expense of power and estimability of higher order interactions (SPSS, 1997). The usability of this program for the generation of fractional designs is critically discussed in Appendix A.

Fractional designs use a reduced number of experimental conditions in a systematic way, so that it allows researchers to estimate main effects and sometimes two-way interaction effects, while higher interaction effects are no longer estimable. Higher order interactions are intentionally confounded with second-order interaction effects and main effects. The confounding is done in such a way that there is a single variance estimate for every set of confounded effects. Main effects are estimable when they are confounded only with higher order effects but not with each other. The second-order effects are estimable when they are not confounded with each other or with the main effects but are confounded only with higher order interaction effects. The main effects and two-way interactions are estimable only when these higher order effects are known. In general, if the higher order effects are known in advance, they can be subtracted from the total of estimated variance. However, in most cases, the higher order effects are assumed to be zero. A fractional factorial design can be particularly useful for the study of main effects of a large number of independent variables.

In a full design, the number of cases that have to be studied can become impossibly large as the number of factors increases. In fractional design, the number of cases is limited by reducing the number of experimental conditions. Instead of examining the total factorial design, a fraction of the factorial design is investigated. For instance, instead of a full  $2^n$  with  $n$  factors of 2 levels, a fractional  $2^{(n-1)}$  design uses half of the cells of the full design.

When the number of cases is restricted, however, a fractional design results in reduced power to reject the null hypothesis ( $H_0$ ) when it is false (Type II error). In addition, main effects can become significant in fractional design when they are confounded with higher interactions. When this higher order interaction is assumed to be zero, while in fact it is significant, the estimate of the main effect is incorrect and will appear to be significant. In such cases, an increased probability of a Type I error (i.e., incorrect rejection of  $H_0$ ) will arise.

Research in the social sciences differs from technological research in at least one aspect: the degree of mea-

surement error. The first question of this simulation study therefore concerned the usability of fractional designs in social science research where the error variance can be large. More specifically, do fractional designs remain sufficiently powerful in situations with increased measurement error? The second question concerned the usability of fractional design when there is a significant higher order interaction. This violation of the assumption that higher order interactions should be zero was also studied under conditions of varying amounts of measurement error.

## DESIGN

To estimate the percentage of rejections in relation to measurement error, we ran two simulation studies to compare full designs with their fractional counterparts. The comparison was between a full  $2^5$  design and a fractional  $2^{(5-1)}$  design. Both designs incorporated five dichotomous variables, A through E. The full design had a total of 32 conditions, with two cases in each condition, resulting in a sample of 64 for each iteration of the simulation. The fractional design consisted of half this design with a total of 16 conditions, but with four cases in each condition. Both designs therefore incorporated a total of 64 cases in each trial. The number of cells in the full design was 32, and the number of cells in the fractional design was 16.

The Franklin algorithm was developed for the generation of fractional designs with variables of two or more levels (Franklin, 1985; Turiel, 1988). This computer algorithm selects fractional designs, which allow at least the estimation of main effects and optionally the estimation of two-way interactions. The fractional  $2^{(5-1)}$  design<sup>1</sup> is created by this algorithm and allows estimation of the five main effects, but not the estimation of two-way interactions. The defining contrast (1) = ABCE gives rise to the following pattern of confounding: (1) = ABCE, A = BCE, B = ACE, C = ABE, D = ABCDE, E = ABC, AB = CE, AC = BE, BC = AE, AD = BCDE, BD = ACDE, CD = ABDE, DE = ABCD, ABD = CDE, ACD = BDE, and BCD = ADE. It is evident that the main effects are confounded with three- or four-factor interactions, and some two-factor interactions are confounded with each other. Effects cannot be estimated separately from the effects with which they are confounded (i.e., their aliases). For instance, A and BCE refer to the same variance component and are therefore aliases of each other. Under the assumption that specific three- and five-factor interactions (e.g., BCE) are zero, all main effects (e.g., A) become estimable. Several two-factor interactions (e.g., AB and CE) are confounded with each other and cannot be estimated separately.

