This is an Accepted Manuscript of an article published by Taylor & Francis in The American Statistician on March 5, 2014, available online at: http://www.tandfonline.com/10.1080/00031305.2014.896829 reprinted by permission of the American Statistical Association, www.amstat.org

Mosaic plots are useful for visualizing low order projections of factorial designs

By Ulrike Grömping

Beuth University of Applied Sciences, Berlin, Germany

Abstract

Factorial experiments are widely used in industrial experimentation and other fields. Whenever a factorial experiment is not designed as a full factorial but as a – regular or non-regular – fraction thereof, choice between competing designs and interpretation of experimental results should take into consideration how the experimental plan will confound experimental effects. This article proposes mosaic plots of low order projections of factorial designs for visualizing confounding of low order effects. Mosaic plots are particularly useful for design and analysis of orthogonal main effects plans. The R code for creation of the plots in this article is available in the supplemental material online.

Key words: orthogonal arrays, orthogonal main effects plans, confounding, strength, resolution

1 INTRODUCTION

Factorial experiments can be used for simultaneously investigating the effects of several experimental factors on a response variable of interest. Each setting of a factor is called a factor level. For gaining as complete information as possible, one would ideally experiment with all factor level combinations: In such full factorial experiments, all possible combinations of the factor levels of k factors are run the same number of times (say, r times); for example, a full factorial in k=4 3-level factors would have 81r experimental runs. A full factorial experiment is completely balanced, and one can estimate the main effects of each factor, all interaction effects between pairs of factors (interaction effects of degree 2) up to all interaction effects of degree k independently from each other. The complete balance even implies that estimates for ~~all estimable contrasts of lower order~~ modeled effects cannot incur bias from wrongly omitting important effects, e.g. higher order interactions, in other factors.

If the number of experimental factors is moderate or large, or some factors have many levels, full factorial experiments become excessively large and are often not feasible, even when resorting to unreplicated designs (i.e. r=1). In such cases, experiments with fewer runs have to be conducted, which leads to confounding (aliasing) among effects. This article considers factorial designs based on orthogonal arrays, for which the concept of strength describes the severity of confounding (see Section 2). The strength of a design is limited by the lowest order of confounding that is attained in the design; however, designs of the same strength can exhibit very different extents and severities of such lowest-order confounding. Mosaic plots (see Section 3) are proposed as a tool for visualizing the ~~degree of~~ lowest-order (=worst case) confounding that is inherent in an experimental design. In particular, these plots are useful for providing a warning against bias from 2-factor interactions in orthogonal main effects plans (e.g. Addelman 1962; Gupta et al. 1982; Suen and Kuhfeld 2005), many of which are available from published design catalogues (e.g. Appendix 8C of Wu and Hamada 2009); practitioners use such catalogued arrays for designing experiments, often without any knowledge about their inherent confounding structure.

The considerations in this article are most useful for designs with qualitative factors, since mosaic plots do not reflect the geometric properties relevant for designs in quantitative factors. Nevertheless, mosaic plots can also be used to visualize confounding severity for designs with quantitative factors (see also Section 4.2). To the author's knowledge, mosaic plots have not been considered elsewhere as tools for diagnosing experimental designs.

The next section introduces orthogonal arrays and the terminology around their confounding properties. Section 3 introduces mosaic plots (Hartigan and Kleiner 1981 and 1984) in general, Section 4 describes their specific properties for strength s designs. Section 5 discusses the use of mosaic plots in experimental practice. The final section briefly summarizes key uses and limitations of mosaic plots and relates back to research on assessing factorial designs. All code for the analyses presented or discussed in this paper is provided in the online supplementary material.

2 ORTHOGONAL ARRAYS AND FACTORIAL DESIGNS

An orthogonal array is a matrix with *n* rows and *k* columns; the entries of the *i*-th column consist of l_i levels (here denoted with 1, …, l_i) each of which occurs the same number of times; furthermore, the term "orthogonal" implies that, in addition, each pair of levels of the *i*-th and *j*-th column ($i \neq j$) occurs the same number of times.

Table 1 shows an example of a particularly popular orthogonal array, the well-known Taguchi L18, which can be used for examining up to one 2-level factor and up to seven 3-level factors in 18 runs (see e.g. NIST/SEMATECH 2012, Section 5.3.3.10).

Table 1: The Taguchi L18 (columns show the 8 factors, rows the 18 experimental runs)

	Factor							
Run	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

The array of Table 1 is widely known among engineering quality practitioners and is often used as the basis for designing experiments, by selecting as many columns as needed and ignoring the other columns. For example, Dhole, Naik and Prabharwalkar (2012) report an experiment on optimizing a milling process for EN33 steel, in which they assigned the 3-level factors A="cutting speed", B="feed rate", C="depth of cut" and D="tool material" to columns 2 to 5 from an L18. Similarly, Kim and Lee (2009) report an experiment on hybrid welding conditions of aluminum alloy in one 2-level factor and six 3-level factors, using columns 1 to 7 from the L18. Based on these published experiments, Section 5 will demonstrate how mosaic plots can help to improve experimental practice.

