On a Heuristic Analysis of Highly Fractionated 2^n Factorial Experiments

by C. Auer[1] and J. Kunert[1]

[1] Fachbereich Statistik, Universität Dortmund, D 44221 Dortmund, Germany

Abstract

The paper deals with a method for the analysis of highly fractionated factorial designs proposed by Raghavarao and Altan (2003). We show that the method will find "active" factors with almost any set of random numbers. Once that an alias set is found active, Raghavarao and Altan (2003) claim that their method can resolve the alias structure of the design and identify which of several confounded effects is active. We show that their method cannot do that. The error in Raghavarao and Altan's (2003) arguments lies in the fact that they treat a set of highly dependent (sometimes even identical) F-statistics as if they were independent.

Key words

Fractional factorial designs, half-normal plot, heuristic arguments, active effects, alias set

Short Title

A Heuristic Analysis of Fractional Factorial Designs

Correspondence to

J. Kunert, Fachbereich Statistik, Universität Dortmund, D 44221 Dortmund, Germany
e-mail: joachim.kunert@udo.edu

1. Introduction

In a seminal paper, Daniel (1959) introduced the half-normal plot for the analysis of unreplicated fractional factorial designs. Since the appearance of Daniel (1959), a lot of work has been done to generalize and/or formalize Daniel's (1959) method. The literature on the analysis of such designs is abundant. Some papers propose alternative estimates for the variance, like e.g. Lenth (1989) and Dong (1993). Others have tried to derive the statistical properties of existing methods and get critical values based on the distribution of the estimates (like e.g. Zahn, 1975a, or Kunert, 1997). Another group of publications compare several proposed methods, mainly with the help of simulations, like e.g. Zahn (1975b) or Haaland and O'Connell (1995).

An interesting group of papers has tried to develop methods of analysis that are based on some formal statistical principle, like e.g. multiple testing (e.g. Voss, 1999) or a Bayesian approach (e.g. Box and Meyer, 1986).

Finally, there is a group of papers that propose methods, the justification of which is purely based on intuition. Although this may lead to sensible procedures, there also is a danger. A statistician has the duty to separate what part of an apparent structure in a given data set is due to a true structure in the data and what part is due to noise. To do this we usually have a model in mind, and we have to have some knowledge what our methods are doing when this model holds. This knowledge can be based on a mathematical theory, or, if the situation gets too complicated, it can be based on simulations.

Simple intuition can lead to highly questionable methods. If the intuition is not backed up by proper thinking, the resulting procedures may become nonsensical. They may lead to conclusions that are entirely arbitrary. An example for such a procedure that leads to arbitrary results is the one proposed by Raghavarao and Altan (2003).

In what follows, we try to point out why the method put forward by Raghavarao and Altan (2003) should not be used. We demonstrate that this method will lead to conclusions that are arbitrary and not backed by the data. One problem is that it will find "active" factors with almost any set of random numbers. The other problem is that even in cases where there are truly active factors, it tries to resolve the confounding structure of the fractional factorial design in a way that is not reasonable.

2. Method of Raghavarao and Altan

Raghavarao and Altan (2003) propose a method to identify active effects in unreplicated 2^n fractional factorial experiments. They describe their method with the help of an example. For this example they use a 2^{5-2} fractional factorial. We begin with a short review of their method and the example.

The confounding pattern of the 2^{5-2} design used is given by the relation ABC=ADE=BCDE. This means that e.g. the main effect of treatment A is confounded with the interaction BC and the interaction DE. On the other hand, the main effect B is confounded with AC, and no other two factor interaction. For an explanation of how the confounding structure can be derived, see e.g. Box, Hunter, Hunter (1978, Section 12).

The 2^{5-2} design has 8 observations. Raghavarao and Altan (2003) in their method propose to consider each of the 10 possible models allowing for two factors and their interaction and neglecting the other factors and interactions. This leaves f = 4 degrees of freedom for each of these models. For each of these models they calculate the three corresponding F-statistics to estimate the three terms in the model, see Table 1. Note that each of the F-statistics in Table 1 has 1 degree of freedom for the numerator and f = 4 degrees of freedom for the denominator. Hence, under the null-hypothesis of no active factors, each F-statistic has the same expectation $f/(f–2)$. Unfortunately, for f = 4 the variance does not exist. In a larger example, where f is greater than 4, however, the variance of the F-statistic exists. Under the null-hypothesis the F-statistics then all have the same variance.

