# Some Investigations on Orthogonal Arrays*

D.V. Chopra
Wichita State University
Wichita, Kansas 67260-0033, U.S.A.

R. Dios
New Jersey Institute of Technology
Newark, New Jersey 07102, U.S.A.

ABSTRACT. In this paper we obtain some results on orthogonal arrays (O-arrays) of strength six by considering balanced arrays (B-arrays) of strength six with $\underline{\mu}' = (\mu - 1, \mu, \mu, \mu, \mu, \mu, \mu - 1)$ which we call Near O-arrays. As a consequence we demonstrate that we obtain better bounds on the number of constraints for some O-arrays as compared to those given by Rao (1947).

## 1 Introduction and Preliminaries

First we state some basic concepts and definitions.

**Definition 1.1.** An *array* $T$ with $s$ symbols (levels), $m$ rows (constraints), and $N$ columns (runs, treatment-combinations) is a matrix $T$ of size $(m \times N)$ with $s$ elements (say; $0, 1, \ldots, s-1$). $T$ is denoted by $T$ $(m, N; s)$. Imposing a combinatorial structure on $T$ leads us to the following definition of a balanced array (B-array).

**Definition 1.2.** A $T(m, N; s)$ is called a *balanced* array (B-array) of *strength* $t$ if for every $(t \times N)$ submatrix $T^*$ of $T$, we have: $\lambda(\underline{\alpha}, T^*) = \lambda(P(\underline{\alpha}), T^*)$ where $\underline{\alpha}$ is a $(t \times 1)$ vector of $T^*$, $\lambda(\underline{\alpha}, T^*)$ is the frequency with which $\underline{\alpha}$ appears in $T^*$, and $P(\underline{\alpha})$ is any vector obtained by permuting the elements of $\underline{\alpha}$.

**Definition 1.3.** A B-array $T$ for which

$$\lambda(\underline{\alpha}, T^*) = \mu \quad (\mu \text{ is a constant})$$

*Paper presented at the Thirteenth Midwestern Conference on Combinatorics, Cryptography and Computing, Normal, Illinois, October 1999.

is called an *orthogonal* array (O-array) where $\underline{\alpha}$ is any $(t \times 1)$ vector in $T^*$.

Clearly an O-array is a special case of a B-array and for an O-array $N = \mu s^t$.

In this paper we restrict ourselves to arrays with $t = 6$ and $s = 2$ (elements 0 and 1). In this special case $w(\underline{\alpha})$, the *weight* of the vector $\underline{\alpha}$, is defined to be the number of non-zero elements in $\underline{\alpha}$. Clearly for $w(\underline{\alpha}) = k (0 \leq k \leq 6)$, we have $\lambda(\underline{\alpha}, T^*) = \mu_k$ for every submatrix $T^*$ of the B-array $T$. The vector $\underline{\mu}'=(\mu_0, \mu_1, \dots, \mu_6)$ is called the index set of the array $T$, and $T$ is sometimes denoted by $T$ $(m, N; \underline{\mu}')$. Clearly

$$N = \sum_{i=0}^{t} \binom{t}{i} \mu_i.$$

**Definition 1.4.** A B-array T with $\underline{\mu}'=(\mu_0, \mu_1, ..., \mu_6)$ satisfying $\mu_i = \mu (i = 1, 2, .., 5)$ and $\mu_i = \mu - 1 (i = 0, 1, \dots, 6)$ is called a near O-array.

**Remark:** Clearly, if we affix to an $m$-rowed near O-array with $\underline{\mu}'=(\mu_0, \mu_1, ..., \mu_6)$ a vector of weight 0 and a vector of weight $m$, we obtain an $m$-rowed O-array with index set $\mu$.

Next we state the following results on O-arrays from C.R. Rao (1947):

**Result:** For O-arrays with $t \geq 2, m$ must satisfy the following inequalities:

$$\mu s^t - 1 \geq \binom{m}{1}(s-1) + \binom{m}{2}(s-1)^2 + \cdots + \binom{m}{u}(s-1)^u \text{ if } t = 2u,$$

$$\mu s^t - 1 \geq \binom{m}{1}(s-1) + \cdots + \binom{m}{u}(s-1)^u + \binom{m-1}{u}(s-1)^{u+1} \text{ if } t = 2u+1.$$

The first inequality for $s = 2$ and $t = 6$ is reduced to

$$384\mu - 6 \geq m(m^2 + 5). \tag{1.1}$$

We use (1.1) later to compare our bounds on $m$ with those obtained from Rao's Inequalities.