The rest of this section introduces some terminology around orthogonal arrays, factorial designs and their confounding structure. An orthogonal array is said to be of strength s (=resolution s+1, denoted by a roman numeral) if all level combinations of s-tuples of factors occur equally often (cf. e.g. Hedayat, Sloane and Stufken 1999). Thus, any orthogonal array has at least strength 2 (=resolution III) because of the orthogonality definition. The array of Table 1 has strength 2: it is orthogonal, but triples of the 3-level columns can of course not be replicates of a full factorial. Any combination of m columns from the array is called an "m-column projection" of the array. Obviously, any m-column projection ($m \geq s$) of a strength s orthogonal array has at least strength s. Factorial designs are often created by choosing a few columns from an orthogonal array for the experimental factors; for these, the expressions "m-factor projection", "strength" and "resolution" are the obvious analogs to the corresponding expressions for orthogonal arrays. Thus, a factorial design has strength s if and only if all its s-factor projections are (replications of) full factorials. The higher the strength of a design, the higher the degree of effects that can be estimated without confounding: in a strength 2 (=resolution III) design, main effects are unconfounded with each other but can be confounded with 2-factor interactions; in a strength 3 design, main effects are unconfounded with each other and with 2-factor interactions, but 2-factor interactions can be confounded with each other, and so forth. One often resorts to assumptions regarding negligibility of $[(s+1)/2]+1$- and higher order interaction effects (where $[\bullet]$ denotes the floor function) in order to justify use of a strength s design, for example negligibility of 2-factor and higher interactions for use of strength 2 designs or of 3-factor and higher interactions for use of strength 3 or strength 4 designs. If such an assumption is not appropriate at least approximately, the experiment can produce misleading results.

For investigating the confounding risk inherent in a factorial design of strength s, s+1-factor projections are the lowest order (=worst) case to be considered, as all projections onto lower dimensions are full factorials. The strength of a design can be no larger than s whenever there is at least one s+1 factor projection that is not a full factorial, even if the deviation from a full factorial is slight only. Thus, designs of strength s can have a different extent of confounding (=proportion of confounded s+1 factor projections), and a confounded s+1 factor projection can have different severities of confounding (from almost full factorial to completely confounded). ~~Therefore, t~~This article proposes to investigate the confounding of strength *s* designs using mosaic plots of the *s*+1 factor projections, which serve to visualize the severity of confounding for each such projection.

3 MOSAIC PLOTS

Mosaic plots are usually attributed to Hartigan and Kleiner (1981, 1984). Friendly (2002) discussed their history and traced them back to even earlier work. Friendly (1994) introduced their use for visualization of log-linear models, Friendly (1995) discussed them as part of a conceptual framework to categorical data visualization, and recently Wickham and Hofmann (2011) embedded them into a larger concept they called product plots. Mosaic plots are an integral part of modern graphical tools for categorical data (cf. e.g. Friendly, 2000; Hofmann 2003; Theus and Urbanek 2008, Section 4.1). Their construction is based on equations of the form

$$P(A=a, B=b) = P(B=b \mid A=a)\, P(A=a)$$

$$P(A=a, B=b, C=c) = P(C=c \mid B=b\,, A=a)\, P(B=b \mid A=a)\, P(A=a) \qquad (1)$$

$$\ldots,$$

where A, B and C are three categorical variables, *a*, *b*, and *c* are possible realizations of these, and the probabilities in (1) are estimated by the respective relative frequencies.

In their most common form, mosaic plots visualize the relative cell frequencies from a two-variable contingency table as areas of rectangles. For example, Figure 1 visualizes the contingency table of hair versus eye colors from Snee (1974; see Table 2; available also in **R** software (R Core Team 2013) as `HairEyeColor`):

- Figure 1a shows the marginal distribution of the Eye Color by the height of the horizontal stripes (e.g.: Brown and Blue are most frequent) and the conditional distribution of Hair Color given Eye Color by the width of the rectangles within the respective horizontal stripe (e.g.: people with brown eyes most often have brown hair, followed by black hair; people with blue eyes most often have blond hair, closely followed by brown hair).
- Figure 1b depicts the marginal distribution of Hair Color by the width of the vertical stripes (e.g.: brown is by far the most frequent) and the conditional distribution of Eye Color given Hair Color by the height of each rectangle within each vertical stripe (e.g.: the majority of blond people have blue eyes).

- Both figures visualize the bivariate distribution through the proportions of the rectangles in the total area; for example, it can be seen that brown hair together with brown eyes is the most frequent combination, while black hair together with green eyes is the rarest combination.

Table 2: The Hair and Eye Color data (Snee data extended by Sex, see Friendly 2000)

(source: R software, datasets package, HairEyeColor)

	Both Sexes				Male				Female			
	Brown	Blue	Hazel	Green	Brown	Blue	Hazel	Green	Brown	Blue	Hazel	Green
Black	68	20	15	5	32	11	10	3	36	9	5	2
Brown	119	84	54	29	53	50	25	15	66	34	29	14
Red	26	17	14	14	10	10	7	7	16	7	7	7
Blond	7	94	10	16	3	30	5	8	4	64	5	8

Mosaic plots of more than two variables introduce further splits of each rectangle (cf. Figure 2a, which includes a further split on Sex for Figure 1b). The figures illustrate that the visual impression from a mosaic plot may strongly depend on the order in which the variables are added to the plot. For example, both plots in Figure 2 enrich Figure 1b by Sex information, either as the last split (Fig. 2a) or as the first split (Fig. 2b). The different factor order draws attention to different aspects of the three-dimensional contingency table, either making it easy to assess sex proportions within eye/hair color combinations together with showing the same aspects as depicted in Fig. 1b (Fig. 2a) or supporting sex comparisons regarding the bivariate distribution of Hair Color and Eye Color (Fig. 2b). Mosaic plots can get quite messy when increasing the number of variables, which is presumably the reason many commercial software products offer them for two variables only.