We selected this design with five factors of two levels each, because it is a relatively simple fractional design and can be analyzed using a regular analysis of variance (ANOVA). The latter is true for all fractional designs of dichotomous variables. Fractional designs of variables

with more than two levels need special computations when estimates of the lower interactions are required, because all interaction effects of variables with more than two levels are partitioned into two or more different components.

To get sufficiently stable results in the simulation studies, the same analysis was reiterated 1,000 times on the same number of cases. With each iteration, a full design was simulated on a random basis with 128 cases (four replications in each cell), of which half of the replications in each cell were used to get the full design with two replications. The fractional design was created by excluding those cases of the 128 that belonged to cells that were not part of the fractional design. The full and fractional designs were optimally comparable in each iteration, since both were derived from the same number of simulated respondents in each sample. The probability of rejecting  $H_0$  incorrectly is  $\alpha = .05$ . Nine error variance conditions with two ANOVAs per iteration and two interaction conditions (no interaction and interaction) gives a total of two times 18,000 ANOVA results to be analyzed. The automated SPSS 8.0 setup for this simulation study is shown in Appendix B.<sup>2</sup>

All five main effects in the first simulation study were kept at the same level, but the extent of error variance (reflected by the MS WITHIN) has been varied from small (1 unit) to large (1,000 units). Though always present, main effects may or may not be significant as a result of unsystematic sampling error due to the error variance. In the first simulation, the data are generated according to the formula (in SPSS syntax): COMPUTE Y = A + B + C + D + E + NORMAL (SD). This formula defines the outcome variable Y, by adding the main effects A to E and some measurement error. The measurement error is defined as unsystematic error, which is normally distributed with a standard deviation SD. SD is varied between 1 and 1,000. NORMAL (SD) means that a normal distributed error is added to the main effects with zero mean and standard deviation SD. Thus, SD represents the standard deviation of the error, which results in an expected MS WITHIN ranging from 1 to 1,000.

The MS BETWEEN depends on the factors A to E (with possible values of -1 and 1 to get orthogonal factors). The use of this formula results in an anticipated contribution to the MS BETWEEN of 64 in the full designs with two replications and 64 in the fractional design with four replications.

Thus, the simulation includes various conditions, from strong effects (in which the extent of systematic variance is substantially larger than the extent of unsystematic error variance) to unpredictable effects (in which unsystematic error variance becomes comparable, though smaller than the systematic variance) and nonsignificant effects (in which the systematic variance is equivalent to the unsystematic error variance or less). In both the full design and the fractional design, occasionally effects can be significant due to sample differences, even in the simulations with a large amount of error. The number of cases is kept

equal at each level of error variance by comparing full designs with two cases per cell with fractional designs with four cases per cell but half of the cells. At each iteration, differences in outcome between both designs are therefore the result of differences between the designs and sampling differences from the normal distributions.

The second simulation study focuses on the effect of a three-way interaction effect in combination with error variance. The data are generated according to the formula (in SPSS syntax) COMPUTE Y = A + B + C + D + A\*B\*C + NORMAL (SD), in which the interaction effect A\*B\*C has been added, but the main effect E has been omitted. The MS BETWEEN of the interaction effect is of the same magnitude as the main effects (64). Since higher order interaction effects in fractional designs are normally assumed to be 0, this formula involves a clear violation of this assumption in the fractional designs. In the fractional design used here, the ABC effect is confounded with main effect E that has been left out. Erroneous results are therefore to be expected when using this fractional design.

## RESULTS

Table 1 shows the results of the first simulation with all main effects present. The smallest degree of error (SS WITHIN = 1) leads in both the full design and the fractional design to 100% identical decisions about the existence of the five main effects in each of the 1,000 iterations.