Table 1[*]. All possible two factor models and corresponding F-statistics for the 2^{5-2}-design

Model Number	Terms in Model	F-Statistics
1	A, B, AB	F_A^1, F_B^1, F_{AB}^1
2	A, C, AC	F_A^2, F_C^2, F_{AC}^2
3	A, D, AD	F_A^3, F_D^3, F_{AD}^3
4	A, E, AE	F_A^4, F_E^4, F_{AE}^4
5	B, C, BC	F_B^5, F_C^5, F_{BC}^5
6	B, D, BD	F_B^6, F_D^6, F_{BD}^6
7	B, E, BE	F_B^7, F_E^7, F_{BE}^7
8	C, D, CD	F_C^8, F_D^8, F_{CD}^8
9	C, E, CE	F_C^9, F_E^9, F_{CE}^9
10	D, E, DE	F_D^{10}, F_E^{10}, F_{DE}^{10}

[*] Table 3 from Raghavarao and Altan (2003)

To identify the active factors, Raghavarao and Altan (2003) calculate an adjusted F-Statistic $adj(\overline{F})$ for each alias set. This is derived as follows.

Assume that e.g. factor A was active. Then we could expect that the F-statistics F_A^1, F_A^2, F_A^3, F_A^4, but also F_{BC}^5 and F_{DE}^{10} should be large. This is due to the fact that the main effect A and the interactions BC and DE are in the same alias set.

Therefore, it looks sensible to take the average $\overline{F}$ of all the F-statistics corresponding to an alias set. From this average the expected value for such an F-statistic, i.e. *f*/(*f*-2), is subtracted. Then this difference is standardized by multiplication with the square root of the number of F-statistics used to derive it. Note that under the null-hypothesis and if all the F-statistics were independent, this would lead to an $adj(\overline{F})$ with expectation 0 and (if the variance exists) with a variance that does not depend on the number of F-statistics used to derive it. If for an alias set, all the F-statistics are large, then the adjusted F-statistic will be larger than for an alias set where just one of the F-statistics is large. Therefore a large $adj(\overline{F})$ looks more convincing than just one large F-statistic. This argumentation sounds intuitively appealing - we will, however, see below that there is a fundamental error in the argumentation.

For the 2^{5-2} design considered in the example, the 7 alias sets, the formulae for the average of the F-statistics and the adjusted F-statistics are given in Table 2.

Table 2*. Alias sets with main effects and two factor interactions along with $\overline{F}$ and adjusted $\overline{F}$ values

Alias Set	Main Effects and 2 Factor Interactions	$\overline{F}$	$adj(\overline{F})$
I	A = BC = DE	$\overline{F}_I = \left(F_A^1 + F_A^2 + F_A^3 + F_A^4 + F_{BC}^5 + F_{DE}^{10}\right)/6$	$\left(\overline{F}_I - 2\right)\sqrt{6}$
II	B = AC	$\overline{F}_{II} = \left(F_B^1 + F_B^5 + F_B^6 + F_B^7 + F_{AC}^2\right)/5$	$\left(\overline{F}_{II} - 2\right)\sqrt{5}$
III	C = AB	$\overline{F}_{III} = \left(F_C^2 + F_C^5 + F_C^8 + F_C^9 + F_{AB}^1\right)/5$	$\left(\overline{F}_{III} - 2\right)\sqrt{5}$
IV	D = AE	$\overline{F}_{IV} = \left(F_D^3 + F_D^6 + F_D^8 + F_D^{10} + F_{AE}^4\right)/5$	$\left(\overline{F}_{IV} - 2\right)\sqrt{5}$
V	E = AD	$\overline{F}_V = \left(F_E^4 + F_E^7 + F_E^9 + F_E^{10} + F_{AD}^3\right)/5$	$\left(\overline{F}_V - 2\right)\sqrt{5}$
VI	BD = CE	$\overline{F}_{VI} = \left(F_{BD}^6 + F_{CE}^9\right)/2$	$\left(\overline{F}_{VI} - 2\right)\sqrt{2}$
VII	BE = CD	$\overline{F}_{VII} = \left(F_{BE}^7 + F_{CD}^8\right)/2$	$\left(\overline{F}_{VII} - 2\right)\sqrt{2}$

* Table 4 from Raghavarao and Altan, corrected for misprints

The so calculated adjusted F-statistics are now plotted in descending order. One expects that the adjusted F-statistic of an alias set containing an active factor or interaction will be large compared to the other F-statistics. Raghavarao and Altan (2003) claim that if there are active factors then there will be an "elbow shape formation". Here elbow shape means that the adjusted F-statistics roughly speaking are arranged along two lines, one with a greater slope with the large F-statistics (presumably the ones corresponding to active factors), one with a smaller slope, with the small F-statistics.