Figure 1: Two mosaic plots of Hair and Eye Colors data from Snee (1974)

Relation to Equation (1): Fig.1a: Eye Color=A, Hair Color=B, Fig.1b: Hair Color=A, Eye Color=B

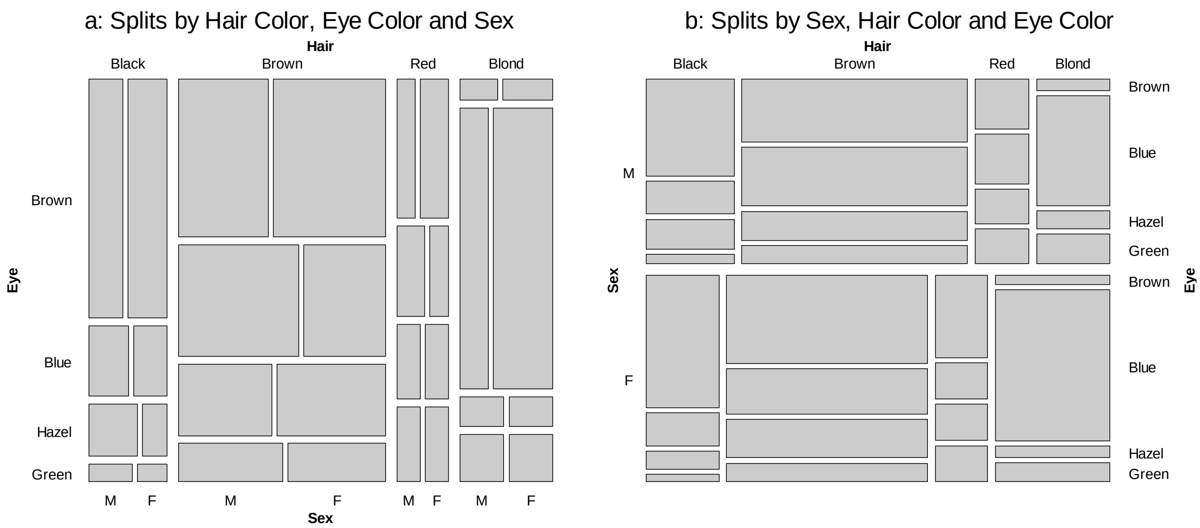

Figure 2: Three-variable mosaic plots (data from Table 2)

Relation to Equation (1):

Fig.2a: Hair Color=A, Eye Color=B, Sex=C, Fig.2b: Sex=A, Hair Color=B, Eye Color=C

4 MOSAIC PLOTS FOR ORTHOGONAL DESIGNS

For the purpose of describing its structure, an *m*-factor projection of a factorial design can be considered as a data set with *m* categorical variables, in the sense that each experimental factor has a small set of factor levels (=categories) only. Hence, mosaic plots can also visualize relative frequencies of *m*-factor projections for factorial designs. It was pointed out in Section 2 that all *s*-factor projections of a strength *s* design are (replications of) full factorials. Any full factorial projection is perfectly balanced; accordingly, regardless of factor order, its mosaic plot conveys a visual impression of perfect balance: it consists of equally-sized rectangles, with all category splits reaching from end to end, like in Figure 3a (this is due to the fact that marginal and conditional probabilities in Equation (1) coincide for any tuple of factors).

For strength *s* designs, as it is known a-priori that *s*-factor projections are full factorials, mosaic plots of *s*-factor projections are uninformative; *s*+1 factor projections are the lowest dimensional projections whose inspection is of interest, as was already mentioned in Section 2. In all mosaic plots of *s*+1 factors, the first *s* splits are from end to end, i.e. equally-sized rectangles from the first *s* splits are sub-divided by the *s*+1st split. This makes *s*+1 factor mosaic plots much easier to grasp for strength *s* orthogonal arrays than for general data sets with *s*+1 categorical variables. This section discusses the specific properties of mosaic plots for strength 2 designs (Subsection 4.1), their relation to estimability and confounding (Subsection 4.2), and the more general strength *s* situation exemplified with *s*=3 (Subsection 4.3).

4.1 Mosaic plots for triples of factors in strength 2 designs

In this section, the L18 from Table 1 is used as an example. According to Schoen (2009, Table III), there are exactly four non-isomorphic triples for 18 run designs in three 3-level factors and exactly three for one 2-level with two 3-level factors; only three of the former and two of the latter can be obtained from the L18 of Table 1, as has been determined by manual isomorphism searches among the projections of the L18 (see supplementary material). Figures 3 and 4 show representatives for all five different confounding patterns that occur among the 8 choose 3 (=56) three-column projections of the Taguchi L18 of Table 1; confounding increases from left to right. The behavior of the third=last split indicates the severity of confounding in the triple of columns. A flat line in the mosaic plot instead of a proper rectangle indicates that the respective combination does not occur at all; for example, in Figure 4c, two of the three possibilities for column 5 do not occur for any combination of columns 2 and 4, i.e. the level combination of columns 2 and 4 completely determines column 5; this is the worst severity of confounding possible in a strength 2 orthogonal array. At the other extreme, the last split can be from end to end like the first two splits (cf. Figure 3a): this indicates that the array for the particular triple (ignoring all other columns) is a (replicate of a) full factorial, i.e. is perfectly balanced.