With a scarcely higher degree of error (SS WITHIN = 4), both the full design and the fractional design lead in 95.4% of the main effects to a simultaneous rejection of  $H_0$ . Opposite decisions, however, arrived at 4.1% of the total of 5,000 main effects. When the SS WITHIN increases, the number of opposite decisions grows rapidly to 33.1% when SS WITHIN is 16. In the intermediate area, with an SS WITHIN of 16 to 21.33 and an SS BETWEEN of 64 for each main effect, the value of  $F(1,58)$  approaches the critical value of 4 more often, as a result of which the random differences more frequently cross the borderline between rejection and no rejection of  $H_0$  at either side. This gives rise to a large number of opposite decisions between the full design and the fractional design. When SS WITHIN becomes larger, the  $F$  value diminishes further, leading to more simultaneous failures to reject  $H_0$ . With SS BETWEEN = 1,000, the full design signals that 94.6% of the main effects are not significant, and the fractional design fails to reject  $H_0$  in 94.4% of all main effects. With this large and unambiguous degree of error, the percentage of opposite decisions given by the full design and the fractional design is reduced to 8.8%.

Table 1 also shows that the power to reject  $H_0$  is about the same in a fractional design and in a full design. The frequency of opposite decisions depends on the design of SS WITHIN and is especially high in the middle area. Although differences occur, the percentage of rejections is about the same for both designs.

**Table 1**  
**Results of the First Simulation Study: Percentages of Rejections of  $H_0$  Concerning the Main Effects of a Full  $2^5$  Design and a Fractional  $2^{(5-1)}$  Design, Both With 64 Cases per Run**

	Full design	Reject	No Reject	Reject	No Reject	Opposite decision
	Fractional design	Reject	Reject	No Reject	No Reject	
SS WITHIN						
1		100.0	0	0	0	0
4		95.4	2.1	2.0	0.4	4.1
8		68.5	11.2	11.0	9.3	22.2
16		33.8	17.0	16.1	33.1	33.1
20		25.1	16.8	15.9	42.3	32.7
21.33		22.2	15.4	17.3	45.1	32.7
32		14.2	14.2	14.1	57.6	28.3
100		4.2	8.1	8.0	79.7	16.1
1,000		1.0	4.6	4.2	90.2	8.8

Note—In both the full design and the fractional design, the five main effects define the systematic variance, whether the amount of SS WITHIN is varied.

Table 2 shows the results of the second simulation study of the effects of assumption violation. The direct effect of the single A\*B\*C interaction effect is that the fractional design results in an erroneous detection of the confounded, but nonexistent main effect E (a Type I error), and the full design concludes correctly that there is no main effect of the factor E.

The full design rejects  $H_0$  concerning all entered main effects (i.e., A, B, C, and D) and does not reject  $H_0$  concerning a main effect E; the fractional design, on the other hand, rejects  $H_0$  for all five main effects. This leads to a large difference in results. The percentage of opposite decisions is 19% in the extreme case with an SS WITHIN of 1, and, interestingly enough, it is reduced to 9% when SS WITHIN is large (1,000). Comparing Table 1 with Table 2, the frequency of opposite decisions is much greater in Table 2 at the extremely low error rates, but the difference diminishes rapidly. It is roughly the same when SS WITHIN is 21.33 or higher. In other words, although the probability of a Type I error for specific main effects increases seriously in a fractional design when confounded higher order interaction effects are present, this probabil-

ity diminishes when the level of error increases and the magnitude of the effect remains the same.

### CONCLUSIONS AND DISCUSSION

When compared with the full design, the fractional design is equally powerful as long as the total number of cases is the same. However, the fractional design is most often employed in order to reduce the minimum number of cases needed to use the design. Precisely in that situation, of course, the power of the design is diminished by the reduced number of cases.