If the plot has an elbow shape formation, the method declares those F-statistics to be significant, which lie along the line with the steeper slope. Again, we will point out in the next section why this approach is highly questionable.

When an alias set is assumed to be active, the method computes adjusted F-statistics for the main effects and 2 factor interactions in that set. These adjusted F-statistics are calculated analogous to the ones for the alias sets, but we take the average only over those F-statistics which test the factor that we are interested in, not the ones which test the alialised factors. For instance, to calculate the adjusted F-statistic for the factor A we would take the mean of the four statistics F_A^1, F_A^2, F_A^3, F_A^4, subtract 2 and multiply by the square root of 4, i.e. 2. The adjusted F-statistic for the interaction BC would be F_{BC}^5 minus 2 (multiplied by 1). For each alias set that was declared active, the method then chooses the factor or interaction with the largest adjusted F-statistic, and claims that this is the one that is active and therefore responsible for the size of the contrast.

3. Discussion of the method

The method of Raghavarao and Altan (2003) is developed using heuristic arguments only. Although there are many good heuristic methods in the literature, intuition sometimes can lead to questionable results. In our experience it is important to examine the mathematical properties of a method or at least do some simulation runs to get a feeling of what a method really does.

For these reasons we found it interesting to have a closer look at the proposed method.

In their 2^{5-2} design, Raghavarao and Altan (2003) consider all possible sub-models with two factors and the corresponding interaction. For each of the factors of each of these models they calculate the F-statistic. This procedure will lead to valid F-tests whenever all contrasts not in the model are not active. There are similar approaches in the literature: e.g. Berk and Picard (1991) as well as Al-Shiha and Yang (1999) would calculate similar F-statistics and compare the largest of those to an appropriate critical value.

To identify which F-statistic indicates significance, Raghavarao and Altan (2003) go another way. They calculate their adjusted F-Statistic. When they do that, however, they do not realise that the F-statistics used for the calculation of an adjusted F-statistic are not independent at all.

In the 2^{5-2} design, there are exactly 7 independent contrasts. If we assume a model with two main effects and their interaction, then there are three contrasts used to estimate the effects in that model. The other four contrasts can be used to estimate the variance. Hence, each of the F-statistics in Table 1 is the quotient of the square of one of these contrasts, divided by the mean of the square of four others. The confounding structure in Table 2 helps to see which contrast plays which role in each of the models. For instance model 1 in Table 1 uses the factors A and B and the interaction AB. From Table 2 we see that the contrasts corresponding to the alias sets I, II and III are used to estimate the parameters of the model, while the alias sets IV, V, VI and VII are used to calculate the variance. Hence, F_A^1 uses the square of the contrast corresponding to alias set I as numerator and the mean square of the alias sets IV, V, VI and VII as denominator. However, model 2 from Table 1 uses the same three contrasts to estimate the parameters, since AC is confounded with B and C is confounded with AB. Therefore, it also uses the same four contrasts to estimate the variance. Consequently, $F_A^2 = F_A^1$.

In Table 3 we list the contrasts used for numerators and denominators of the 30 F-statistics in Table 1. We find that quite a few of them are identical.

Table 3

Contrasts used to calculate the F-statistics of the two factor models proposed by Raghavarao and Altan (2003)

Model No	Terms in the model	F-statistics with corresponding contrasts in the numerator	contrasts used for the denominator
1	A, B, AB	F_A^1 with I, F_B^1 with II, F_{AB}^1 with III	IV, V, VI, VII
2	A, C, AC	F_A^2 with I, F_C^2 with III, F_{AC}^2 with II	IV, V, VI, VII
3	A, D, AD	F_A^3 with I, F_D^3 with IV, F_{AD}^3 with V	II, III, VI, VII
4	A, E, AE	F_A^4 with I, F_E^4 with V, F_{AE}^4 with IV	II, III, VI, VII
5	B, C, BC	F_B^5 with II, F_C^5 with III, F_{BC}^5 with I	IV, V, VI, VII
6	B, D, BD	F_B^6 with II, F_D^6 with IV, F_{BD}^6 with VI	I, III, V, VII
7	B, E, BE	F_B^7 with II, F_E^7 with V, F_{BE}^7 with VII	I, III, IV, VI
8	C, D, CD	F_C^8 with III, F_D^8 with IV, F_{CD}^8 with VII	I, II, V, VI
9	C, E, CE	F_C^9 with III, F_E^9 with V, F_{CE}^9 with VI	I, II, IV, VII
10	D, E, DE	F_D^{10} with IV, F_E^{10} with V, F_{DE}^{10} with I	II, III, VI, VII