The key aspect conveyed by a mosaic plot is the severity of (im)balance in a design. This point will be discussed in more detail in the next section.

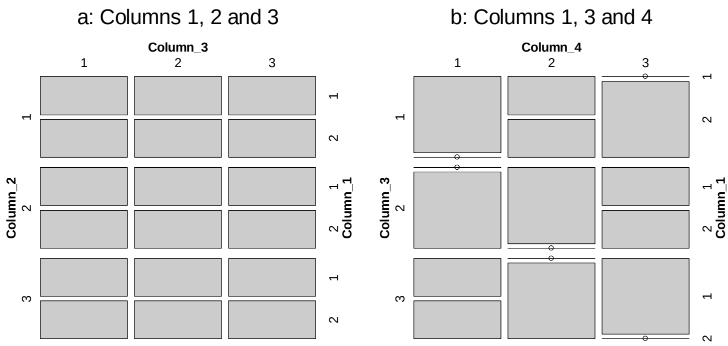

Figure 3: The two types of 3-column projections involving the 2-level column in the Taguchi L18 (there are 12 triples of the full factorial type and 9 triples of the right-hand side nature)

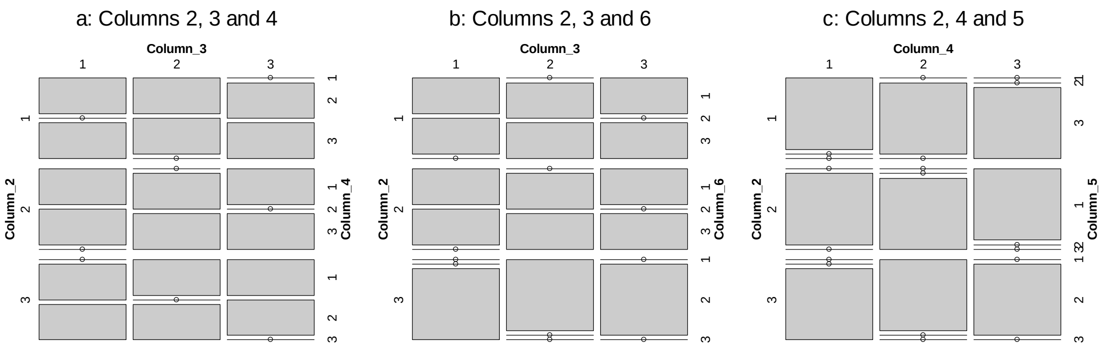

Figure 4: The three types of 3-column projections with only 3-level columns in the Taguchi L18 (there are 28 triples, 6 triples and 1 triple of the respective sort, from left to right)

4.2 Confounding, effect estimation, and mosaic plots

To a pure mathematician, strength of a design is beautiful in itself. To an experimenter, it is important because strength is related to estimability and bias risk. Figure 5a shows the best possible structure for three 2-level factors: a full factorial design, i.e. a design of strength 3; all effects up to the highest interaction are estimable, and the naïve main effect estimator (difference of means) for each factor remains unbiased, even if the other two factors interact. Estimability can be harmed by level combinations that do not occur in the experiment (the flat lines in a mosaic plot). In extreme cases, effects can be completely confounded; this is for example well-known for main effects and 2-factor interactions in some triples of factors in regular resolution III fractional factorial 2-level designs. Figure 5c shows a (replicate of a) regular fractional factorial 2-level design in 4 runs with 3 factors, for which the combination of factors A and B completely determines the level of factor C, i.e. only one level of factor C occurs for any combination of factors A and B; note that the mosaic plot and the consequences of aliasing in the triple are the same, regardless of the number of runs this triple is a projection of. Figure 5b depicts any triple of 2-level factors from a 12 run Plackett-Burman design (Plackett and Burman 1946; note that all three-factor projections of a 12-run Plackett-Burman design consist of a full factorial augmented with a resolution III half fraction thereof, see e.g. Lin and Draper 1992): the figure shows that both levels of factor C occur for each level combination of factors A and B, however with unequal frequencies (ratio 2:1; in general, the exact proportions are difficult to read off the mosaic plot). For this more balanced triple of factors, main effects and 2-factor interactions (and even the three-factor interaction, provided it is not confounded with some other effect from the overall design) are all estimable, but nevertheless confounded; here, the milder consequence of confounding is that each naïve main effect estimator (difference of means) will be biased if there is a non-negligible 2-factor interaction between the other two factors. The size of a potential bias of naïve main effects estimators depends on two aspects: (a) the larger the confounded interactions the larger the bias; (b) the more severe the confounding in the experimental plan, the larger the bias (e.g. bias would of course be larger in Figure 5c than in Figure 5b). While the size of the interaction is usually unknown, the confounding structure of the experimental plan can be assessed and – as far as resource constraints permit – influenced.