The first simulations show that, when the presence and size of the effects are equal and the total number of cases is the same, the chances that the same decision of rejection of  $H_0$  will be made with a fractional design and a full design depend on the degree of error. The number of equal decisions can be fairly low, especially in cases where measurement error is intermediate. When SS WITHIN is small, the divergence between the full design and the fractional design is relatively small. When measurement error is intermediate, the divergence reaches a maximum

**Table 2**  
**Results of the Second Simulation Study: Percentages of Rejections of  $H_0$  Concerning the Main Effects of a Full  $2^5$  Design and a Fractional  $2^{(5-1)}$  Design, Both With 64 Cases per Run**

	Full design	Reject	No Reject	Reject	No Reject	Opposite decision
	Fractional design	Reject	Reject	No Reject	No Reject	
SS WITHIN						
1		81.0	0.0	19.0	0.0	19.0
4		77.5	1.8	20.2	0.6	22.0
8		55.2	9.4	24.1	11.3	33.5
16		27.5	14.0	22.8	35.7	36.8
20		21.1	14.7	19.9	44.3	34.7
21.33		19.6	13.2	20.4	46.8	33.6
32		11.0	11.9	17.4	59.8	29.3
100		3.4	6.8	9.0	80.7	15.8
1,000		0.9	4.6	4.4	90.0	9.0

Note—In both designs, the systematic variance is defined by the four main effects A, B, C, and D and the interaction effect ABC, whereas the amount of SS WITHIN is varied.

when the  $F$  value becomes critical and rejection of  $H_0$  becomes strongly dependent on sample differences. When measurement error is high, the divergence between the full design and the fractional design diminishes. This result is clearly illustrated by the first simulation. Fractional designs are the most vulnerable when the  $F$  value is very near the borderline between a significant effect and a non-significant effect, but the same holds true for full designs.

The second simulation shows that assumption violations can easily lead to failures to reject  $H_0$  for confounded effects and therefore to a large Type I error. In general, these violations have to be examined with care, and research using fractional designs can be considered safe only when previous research has demonstrated that the higher interactions in the design can be neglected. The second simulation study shows that when there is no information from other studies to show the nonexistence of higher order interaction effects, there is a real danger of misinterpreting analysis results. In such cases, confounded significant main effects in fractional designs have to be distrusted, especially when the error variance is small. Such a main effect can be due to confounded interaction effects. Of course, we can assume that higher order interactions become more exceptional as the order gets higher, but it is better to have information available about the existence of higher order interactions from previous research. However, when a main effect is not statistically significant, we can be fairly certain that such a specific main effect really is absent. The too small SS for the total set of confounded effects cannot be the result of the interaction effect. Therefore, fractional designs are especially useful for the exclusion of clearly irrelevant factors.

In conjoint analysis, fractional designs are used in repeated measurement designs, with several or all conditions of the design assigned to every individual case. The power of these designs is essentially the same as the designs simulated here. However, with conjoint analysis the confounding pattern has seldom been presented, and, in many cases, it is wholly undefined. When there is any chance that higher order interactions may be significant, it is unclear which main effects are vulnerable. This makes the argument even stronger that the value of higher interaction effects should be known in advance from previous research. Having said that, clear failures to reject  $H_0$  can be trusted, even where some higher interactions are suspected. Again, fractional designs can best be used as a screening device to exclude clearly nonsignificant effects.

Both simulation studies clearly illustrate the features of the fractional design. As these studies are easily accomplished, they are highly recommended for researchers who want to acquaint themselves with the features of fractional designs. Our approach is suited for all fractional designs with factors of two levels, because, in that case, common ANOVA is suitable for the full analysis of these fractional designs, and SPSS MANOVA or GLM can be used for analysis of both full and fractional designs. When the researcher is interested in main effects only, this simulation approach can also be applied to factors with three lev-

els or more, since common ANOVA is also appropriate for the analysis of main effects of fractional designs with these factors.