It follows from Table 3 that the F-statistics

F_A^1, F_A^2 and F_{BC}^5,

F_B^1, F_{AC}^2 and F_B^5,

F_{AB}^1, F_C^2 and F_C^5,

F_A^3, F_A^4 and F_{DE}^{10},

F_D^3, F_{AE}^4 and F_D^{10},

F_{AD}^3, F_E^4 and F_E^{10}

are respectively the same, whatever observations we might have. Hence there are only 18 different F-statistics possible and e.g. the adjusted F-statistics for the alias set I, in reality is based on the sum of just two F-statistics, not 6. The adjusted F-statistic for the alias set II is based on the sum of three F-statistics, where one is weighted three times higher than the others. This shows that the argumentation in Raghavarao and Altan (2003) is untenable when claiming that the adjusted F-statistics for the alias set I should be multiplied by $\sqrt{6}$ and that for the alias set II should be multiplied by $\sqrt{5}$.

Things are even worse than that. Note that even those F-statistics for a given alias set which are not identical are clearly not independent: they all use the same numerator. Hence, they will not be uncorrelated, but highly positively correlated (whenever the correlation exists). In fact, if the correlation exists, it will be close to 1. This implies that the means $\bar{F}$ of the F-statistics corresponding to a given alias set, will have a variance that almost is the same as the variance of the single F-statistics (provided these variances exist). For this it follows that the adjusted F-statistics will not have identical variances, but that the variance of an adjusted F-statistic will equal k times the variance of a single F-statistic, where k is the number of F-statistics used to calculate it. In the 2^{5-2} design considered to explain the method, this implies that compared to the alias sets II, III, IV and V, the adjusted F-statistic for alias set I will be pushed away from 0, while the adjusted F-statistics for alias sets VI and VII will be pushed towards 0, whatever observations we might have.

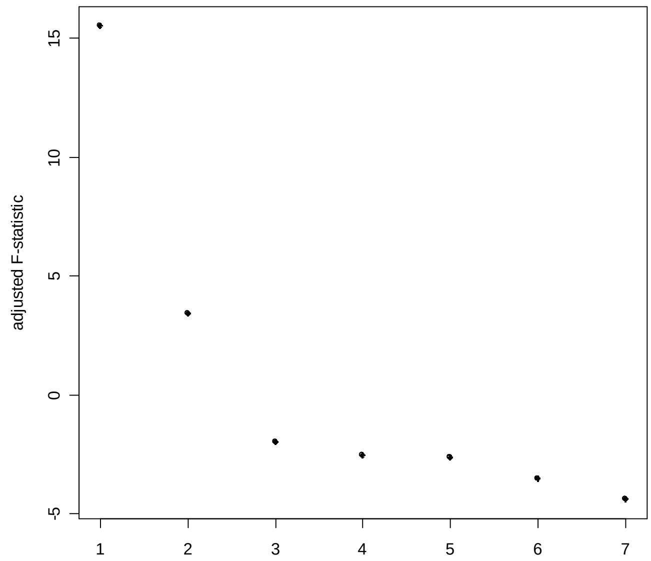
Now assume we observe a set of random numbers. Note that this may happen in practice, even with industrial experiments. In our practical experience, we have seen more than one expensive experiment, where the measuring device did not work like we had expected, or where there were such strong uncontrolled influences on the observations that any possible effect of the factors could not be seen. In some of these cases, the problems were only realized in retrospect after the experiment had been done, when a global test could not reject the null-hypothesis that none of the factors had any effect.

We therefore think that any reasonable method should allow for such a test and should make sure that the probability to find "active" effects in a set of random numbers is controlled.