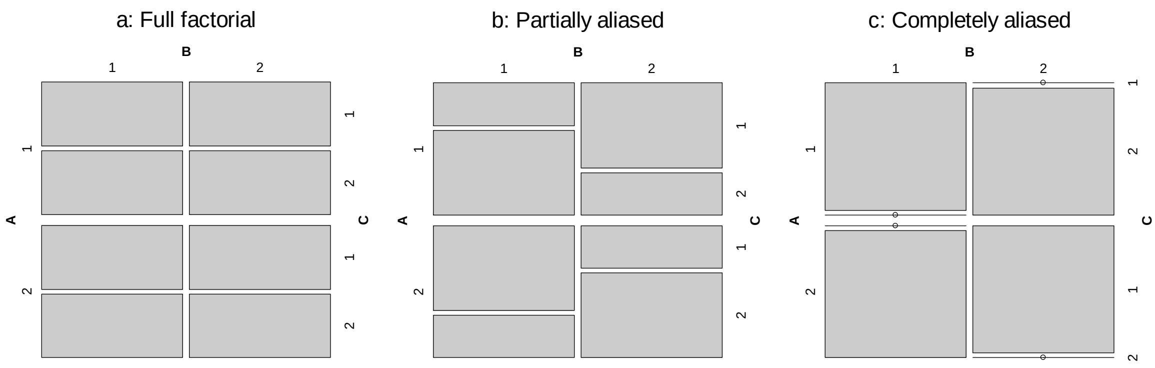

Figure 5: Mosaic plots of two strength 2 triples of three 2-level factors

Figures 4c and 5c show complete confounding, Figures 3a and 5a show full factorial situations, and all other mosaic plots from Figures 3 to 5 show partial confounding. The partial confounding of the 3-factor projection in Figure 5b is relatively mild only, in the sense that all level combinations occur at least once and the imbalance is weak. All confounded cases in Figures 3 and 4 show at least some empty cells; hence, it will generally not be possible to estimate all effects up to the 3-factor interaction, even if no further factors have to be considered; usually, even the main effects and 2-factor interaction estimates are not completely obtainable from such designs. The plots above (Figures 3 to 5) show 3-factor projections of relatively small designs. Larger non-regular orthogonal arrays may contain projections that look almost but not quite like a full factorial design and thus yield almost unbiased naïve main effects estimates even in case of moderate interactions.

As was mentioned before, the key feature of a mosaic plot for diagnosing experimental designs is the amount of balance. This is best reflected by the deviation from the ideal full factorial structure, which is a rectangular grid for any combination of factor levels (e.g. Figure 5a). The comparison to this ideal provides a good impression of the imbalance present in the design, as is for example observed when comparing Figures 5b and 5c to Figure 5a: Figure 5c shows only half as many proper rectangles of twice the expected size, which indicates strong imbalance. Figure 5b shows far less imbalance, since all 8 rectangles are present with much more moderate size differences. Thus, the mosaic plots provide an assessment of structure that can be quickly grasped by statisticians and non-statisticians alike and can be easily used for sensitizing non-statisticians to imbalance issues, once the full factorial pattern of equally-sized rectangles has been established as the ideal.

It is also important to realize which features of a mosaic plot are not relevant for the quality of a design. For this purpose, Figure 6 illustrates two mosaic plots of the same strength 2 design, with the factor levels arranged in two different orders. The visual impression is quite different, but the severity of imbalance vs. a plot of a full factorial design (which would have twice as many equally-sized rectangles and no flat lines) is the same for both plots. Note that this consideration is valid for qualitative factors, where the order of the factor levels is arbitrary, so that plots a and b of Figure 6 show equivalent designs. For quantitative factors, the designs would not be considered equivalent, and the design of Figure 6a would be considered worse than design of Figure 6b, because it completely confounds the low=(1,2) vs high=(3,4) comparison of any factor with the 2-factor interaction of the other two low/high comparisons, which would be considered more serious than confounding the (1,3) vs. (2,4) comparisons like in Figure 6b.