## REFERENCES

- BERGER, P. D., & MAGLIOZZI, T. L. (1993). Experimental design in direct mail and the application of Taguchi methods. *Journal of Direct Marketing*, 7, 44-55.
- BOX, G. E. P., HUNTER, W. G., & HUNTER, J. S. (1978). *Statistics for experimenters*. New York: Wiley.
- CARROLL, J. D., & GREEN, P. E. (1995). Psychometric methods in marketing research: I. Conjoint analysis. *Journal of Marketing Research*, 32, 385-391.
- CESTRE, G., & DARMON, R. Y. (1998). Assessing consumer preferences in the context of new product diffusion. *International Journal of Research in Marketing*, 13, 139-161.
- DAVIES, O. L. (1956). *The design and analysis of industrial experiments*. Edinburgh: Oliver & Boyd.
- DODGE, Y., FEDOROV, V. V., & WYNN, H. P. (1988). Optimal design of experiments: An overview. In Y. Dodge, V. V. Fedorov, & H. P. Wynn (Eds.), *Optimal design and analysis of experiments* (pp. 1-11). Amsterdam: North-Holland.
- EDWARDS, L. K. (Ed.) (1993). *Applied analysis of variance in behavioral science*. New York: Marcel Dekker.
- FRANKLIN, M. F. (1985). Selecting defining contrasts and confounded effects in  $p^n - m$  factorial experiments. *Technometrics*, 27, 165-172.
- HAHN, G. J. (1977). Some things engineers should know about experimental design. *Journal of Quality Control*, 8, 375-384.
- KAUL, A., & RAO, C. (1995). Research for product positioning and design decisions: An integrative review. *International Journal of Research in Marketing*, 12, 293-320.
- KEMPTHORNE, O. (1973). *The design and analysis of experiments*. Huntington, NY: Krieger. (Original work published 1952)
- KIRK, R. E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole.
- LOUVIERE, J. J. (1988). *Analyzing decision making: Metric conjoint analysis*. Newbury Park, CA: Sage.
- MCLEAN, R. A., & ANDERSON, V. L. (1984). *Applied factorial and fractional designs*. New York: Marcel Dekker.
- MESENBRINK, P., & LU, J. C. (1994). Characterization and optimization of a wave-soldering process. *Journal of the American Statistical Association*, 89, 1209-1218.
- NACHTHEIM, C. J. (1987). Tools for computer-aided design of experiments. *Journal of Quality Control*, 19, 132-160.
- NAIR, V. N. (1992). Experimental strategy and planning techniques. *Technometrics*, 34, 142-147.
- SPSS (1997). *Trial Run 1.0 user's guide*. Chicago: SPSS.
- TURIEL, T. P. (1988). A FORTRAN program to generate fractional factorial experiments. *Journal of Quality Technology*, 20, 63-72.
- UPPERMAN, P. M. (1993). *Designs with a small number of runs for factorial experiments*. Unpublished doctoral dissertation, Technical University Eindhoven.
- VRIENS, M. (1995). *Conjoint analysis in marketing: Developments in stimulus representation and segmentation methods*. Unpublished doctoral dissertation, Groningen University.
- WINER, B. J., BROWN, D. R., & MICHELS, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

## NOTES

1. The Windows 95/98 program to generate these and other fractional designs is available in a beta version from the first author. It will be published once it has been fully debugged and will be made available at no charge.

2. A simple Windows 95/98 program to select the relevant data from the draft output (text-only) of SPSS is available from the first author. The program is in Delphi Pascal source code and needs to be adapted for specific purposes, and some Pascal programming experience is necessary for its use.