With the method of Raghavarao and Altan (2003), however, we will almost surely identify "active" factors. To demonstrate this we have done a simulation study where we used a formalized version of Raghavarao and Altan's method. In this simulation study, we produced 10,000 sets of 8 independent and identically normally distributed observations, which we analysed as if they were the response of the 2^{5-2} design. We analysed each of the data sets with the proposed method. Figure 1 displays the adjusted F-statistics for one of these random samples. We think that this is one instance where the method would identify an elbow-structure.

To produce Figure 1, we sorted the adjusted F-statistics according to size, and plotted them such that the x-coordinate of the i-th point was i, the y-coordinate of the i-th point was the i-th largest adjusted F-statistic. If in Figure 1 we would fit a straight line to the smallest four points, then the largest two points would be clearly above the line. We therefore might fit another line to the three largest points. The two lines would meet approximately at the third point. We think that this is what Raghavarao and Altan (2003) termed elbow shape.

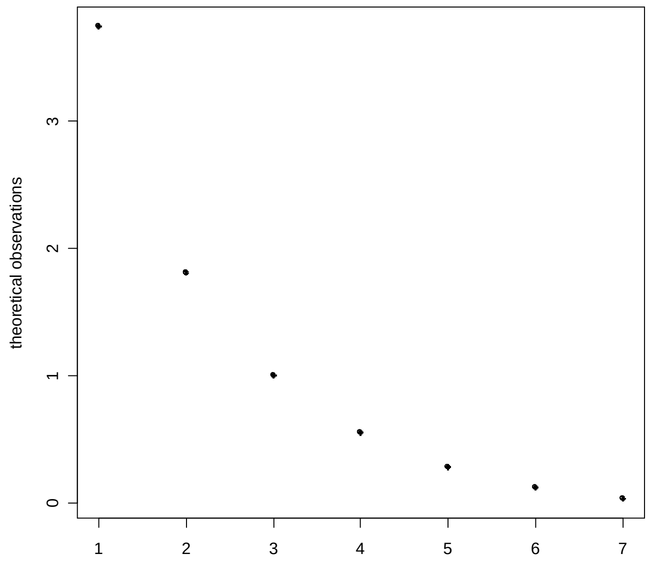
Figure 1: An example of adjusted F-statistics with an elbow structure



To formally identify whether there was an elbow shape in a given sample, we always fitted a straight line to the smaller 4 of the 7 adjusted F-statistics. If the largest of the F-statistics was more than 3 times larger than the value of the straight line at $x = 1$, we concluded that there was an elbow shape. This happened in more than 85% of all random data sets in our study. Hence, the method clearly does not help to see whether there are truly active factors or not. An easier way to analyse (with almost the same statistical properties) would be to simply declare the two or three largest of the contrasts as active.

To explain why such a high fraction of plots with an elbow shape had to be expected, we might consider the theoretical distribution of an F-statistic with 1 and 4 degrees of freedom. If each point of the empirical distribution function lay on the theoretical distribution function, then the *i*-th largest observation would equal $F_{1,4}^{-1}\left(\frac{8-i}{7}\right)$, where $F_{1,4}^{-1}$ is the inverse of the F-distribution with 1 and 4 degrees of freedom. Since this would mean that the largest observation had to be ∞, it is customary to plot $F_{1,4}^{-1}\left(\frac{8-i}{8}\right)$ against *i*. This produces Figure 2 which clearly shows the elbow shape.

Figure 2: Theoretical observations from a F distribution with 1 and 4 degrees of freedom



However, what makes Raghavarao and Altan's (2003) method particularly questionable is their claim that they can resolve the alias structure of the design.

Raghavarao and Altan (2003) claim, that their method can identify which of the effects in an alias set is active. This certainly would be nice, but of course it is not possible. To give an example, we considered the 2^{5-2} design and the situation where the factors A and B both have a main effect but there is no interaction. All other factors had no effect.

To get data with an effect of A and of B, we simulated 8 independent normally distributed observations, each with expectation 0 and variance 8. To those observations corresponding to runs with factor A at +1, we added 10, from those observations corresponding to runs with factor A at –1 we subtracted the same number. Further, we added 3 to all observations with B at +1 and subtracted 3 from all observations with B at –1. It is easy to see that then the squared estimate for the half-effect of A has expectation 101, while the squared estimate for B has expectation 10. The squared estimates for the factors in other alias sets have expectation 1.