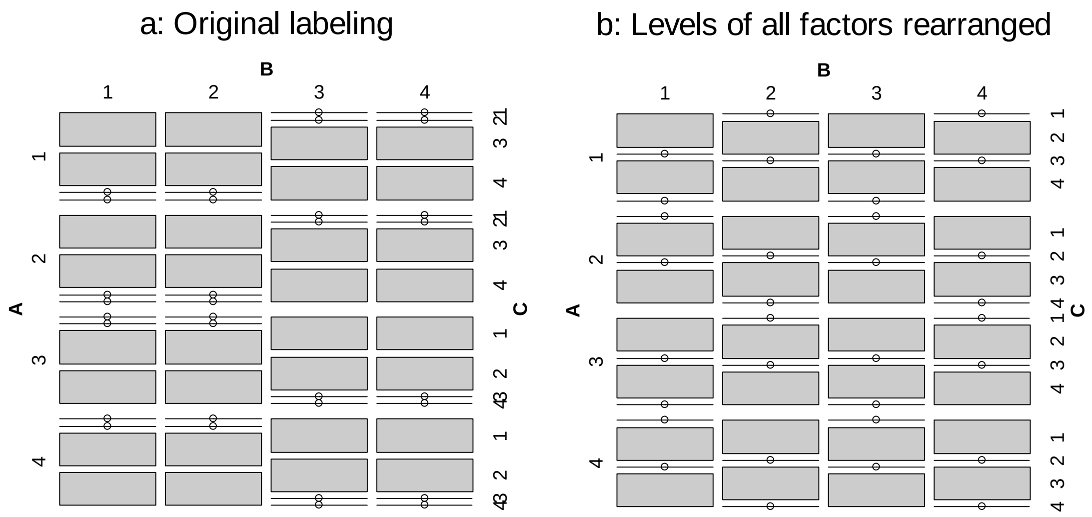

Figure 6: Two mosaic plots for the same triple of 4-level factors

4.3 Mosaic plots for quadruples of factors in strength 3 designs

Figure 7 shows mosaic plots for three strength 3 quadruples of 2-level factors. Now all four margins are used for labeling factor categories. The strength is reflected by the fact that the separation lines are end to end for the first three factors (A, B and C). The designs are the full factorial reference design (Fig.6a, number of runs a multiple of 16) and two four-factor projections obtainable from a 24 run Plackett-Burman design (see e.g. Schoen and Mee 2012; according to Eendebak and Schoen 2013, these are the only two non-isormorphic strength 3 orthogonal arrays in four 2-level factors and 24 runs): Figure 7b shows the best possible strength 3 projection from a 24 run Plackett-Burman design, which is a full factorial augmented with a resolution IV 8 run design; the figure shows that – within each combination of factors A, B and C – both levels of factor D occur, however not with the same frequency, but a ratio of 2:1 (again, the exact ratio is difficult to read from the figure). The other strength 3 projection of the same design in Figure 7c is a triplicate of the resolution IV 8 run design; the figure reveals that any triple of the factors A, B and C completely specifies the level of factor D, i.e. only one level of factor D occurs within each combination of A, B and C. The severity of imbalance obviously increases from left to right, going along with an increase in bias of the naïve main effects estimator in the presence of 3-factor interactions (or of the naïve estimator for e.g. the 2-factor interaction of factors A and B in case a 2-factor interaction of factors C and D is present).

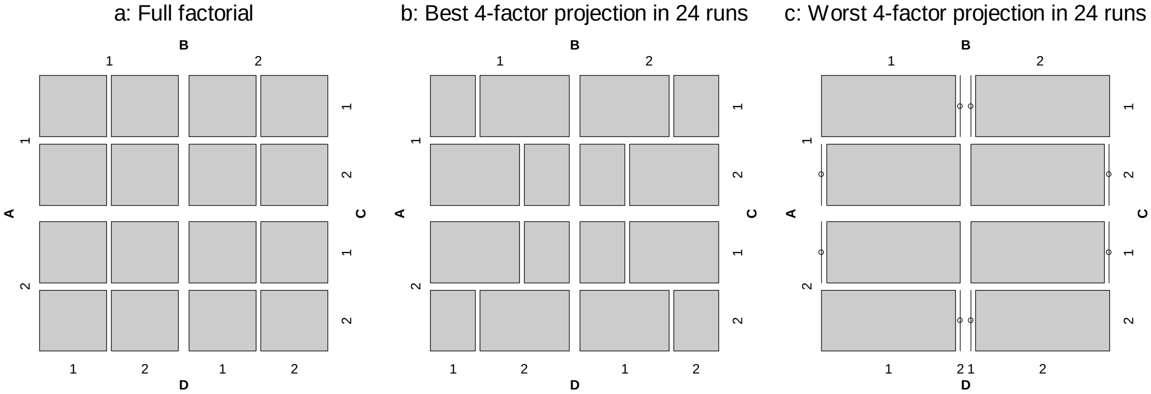

Figure 7: Mosaic plots of strength 3 quadruples of 2-level factors

Conceptually, mosaic plots for *s*+1 factors in strength *s* designs can be used for any *s*; in practice, the idea is limited by space constraints, especially for accommodating labels for the factor levels. All four margins are used for four-factor projections; with the next dimension, one margin has to be used for two factors. In practice, one will rarely consider mosaic plots for more factors than four at a time.

5 MOSAIC PLOTS IN EXPERIMENTAL PRACTICE

Orthogonal main effects plans (e.g. Addelman 1962; Gupta et al. 1982; Suen and Kuhfeld 2005) have been proposed for ensuring that main effects do not bias *each other*; they are strength 2 plans. If interest is chiefly in estimation of main effects, they are considered useful, because they are parsimonious in the number of runs and permit uncorrelated estimation of all main effects. Some effort has gone into identifying minimum size orthogonal main effects plans (e.g. Jacroux 1992, Mukerjee and Wu 1995, Gallant 1997). Orthogonal main effects designs have been used in initial phases of industrial experimentation or in market research (cf. e.g. Kuhfeld and Tobias 2005), for example. It was mentioned in the introduction that such plans are used in experimental practice, based on published catalogues. However, estimates from orthogonal main effects plans are unbiased only *provided that two-factor and higher interactions are negligible*. Experimenters have not always been aware, how severely analysis results from such plans may depend on this assumption. Mosaic plots can help create such awareness, which is helpful for both design creation and analysis.

5.1 Design creation

Many engineers have access to books or software with orthogonal arrays, among others the orthogonal array of Table 1, and it is customary to select a few columns from such an array for a smaller experiment, usually the first few that accommodate the required levels. The example designs reported in Section 2 are of this nature: Dhole et al. (2012) used the first four 3-level columns of the L18 for accommodating the four factors A="cutting speed", B="feed rate", C="depth of cut" and D="tool material", Kim and Lee (2009) used the only 2-level column and the first six 3-level columns of the L18. Both designs could have been made more robust against bias of main effects from 2-factor interactions by avoiding the completely aliased triple of columns 2, 4 and 5 of the L18 that is depicted in Figure 4c.