**APPENDIX A**  
**Puzzling Aspects in the Construction and Analysis of**  
**Fractional Factorial Designs With Computer Programs From SPSS**

As the number of factors increases, the construction of fractional designs becomes more complex. Furthermore, when the design involves a nonprime number of levels or when the number of levels varies per factor, the construction and the analysis of these designs become quite complicated (Upperman, 1993, offers a number of solutions in these special cases). The construction of a fractional design is therefore a complicated task, calling for a dedicated computer program. SPSS offers such a program called *Trial Run 1.0*, which is a tool for generating experimental designs. Trial Run uses the SPSS module GLM for analyzing data, obtained in these experiments. The program offers solutions in even the most difficult cases and is very complete with a huge number of often complicated designs. The effort of SPSS to create such a complete program and to make it available to the general public is highly commendable. The specialized knowledge of designs involved puts it at the cutting edge. However, though it is a program that offers much for experimenters, we also have found some shortcomings, which call for careful usage.

**Notational Problems**

In a fractional design of factors of three or more levels, the interaction effects are split into different components, where each of these components can be used as an estimator of the interaction effect. For instance, when a  $3^3$  is split into three equal fractions (i.e., a  $3^{3-1}$  design is constructed), the  $A * B$  effect is split into two orthogonal components,  $AB^1$  and  $AB^2$ . The superscript merely indicates the way the factor is fractionalized. To prevent confusion, it is customary to keep the letters in alphabetical order, and the superscript of the first letter is kept to unity (Kempthorne, 1952/1973, p. 295). Because it is defined that  $A^1 B^1 = A^2 B^1$  and  $A^1 B^2 = A^2 B^2$ , the usual notation is respectively  $AB^1$  and  $AB^2$ : The first factor in every effect component has no superscript. Both of these fractionalized effects are sometimes regarded as estimators of the interaction effect  $A \times B$ . In a full design, their sum equals the interaction SS of the  $A \times B$  effect (Winer et al., 1991).

Regardless of the number of levels, a letter alone without a superscript number indicates the first factor. This also holds true for main effects  $A^1$  and  $A^2$ , which indicate one and the same main effect,  $A$ ; hence, it makes no sense to add a number. Sometimes, all superscripts of unity are discarded—for instance,  $AB^1$  is noted as  $AB$ —but this notation is easily confused with an ordinary  $A * B$  effect of a full design. Kirk (1968) places the fractionalized interaction components between parentheses—for instance, using the notation  $(AB)$ .

However, Trial Run uses the superscript notation for other purposes, as is reflected in the manual (SPSS, 1997) on page 78. The manual says that the superscript numbers used in the effect notation point to the respective levels of the different factors: “The effects are labeled with a letter indicating the factor involved, and a number indicating the level of the factor” (p. 78). The term  $A^2$  would then mean the second level of factor  $A$ . This notation is confusing in this context, and it remains unclear how the different fractional interaction effect components can be distinguished. The text continues: “Thus for a three-level factor, the effects are  $A$  and  $A^2 \dots$ ”; normally, only a single effect of  $A$  is considered in the context of fractional design. There is a possibility that the writers of the manual want to indicate different contrasts; however, especially in the case of fractional de-

signs, doing it this way creates insurmountable problems of interpretation.

Regrettably, this is not merely a confusing notation that affects only the documentation; it is a mistake that affects the core of fractional design. The program produces erroneous results, which become clear when we look at the produced tables of confounded effects. For instance, when generating a resolution IV design of four factors each with four levels, the identity generator reports effects such as  $A^2 B^2 C^2 D^1$ , which is identical to  $AB^2 C^2 D^1$  and therefore confusing. Bizarrely, it also reports  $A^1 A^2 B^1 C^1 D^2$ , which is a nonexistent effect in every notation, since a variable does not interact with itself. Especially when the design involves a nonprime number of levels, the results of the identity generator are incorrect. In other cases, the presented results are, at the very least, confusing.

**Data Analysis**

Data obtained by fractional designs will usually be analyzed by ANOVA (e.g., the GLM module in SPSS). For designs that incorporate only two levels [ $2^{(n-m)}$  designs], full analysis is possible with ANOVA, because all effects in these fractional designs have a single degree of freedom, the same as in a regular design. Confounded effects are indicated simply by the first of the confounded effects. In fractional designs with factors of three levels or more, main effects can be estimated by ANOVA, but interaction effects are split into various components. The estimation of these interaction components requires special facilities that are not provided by regular ANOVA.