We then calculated the adjusted F-statistics for the factors and interactions within alias sets I and II. Note that the method would correctly decide that A and B are active only if both

$$adj(\overline{F}_A) \geq \max\left\{adj(\overline{F}_{BC}), adj(\overline{F}_{DE})\right\},$$

and

$$adj(\overline{F}_B) \geq adj(\overline{F}_{AC}).$$

We did this simulation 10,000 times and counted the number of simulation runs where the inequalities were true. It turned out that the first inequality was true in all runs, but the second inequality only held in 2 out of 10,000 runs. So the estimated probability is 0.9998 that the method falsely concludes that AC is active and B is not active. A very poor performance, indeed.

Table 4. Expected Sum of Squares due to the alias sets

Alias Set	Main Effects and 2 Factor Interactions	Sum of Squares
I	A = BC = DE	$SSI = x$
II	B = AC	$SSII = 10$
III	C = AB	$SSIII = 1$
IV	D = AE	$SSIV = 1$
V	E = AD	$SSV = 1$
VI	D = CE	$SSVI = 1$
VII	BE = CD	$SSVII = 1$

To get some more insight in what is going on, we let the effect of A vary and consider the simplified situation where each squared estimate is equal to its expectation. Assume A has the effect $\sqrt{x-1}$ (instead of 10) and B has the effect 3 as before. Table 4 lists the expected sum of squares of the 7 contrasts for this situation.

Table 5. F-statistics for the data in Table 4

Model Number	Terms in Model	F Statistics
1	A, B, AB	$F_A^1 = x$, $F_B^1 = 10$, $F_{AB}^1 = 1$
2	A, C, AC	$F_A^2 = x$, $F_C^2 = 1$, $F_{AC}^2 = 10$
3	A, D, AD	$F_A^3 = \frac{4x}{13}$, $F_D^3 = \frac{4}{13}$, $F_{AD}^3 = \frac{4}{13}$
4	A, E, AE	$F_A^4 = \frac{4x}{13}$, $F_E^4 = \frac{4}{13}$, $F_{AE}^4 = \frac{4}{13}$
5	B, C, BC	$F_B^5 = 10$, $F_C^5 = 1$, $F_{BC}^5 = x$
6	B, D, BD	$F_B^6 = \frac{40}{x+3}$, $F_D^6 = \frac{4}{x+3}$, $F_{BD}^6 = \frac{4}{x+3}$
7	B, E, BE	$F_B^7 = \frac{40}{x+3}$, $F_E^7 = \frac{4}{x+3}$, $F_{BE}^7 = \frac{4}{x+3}$
8	C, D, CD	$F_C^8 = \frac{4}{x+12}$, $F_D^8 = \frac{4}{x+12}$, $F_{CD}^8 = \frac{4}{x+12}$
9	C, E, CE	$F_C^9 = \frac{4}{x+12}$, $F_E^9 = \frac{4}{x+12}$, $F_{CE}^9 = \frac{4}{x+12}$
10	D, E, DE	$F_D^{10} = \frac{4}{13}$, $F_E^{10} = \frac{4}{13}$, $F_{DE}^{10} = \frac{4x}{13}$

According to Raghavaraos and Altan's (2003) method we start by computing the F-statistics for each of the 10 models in Table 1. For example, for model 3 we get as the denominator SSE of the F-statistic

$$SSE = \frac{1}{4}(SSII + SSIII + SSVI + SSVII) = \frac{1}{4}(10+1+1+1) = \frac{13}{4}.$$

Thus we get the following F-statistics

$$F_A^3 = \frac{SSI}{SSE} = \frac{4x}{13},\ F_D^3 = \frac{SSIV}{SSE} = \frac{4}{13} \text{ and } F_{AD}^3 = \frac{SSV}{SSE} = \frac{4}{13}.$$

The F-statistics for each of the 10 models are given in Table 5.

From these F-statistics the $\overline{F}$ -values and the adjusted F-statistics can be derived, see Table 6.