In the four-factor design, it would have been easy to look at the four mosaic plots of triples of factors before accepting the design; this could have been done manually or supported by a software option to browse all 3-factor projections (available in R package `DoE.base` by Grömping (2013a), see supplementary material). In either case, the plot of design factors A, C and D would have revealed the strong confounding of each main effect with the two-factor interaction of the other two factors. In this simple case, replacing any of the three columns behind factors A, C or D with one of the unused 3-level columns would have improved the design; thus, even equipped with only mosaic plots and common sense, Dhole et al. might have ended up with column 6 instead of 5 (which would have been improved but not ideal, yielding three triples like Fig. 4a and one like Fig. 4b) or with the L18's last four 3-level columns (5 to 8) instead of the first ones (which would have been one of the optimum choices with four triples like in Fig. 4a, like any choice without column 2). For situations with more than very few factors, a strategy based solely on looking at all mosaic plots becomes impractical. For example, Kim and Lee would have had to look at at least 35 mosaic plots (all triples of the seven 3-level columns for the L18) in order to safely avoid the one particularly badly biased triple when manually trying to improve their design. With the help of numerical criteria for selecting relevant plots to look at, mosaic plots can still be very useful for designs like the one used by Kim and Lee; the R package `DoE.base` allows to look at all the completely confounded projections or at a user-specified percentage of worst projections. Applying the latter tool to the L18, it would also have been easy to improve Kim and Lee's design (again, see supplementary online material). Nevertheless, manual creation of good designs with such simple tools is limited, especially as for more complicated designs things will not be as obvious as avoidance of a few badly aliased triples.

Generally, it will be necessary to determine a good or the best design by numerical criteria implemented in software or applied by a statistical expert. Identification of good orthogonal arrays is a current research topic (see also the final section). The R package `DoE.base` (Grömping 2013) offers an experimental automated column allocation optimization according to criteria proposed in Xu, Cheng and Wu (2004) and Schoen (2009) and extended in Grömping (2011). The development and implementation of further quality criteria is in progress. The theoretical confounding criteria are not easily accessible to statistical lay persons; statistical experts or software may use mosaic plots as a means of communicating the strength of the confounding to experimenters and decision makers. This communication may support interdisciplinary decision processes about whether a design appears so badly confounded that it may be better to decide on a larger design beforehand, or which factors to allocate to a badly confounded triple of array columns.

5.2 Mosaic plots can support the analysis of data from an orthogonal main effects plan

In regular fractional factorial 2-level designs, it is considered good practice to look at the confounding structure of a design, which is easily given as a list of aliased effects. For orthogonal main effects plans, which are often non-regular designs, such listings are not routinely considered and would be much more complicated to obtain. It is useful to inspect the design for particularly severe confounding, because this will have the most detrimental consequences in terms of biased estimates. Therefore, the afore-mentioned possibility to show mosaic plots for a user-specified proportion of worst-confounded projections can also help in analyzing experimental results. For example, the procedure will show up the triple of factors A, C and D as completely confounded in the experiment conducted by Dhole et al. (2012). Suppose (completely fictitious) that these authors encounter unexpected results for the main effect of factor “cutting speed” (A), or that they obtain subsequent confirmation runs that do not react as expected to changes in this factor. Then, being warned about the confounding issue in the triple A, C and D will suggest to consider the interaction of “depth of cut” (C) and “tool material” (D) as a potential cause of the unexpected result.

6 FINAL REMARKS

Severe confounding like in Figures 4c, 5c, or 6c should be avoided whenever possible. Such avoidance can be supported by mosaic plots, as was illustrated in Section 5.1. Whenever confounding as severe as in these figures appears unavoidable, allocation of design factors to array columns should be carefully managed, and interpretation of results requires awareness of the strength and nature of the confounding; a mosaic plot can help with both these needs. Awareness of the confounding is particularly necessary, when orthogonal main effects plans (cf. Section 5) are used; it appears that their users sometimes feel overly safe-guarded against bias in main effects estimates. Awareness of the confounding pattern – supported by mosaic plots – can also help interpret experimental results correctly in moderate confounding situations. Good interpretability of mosaic plots is limited to three- or at most four-factor projections, which is sufficient for the purpose of checking main effects against bias risk from low order interactions.

Some recent research deals with finding optimal experimental plans based on orthogonal arrays: systematic catalogues of non-isomorphic arrays have been created and investigated according to various criteria for comparing designs (e.g. Evangelaras, Koukouvinos and Lappas 2007, Schoen 2009). The so-called projection frequency tables proposed by Xu, Cheng and Wu (2004) are particularly closely related to mosaic plots, as both focus on s+1 tuples of factors for strength s designs. Projection frequency tables are based on generalized aberration as introduced by Xu and Wu (2001), and Grömping (2013b) proposed a modification and further types of tables for mixed level designs. The search for adequate criteria for ranking orthogonal arrays – especially with mixed levels – is an active field of research. Whatever criteria the future brings, mosaic plots are likely to remain useful tools for visualizing a design's low-order confounding properties.