B-arrays and O-arrays tend to unify various areas of combinatorial mathematics, and have been extensively used in constructing symmetrical as well as asymmetrical fractional factorial designs. It is quite obvious that these arrays for an arbitrary set of parameters may not exist. To construct such arrays with the maximum possible number of constraints is very important both in combinatorial mathematics and experimental designs. This problem for O-arrays has been studied, among others, by Bose and/or Bush (1952, 1952), Rao (1946, 1947), Seiden and Zemach (1966), Yamamoto et al. (1993), etc., while the corresponding problem for B-arrays has been studied, among others, by Chopra and Dios (1997), Saha et al. (1988), Yamamoto et al. (1985), etc. To gain further insight into the importance of O-arrays and B-arrays to statistical designs and combinatorics, the interested reader may consult the list of references (absolutely not an exhaustive list) at the end of this paper, and also further references listed therein.

## 2 Main Results

In what follows we consider near O-arrays, i.e., B-arrays with $\underline{\mu}'=(\mu - 1, \mu, \mu, \mu, \mu, \mu, \mu - 1)$ which can be written as $\mu\ J_7 - (J_1, 0J_5, J_1)$ where $J_k$ is a $(1 \times k)$ vector with each element being unity. The following results can be easily established.

**Lemma 2.1.** *A near O-array $T$ with index set $\underline{\mu}'= \mu\ J_7 - (J_1, 0J_5, J_1)$ is also of strength $t'$ where $t'$ is such that $0 < t' \leq t = 6$. Considered as an array of strength $t'$, its index set $(A_0^{t'}, A_1^{t'}, A_2^{t'}, .., A_{t'}^{t'})$ is given by $2^{t-t'}\mu\ J_{t'+1} - (J_1, 0J_{t'-1}, J_1)$. Clearly $A_{t'}^{t'} = 2^{t-t'}\mu - 1$.*

**Remark:** It is quite obvious from the above lemma that the values of $A_{t'}^{t'}(t' = 6, 5, 4, 3, 2,\text{and } 1)$ in terms of the index parameter $\mu$ are $A_6^6 = \mu - 1$, $A_5^5 = 2\mu - 1$, $A_4^4 = 2^2\mu - 1$, $A_3^3 = 2^3\mu - 1$, $A_2^2 = 2^4\mu - 1$, and $A_1^1 = 2^5\mu - 1$. Here $A_k^{t'}(k = 0, 1, 2, .., t'$ and $0 < t' \leq t = 6)$ is the number of times $(t' \times 1)$ vector of weight $k$ appears in every submatrix $T^*$ $(t' \times N)$ of $T$. Here $N = 64\mu - 2$.

**Lemma 2.2.** *Let $x_j$ be the number of columns of weight $j(0 \leq j \leq m)$ in a near O-array of strength 6 with $m$ rows. Then the following results hold:*

$$
\begin{aligned}
N &= \sum_{j=0}^{m} x_j = 64\ \mu - 2 \\
\sum j x_j &= m(32\mu - 1) = m\ A_1^1 \\
\sum j^2 x_j &= m_2\ A_2^2 + m\ A_1^1 \qquad\qquad (2.1)\\
\sum j^3 x_j &= m_3\ A_3^3 + 3m_2\ A_2^2 + m\ A_1^1 \\
\sum j^4 x_j &= m_4\ A_4^4 + 6m_3\ A_3^3 + 7m_2\ A_2^2 + m\ A_1^1 \\
\sum j^5 x_j &= m_5\ A_5^5 + 10m_4\ A_4^4 + 25m_3\ A_3^3 + 15m_2\ A_2^2 + m\ A_1^1 \\
\sum j^6 x_j &= m_6\ A_6^6 + 15m_5\ A_5^5 + 65m_4\ A_4^4 + 90m_3\ A_3^3 + 31\ m_2\ A_2^2 + m\ A_1^1
\end{aligned}
$$

*where/ $m_r$ stands for/ $m(m-1)(m-2)....(m-r+1)$, and values of/ $A_k^k$ (a linear function of/ $\mu$) are as defined in the last lemma.*

**Remark:** The above equalities can be easily derived by applying Lemma 2.1 and counting in two ways (through columns and through rows) the number of vectors of different weights $k$ $(k = 1, 2, .., 6)$. These equalities clearly represent moments of order $k$ of the weights of the column vectors of $T$ in terms of the parameters $\mu$ and $m$.