Table 6. The adjusted F-statistics

Alias Set	Main Effects and 2 Factor Interactions	$\overline{F}$	$adj(\overline{F})$
I	A = BC = DE	$\frac{17}{26}x$	$\left(\frac{17}{26}x - 2\right)\sqrt{6}$
II	B = AC	$6 + \frac{16}{x+3}$	$\left(4 + \frac{16}{x+3}\right)\sqrt{5}$
III	C = AB	$\frac{3}{5} + \frac{8}{5(x+12)}$	$\left(-\frac{2}{5} + \frac{8}{5(x+12)}\right)\sqrt{5}$
IV	D = AE	$\frac{12}{65} + \frac{4}{5(x+3)} + \frac{4}{5(x+12)}$	$\left(-\frac{53}{65} + \frac{4}{5(x+3)} + \frac{4}{5(x+12)}\right)\sqrt{5}$
V	E = AD	$\frac{12}{65} + \frac{4}{5(x+3)} + \frac{4}{5(x+12)}$	$\left(-\frac{53}{65} + \frac{4}{5(x+3)} + \frac{4}{5(x+12)}\right)\sqrt{5}$
VI	BD = CE	$\frac{2}{x+3} + \frac{2}{x+12}$	$\left(\frac{2}{x+3} + \frac{2}{x+12} - 2\right)\sqrt{2}$
VII	BE = CD	$\frac{2}{x+3} + \frac{2}{x+12}$	$\left(\frac{2}{x+3} + \frac{2}{x+12} - 2\right)\sqrt{2}$

To decide whether B or AC is active, the method computes and compares the corresponding adjusted F-statistics as follows

$$adj(\overline{F}_B) = \left\{\frac{F_B^1 + F_B^5 + F_B^6 + F_B^7}{4} - 2\right\}\sqrt{4} = \left\{\frac{10 + 10 + \frac{40}{x+3} + \frac{40}{x+3}}{4} - 2\right\}2 = 6 + \frac{40}{x+3}$$

$$adj(\overline{F}_{AC}) = F_{AC}^2 - 2 = 10 - 2 = 8.$$

We find that

$$adj(\overline{F}_B) \geq adj(\overline{F}_{AC}) \Leftrightarrow 6 + \frac{40}{x+3} \geq 8 \Leftrightarrow x \leq 17.$$

Hence, Raghavarao and Altan (2003) will falsely conclude that the interaction AC is active whenever $x > 17$. They will only see that B is active if $x \leq 17$.

To resolve the alias set A=BC=DE one has to compute the adjusted F-statistics for A, BC and DE, which give $adj\left(\overline{F}_A\right)=\frac{17}{3}x-4$, $adj\left(\overline{F}_{BC}\right)=x-2$ and $adj\left(\overline{F}_{DE}\right)=\frac{4}{13}x-2$.

Comparing these values, it is easy to see that again the decision only depends on the size of *x*:

we get a decision for $\begin{cases} A, & \text{if } x \geq 6.5, \\ BC & x < 6.5. \end{cases}$

So we get the correct conclusion that A and B are active only if the effect $\sqrt{x-1}$ of A is between 2.1 and 4. Note that the situation in our simulation corresponds to the case *x*=101.

To understand what is going on here, we should realize that e.g. the interaction AC gets tested only in models where factor A is present. Hence, the F-statistics for AC will never have A in the denominator. The main effect of B, however, gets tested with some F-statistics that do have A in the denominator (and therefore are rather small). This is why, for a large effect of A, the method thinks that AC and not B is active. Obviously, when there are two active main effects, the method tries to explain the smaller of these as an interaction of the larger factor with some other factor.

We therefore can see that the complicated method of resolving the alias structure breaks down to the following simple algorithm: If there is an active alias group containing a main effect M, say, and some two factor interactions, then check whether one of the interactions contains a factor L, say, whose main effect is in another active alias set. If the estimate for the effect of L is "much" larger than M, then declare the interaction as active, if L is about the same size or smaller than M, then declare M itself as active.

In their paper, Raghavarao and Altan (2003) do a reanalysis of an example from Box and Meyer (1993). Box and Meyer (1993) identified the main effects of factors B, D and E as active, where the estimated effect of B was largest of the three. We were not surprised to read that Raghavarao and Altan (2003) claim that their method identified B and the two factor interactions BD and DE instead.