Trial Run automatically generates GLM statements for analysis of the fractional designs. The GLM statements generated by Trial Run always allow the estimation of main effects. Often, fractional designs offer the possibility of estimating some or all two-way interaction effects. In most cases, the GLM statements generated by Trial Run ignore these estimable interaction effects. It appears that interactions are estimated only when all effect components are confounded with third-order or higher order interaction effects.

A serious problem encountered when using GLM for analysis of fractional designs is that GLM does not make use of the usual notation for fractional designs. For instance, GLM uses only the

**Table A1**  
**Comparison of the Analysis of the Example as Offered by Kirk (1968) and the Analysis Offered by the SPSS Module GLM**

Kirk			SPSS		
Source	SS	df	Source	Type III SS	df
			Corrected Model	236.296	26
			Intercept	1095.704	1
A	4.52	2	A	4.519	2
B	103.19	2	B	103.185	2
C	90.74	2	C	90.741	2
D	4.74	2	D	4.741	2
(AB <sup>2</sup> )	.30	2	A * B	.296	2
(AC <sup>2</sup> )	1.19	2	A * C	1.185	2
(AD)	2.74	2	A * D	2.741	2
(BC <sup>2</sup> )	7.41	2	B * C	7.407	2
(BD)	5.86	2	B * D	5.852	2
(CD)	1.86	2	C * D	1.852	2
Residual	13.75	6	Error	.000	0
Total	236.60	26	Total	1332.000	27

## APPENDIX A (Continued)

notation  $A * B$  for a two-way interaction, and an  $AB^1$  interaction component cannot be distinguished from an  $AB^2$  component.

However, the GLM module of SPSS is very elaborate and offers more possibilities than are used by Trial Run. By specifying the effects that can be estimated with the effect of the lowest order of a series of confounded effects, GLM can produce the correct sum of squares (SS) and mean squares (MS); however, without the appropriate notation, it remains guesswork as to whether  $AB^2$  or  $AB^1$  has been estimated. As an example, we have reanalyzed the  $3^{(4-1)}$  fractional design of Kirk (1968, p. 399) with main effects and the six estimable two-way interactions, as shown in Table A1. Kirk's total of 236.60 is concordant with SPSS Type III SS for the corrected model.

With one exception, all desired effects are produced, albeit without the desired notation. The residual is not estimated. An error SS of 0 is obtained, with  $df = 0$ ; and GLM does not show

the residual effect SS of 13.75 with  $df = 6$  for the remaining (nonestimable) interaction effects. It is customary to add this residual to the error, but we have not found a way to do this with GLM. The disputable SS error does not show up because of the type of SS chosen for the analysis, and other types do not produce the results desired here. More disturbingly, 6  $dfs$  have disappeared from the analysis without comment, thereby making it impossible to test the estimable main effects and interaction effects. It seems that GLM is not very well suited for the analysis of fractional designs. Of course, we have reported these inadequacies to SPSS, Inc., but we feel that users and potential users of Trial Run should be warned as well. Regrettably, we have received no response from SPSS, Inc., other than a request that we resend the remarks as e-mail. As far as we know, Trial Run 1.0 is still sold, and we could not find any information on this Website ([www.SPSS.com](http://www.SPSS.com)) about possible updates.

## APPENDIX B

SPSS 8.0 Set-Up for Comparison of a Full  $2^5$  Design and a Fractional  $2^{(5-1)}$  Design

(A comment is preceded by an asterisk [\*] and concerns the SPSS command following the comment.)

\*Definition of how the data must be read. Variable Run is the case number, A through E are the five variables.

DATA LIST RECORDS 1 /1 run 1-3 A 5-6 B 7-8 C 9-10 D 11-12 E 13-14.