Next we state some results, without proofs, from statistics involving central moments which are used to study empirical and theoretical distributions (See Chakrabarti (1946), and Lakshmanamurti (1950)).

**Result** Let $Z_i (i = 1, 2, .., N)$ be reals satisfying $\sum_{i=1}^{N} Z_i = 0$ and $\sum_{i=1}^{N} Z_i^2 = 1$. Set $\alpha_k = \frac{1}{N} \sum Z_i^k$. Then the following result holds:

$$\alpha_6 \geq \alpha_4^2 + \alpha_3^2 \tag{2.2}$$

**Note:** $\alpha_4$ and $\alpha_3$, called kurtosis and skewness, have been extensively used in statistics to study "Peakedness" and "Symmetry". We use (2.1) and (2.2) to obtain some new results on the existence of near O-arrays and O-arrays.

**Theorem 2.1.** *Consider a near O-array $T(m \times N)$ with $t = 6$ and $\underline{\mu}' = \mu\, J_7 - (J_1, 0J_5, J_1)$. Then the following inequality is true:*

$$L_2 L_6 \geq L_4^2 + L_2\, L_3^2 \tag{2.3}$$

*where*

$L_2 = N \sum j^2 x_j - (\sum j x_j)^2$

$L_3 = N^2 \sum j^3 x_j - 3N \sum j^2 x_j \sum j x_j + 2(\sum j x_j)^3$

$L_4 = [N^3 \sum j^4 x_j - 4N^2 \sum j^3 x_j \sum j x_j + 6N \sum j^2 x_j (\sum j x_j)^2 - 3(\sum j x_j)^4]$

$$L_6 = \left[ \begin{array}{c} N^5 \sum j^6 x_j - 6N^4 \sum j^5 x_j \sum j x_j + 15N^3 \sum j^4 x_j (\sum j x_j)^2 - \\ -20N^2 \sum j^3 x_j (\sum j x_j)^3 + 15N \sum j^2 x_j (\sum j x_j)^4 - 5(\sum j x_j)^6 \end{array} \right].$$

**Proof:** Set

$$\frac{\sum j x_j}{N} = M \tag{2.4a}$$

and

$$\frac{1}{N} \sum (j - M)^2 x_j = s^2. \tag{2.4b}$$

Clearly, $\sum \left(\frac{j-M}{s}\right) x_j = 0$, and $\sum \left(\frac{j-M}{s}\right)^2 x_j = N$.

Let $\alpha_k = \frac{1}{N} \sum \left(\frac{j-M}{s}\right)^k x_j$. Using (2.2), we obtain

$\frac{1}{N} \sum \left(\frac{j-M}{s}\right)^6 x_j \geq \frac{1}{N^2 s^8} \left[\sum (j-M)^4 x_j\right]^2 + \frac{1}{N^2 s^6} \left[\sum (j-M)^3 x_j\right]^2.$

This leads to

$N\, s^2 \left[\sum (j-M)^6 x_j\right] \geq \left[\sum (j-M)^4 x_j\right]^2 + s^2 \left[\sum (j-M)^3 x_j\right]^2.$