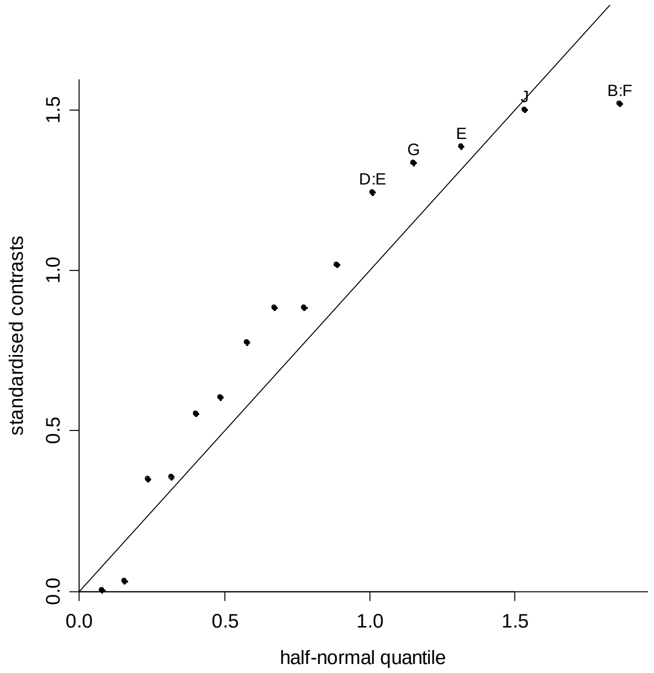
Raghavarao and Altan (2003) motivated their method with the data from a 2^{9-5} design. For the data, see Table 1 in Raghavarao and Altan (2003). They had used the Bayesian analysis of Box and Meyer (1993) and a method implemented in SAS to analyse these data. (They write in their paper that they used the macro ADXCODE, but this macro cannot be used not for the analysis of data. Presumably they used the macro ADXFFA.) They found that both standard methods did not find any active factors. So they came to the conclusion that these methods are not able to detect active contrasts and decided to develop a new method.

Applying their method, they claim that the factors and interactions BF, J, E, G and DE should be active.

For data like these, we would start the analysis with a half normal plot. Note that SAS, in their explanation of the ADXFFA-macro, recommend to use a normal probability plot, see Box, Hunter and Hunter (1978, section 10.9). We prefer the half-normal plot for various reasons, but this is a

matter of taste. In this example both the normal plot and the half normal plot would lead to very similar conclusions.
When we did the half normal plot for Raghavarao and Altan's data, we got the graph in Figure 3.

Figure 3. Half-Normal-Plot for the data in Raghavarao and Altan's Table 1



This half-normal plot looks exactly like what we would expect from a set of random numbers. Whenever we have seen a half-normal plot like that in our consulting work, we could verify in discussions with the experimenters that we had the following situation: There were no active factors in this experiment, the random error was too large compared to any possible signal.

4. Summary

We think that the method of Raghavarao and Altan (2003) is fundamentally wrong and should not be used. The heuristic arguments used to motivate the method cannot work. The fact that they cannot work is related to a lesson that we try to teach to our students in their first year:
While the mean over several numbers intuitively seems more convincing than a single number, this only is true if the numbers were derived in separate, preferably independent, measurements. Writing down the same number 5 times and taking the average, is by no means more convincing than just one single number.

Acknowledgement
Research for this paper was supported by the Deutsche Forschungsgemeinschaft (SFB 475 - Komplexitätsreduktion in multivariaten Datenstrukturen).

References

Al-Shiha AA, Yang SS (1999) A multistage procedure for analysing unreplicated factorial experiments. Biometrical Journal 41:659-670

Berk KN Picard RR (1991) Significance tests for saturated orthogonal arrays. Journal of Quality Technology 23:79-89

Box GEP, Hunter WG, Hunter JS (1978) Statistics for experimenters. John Wiley & Sons, New York, NY

Box GEP, Meyer RD (1986) An analysis for unreplicated fractional factorials. Technometrics 28:11-18

Box GEP, Meyer RD (1993) Finding the active factors in fractionated screening experiments. Journal of Quality Technology 25:94-105

Daniel C (1959) Use of half-normal plots in interpreting factorial two-level experiments. Technometrics 1:311-341

Dong F (1993) On the identification of active contrasts in unreplicated fractional factorials. Statistica Sinica 3:209-217

Haaland PD, O'Connell MA (1995) Inference for effect-saturated fractional factorials. Technometrics 37:82-93

Kunert J (1997) On the use of the factor-sparsity assumption to get an estimate of the variance in saturated designs. Technometrics 39:81-90

Lenth RV (1989) Quick and easy analysis of unreplicated factorials. Technometrics 31:469-473

Raghavarao D, Altan S (2003) A heuristic analysis of highly fractionated 2^n factorial experiments. Metrika 58: 185-191

Voss DT (1999) Analysis of orthogonal saturated designs. Journal of Statistical Planning and Inference 78:111-130

Zahn DA (1975a) Modifications of and revised critical values for the half-normal plot. Technometrics 17:189-200

Zahn DA (1975b) An empirical study of the half-normal plot. Technometrics 17:201-211