\*Data of the full  $2^5$  design, which is replicated four times, for a total of 128 cases.

BEGIN DATA

```

1 0 0 0 0
2 1 0 0 0
3 0 1 0 0
4 1 1 0 0
5 0 0 1 0
6 1 0 1 0
7 0 1 1 0
8 1 1 1 0
9 0 0 0 1
10 1 0 0 1
11 0 1 0 1
12 1 1 0 1
13 0 0 1 1
14 1 0 1 1
15 0 1 1 1
16 1 1 1 1
17 0 0 0 1
18 1 0 0 1
19 0 1 0 1
20 1 1 0 1
21 0 0 1 1
22 1 0 1 1

```

## APPENDIX B (Continued)

23 0 1 1 0 1  
 24 1 1 1 0 1  
 25 0 0 0 1 1  
 26 1 0 0 1 1  
 27 0 1 0 1 1  
 28 1 1 0 1 1  
 29 0 0 1 1 1  
 30 1 0 1 1 1  
 31 0 1 1 1 1  
 32 1 1 1 1 1

REM Copy these 32 cases 3 more times to get a total of 128 cases.  
 END DATA.

\*Automated simulation with 1,000 iterations.

\*Define SPSS subroutine.

DEFINE !SIM ().

\*Repeat 1,000 times.

!DO !I = 1 !TO 1000.

\*Give each run a label with its number.

SUBTITLE Run !I.

\*Compute a separate counter k.

COMPUTE k=k+1.

\*Show where we are by showing k and ADSD in a frequency table.

FREQUENCIES VARIABLES = k ADSD

/FORMAT=NOTABLE

/STATISTICS=MEAN

/ORDER ANALYSIS .

\*Recode variables A through E to get values created by an orthogonal design: Value 0 becomes -1.

RECODE A B C D E (0 = -1).

\*Computation of outcome variable Y for simulation 1.

COMPUTE y = A + B + C + D + E + normal (ADSD).

\*Computation of outcome variable Y for simulation 2; for use, remove the asterisk of the following line.

COMPUTE y = A + B + C + D + A\*B\*C + normal (ADSD).

\*Recode value -1 back to 0 for variables A through E, so that SPSS MANOVA can handle the design.

RECODE A B C D E (-1=0).

\*SPSS command to make the next selection temporary instead of permanent.

TEMPORARY.

\*Select the full design for cases 1 through 64.

SELECT IF (\$Casenum <= 64).

\*Calculate MANOVA for full design for all main effects.

MANOVA Y BY A (0,1) B (0,1) C (0,1) D (0,1) E (0,1) /DESIGN A B C D E .

\*SPSS command to make the next selection temporary instead of permanent.

TEMPORARY.

\*Select four times 16 cells out of 32 for the fractional  $2^{5-1}$  design, by selecting the appropriate runs.



---

**APPENDIX B (Continued)**

---

```
SELECT IF (run=1 or run=4 or run=6 or run=7 or run=9
or run=12 or run=14 or run=15 or run=18
or run=19 or run=21 or run=24 or run=26 or run=27
or run=29 or run=32) .
*Calculate MANOVA for fractional design.
MANOVA Y BY A (0,1) B (0,1) C (0,1) D (0,1) E (0,1) /DESIGN A B C D E .
!DOEND.
*End of subroutine SIM.
!ENDDFINE.

*Give the iteration an appropriate label.
TITLE 'SS within 1'.
*Reset k to 0.
COMPUTE k =0.
*Calculate the added standard deviation ADSD of the unsystematic measurement error.
COMPUTE ADSD=SQRT(1).
*Call subroutine SIM.
!SIM.
TITLE 'SS within 4'.
COMPUTE k =0.
*Calculate another value for the added standard deviation ADSD of the unsystematic measurement
error.
COMPUTE ADSD=SQRT(4).
!SIM.
*Repeat previous four lines with other ADSD values.
```

---