7 SUPPLEMENTARY MATERIAL

The file "Online Supplement.zip" contains a ReadMe file and five files with R code for creation of the mosaic plots in the article, for the two examples, and for the isomorphism checks for the three-column projections of the L18. The code files make use of R packages `DoE.base` (Grömping 2013) and `vcd` (Meyer, Zeileis and Hornik 2006).

8 REFERENCES

Addelman, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics* **4**, 21-46.

Dhole, N.S., Naik, G.R and Prabhawalkar, M.S. (2012). Optimization of milling process parameters of EN33 using Taguchi parameter design approach. *Journal of Engineering Research and Studies* III, 70-74.

Eendebak, P. and Schoen, E. (2013). Complete series of non-isomorphic orthogonal arrays. URL: http://pietereendebak.nl/oapage/ Accessed May 24, 2013.

Evangelaras, H., Koukouvinos, C. and Lappas, E. (2007). 18-run nonisomorphic three level orthogonal arrays. *Metrika* **66**, 31-37.

Friendly, M. (1994). Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association* **89**, 190–200.

Friendly, M. (1995). Conceptual and visual models for categorical data. *The American Statistician* **49**, 153–160.

Friendly, M. (2000). *Visualizing Categorical Data*. SAS Publishing, Cary, NC.

Friendly, M. (2002). A brief history of the mosaic display. *Journal of Computational and Graphical Statistics* **11**, 89-107.

Gallant, R. (1997). Tight orthogonal main effects plans. PhD thesis, University of Waterloo, Ontario, Canada.

Grömping, U. (2013a). `DoE.base`: Full factorials, orthogonal arrays and base utilities for DoE packages. R package version 0.25-2. In: R Core Team (2013).

Grömping, U. (2013b). Frequency tables for the coding invariant ranking of orthogonal arrays. *Report 2/2013, Reports in Mathematics, Physics and Chemistry*, Department II, Beuth University of Applied Sciences Berlin.

Hartigan, J. A., and Kleiner, B. (1981). Mosaics for contingency tables. In W. F. Eddy (Ed.), *Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface*. New York: Springer-Verlag.

Hartigan, J. A., and Kleiner, B. (1984). A mosaic of television ratings. *The American Statistician* **38**, 32-35.

Hedayat, A.S., Sloane, N.J.A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications*. Springer, New York.

Hofmann, H. (2003). Constructing and Reading Mosaicplots. *Computational Statistics and Data Analysis* **43**, 565–580.

Jacroux, M. (1992). A note on the determination and construction of minimal orthogonal main-effect plans. *Technometrics* **34**, 92-96.

Kim, H.R. and Lee, K.Y. (2009). Application of Taguchi method to determine hybrid welding conditions of aluminium alloy. *Journal of Scientific and Industrial Research* **68**, 296-300.

Kuhfeld, W.F. and Tobias, R.D. (2005). Large Factorial Designs for Product Engineering and Marketing Research Applications. *Technometrics* **47**, 132–141.

Lin, D.K.J. and Draper, N.R. (1992). Projection properties of Plackett and Burman designs. *Technometrics* **34**, 423-428.

Meyer, D., Zeileis, A. and Hornik, K. (2006). The Strucplot Framework: Visualizing Multi-way Contingency Tables with `vcd`. *Journal of Statistical Software* **17**, Issue 3, 1:48.

Mukerjee, R. and Wu, C.F.J. (1995). On the existence of saturated and nearly saturated asymmetrical orthogonal arrays. *Annals of Statistics* **23**, 2102-2115.

NIST/SEMATECH (2012). *e-Handbook of Statistical Methods*. URL: http://www.itl.nist.gov/div898/handbook/, June 18 2012.

Plackett, R.L. and Burman, J.P. (1946). The Design of Optimal Multifactorial Experiments. *Biometrika* **33**, 305-325.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Sartono, B., Goos, P. and Schoen, E. (2012). Classification of three-level strength-3 arrays. *Journal of Statistical Planning and Inference* **142**, 794-809.

Schoen, E. (2009). All OAs with 18 runs. *Quality and Reliability Engineering International* **25**, 467-480.

Schoen, E. and Mee, R. (2012). Two-level designs of strength 3 and up to 48 runs. *Applied Statistics* **61**, 163-174.

Snee, R.D. (1974). Graphical display of two-way contingency tables. *The American Statistician* **28**, 9-12.

Suen, C. and Kuhfeld, W.F. (2005). On the construction of mixed orthogonal arrays of strength two. *Journal of Statistical Planning and Inference* **133**, 555-560

Theus, M. and Urbanek, S. (2008). *Interactive Graphics for Data Analysis: Principles and Examples*. Chapman & Hall / CRC, Boca Raton.

Wickham, H. and Hofmann, H. (2011). Product Plots. *IEEE Transactions on Visualization and Computational Graphics* **17**, 2223-2230.

Wu, C.F.J. and Hamada, M. (2009, 2nd ed.). *Experiments: Planning, Analysis and Optimization*. Wiley, New York.

Xu, H. and Wu, C.F.J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Annals of Statistics* **29,** 1066-1077.

Xu, H., Cheng, S.W. and Wu, C.F.J. (2004). Optimal projective three-level designs for factor screening and interaction detection. *Technometrics* **46**, 280–292.