Next, we expand each term in the above inequality in terms of $\sum j^k x_j$ as follows:

$$s^2 = \frac{1}{N^2} \left[N \sum j^2 x_j - (\sum j x_j)^2\right]$$

$$\sum(j-M)^3x_j = \sum j^3x_j - 3M\sum j^2x_j + 2NM^3$$
$$= \frac{1}{N^2}\left[\begin{array}{c} N^2\sum j^3x_j \\ -3N\sum j^2x_j\sum jx_j + 2(\sum jx_j)^3 \end{array}\right]$$

$$\sum(j-M)^4x_j = \sum j^4x_j - 4M\sum j^3x_j + 6M^2\sum j^2x_j - 3NM^4$$
$$= \frac{1}{N^3}\left[\begin{array}{c} N^3\sum j^4x_j - 4N^2\sum j^3x_j\sum jx_j \\ +6N\sum j^2x_j(\sum jx_j)^2 - 3(\sum jx_j)^4 \end{array}\right]$$

$$\sum(j-M)^6x_j = \sum j^6x_j - 6M\sum j^5x_j + 15M^2\sum j^4x_j$$
$$- 20M^3\sum j^3x_j + 15M^4\sum j^2x_j - 5NM^6$$

$$= \frac{1}{N^5}\left[\begin{array}{c} N^5\sum j^6x_j - 6N^4\sum j^5x_j\sum jx_j + 15N^3\sum j^4x_j(\sum jx_j)^2 - \\ -20N^2\sum j^3x_j(\sum jx_j)^3 + 15N\sum j^2x_j(\sum jx_j)^4 - 5(\sum jx_j)^6 \end{array}\right].$$

Substituting the value of $s^2$ and after some simplification we obtain the desired result. □

Next we give two illustrative examples to demonstrate that the maximum number of constraints for O-arrays (obtained from near O-arrays) by using (2.3) are better than those given by C.R. Rao. We also include a table which compares, for various values of $\mu$, the number of constraints obtained using Rao's results with those obtained using the results of this paper. To accomplish all this, a computer program was prepared to check (2.3) for different values of $\mu$ to make a comparison with Rao's bounds.

**Example 1.** Consider an O-array with $\mu = 1$. Using Rao's result (1.1), we obtain $m \le 7$. The near O-array (0,1,1,1,1,1,0) has $\underline{\mu}' = (0, J_5, 0)$. It was found that for $m = 7$, L.H.S. of (2.3) is 5.508313E+14, while the R.H.S. = 3.053855E+15 which is a contradiction. Hence, according to our result, $m \le 6$. Thus the O-array obtained from the near O-array by adding a $(6 \times 1)$ vector of weight 0 and a $(6 \times 1)$ vector of weight 6 has $m \le 6$. Clearly such an O-array exists for $m = 6$.

**Example 2.** Consider an O-array with $\mu = 9$. Using Rao's inequality, we obtain $3450 \ge m(m^2 + 5)$ implying $m \le 15$. For the near O-array we have $\underline{\mu}' = (8,9,9,9,9,9,8)$. Using (2.3) with $m = 13$, we get L.H.S. = 6.511871E+24, R.H.S. = 6.758249E+24 which is a contradiction. Thus $m \le 12$. The corresponding O-array obtained from the near O-array by adding a vector of all zeros and a vector of all 1's is such that $m \le 12$. Therefore, our bound is better.

| *Value of $\mu$* | *Rao's Bound* | *Bound from* (2.3) |
|---|---|---|
| 2 | $m \leq 8$ | $m \leq 8$ |
| 3 | $m \leq 10$ | $m \leq 9$ |
| 4 | $m \leq 11$ | $m \leq 9$ |
| 5 | $m \leq 12$ | $m \leq 10$ |
| 6 | $m \leq 13$ | $m \leq 11$ |
| 7 | $m \leq 13$ | $m \leq 11$ |
| 8 | $m \leq 14$ | $m \leq 12$ |
| 9 | $m \leq 15$ | $m \leq 12$ |
| 10 | $m \leq 15$ | $m \leq 13$ |
| 11 | $m \leq 16$ | $m \leq 13$ |
| 12 | $m \leq 16$ | $m \leq 14$ |
| 13 | $m \leq 16$ | $m \leq 14$ |
| 14 | $m \leq 17$ | $m \leq 14$ |
| 15 | $m \leq 17$ | $m \leq 15$ |
| 16 | $m \leq 18$ | $m \leq 15$ |
| 17 | $m \leq 18$ | $m \leq 15$ |
| 18 | $m \leq 18$ | $m \leq 16$ |
| 19 | $m \leq 19$ | $m \leq 16$ |
| 20 | $m \leq 19$ | $m \leq 16$ |

The above Table shows that we have improvement on the number of constraints in almost all cases.

**Remark:** For values of $\mu > 20$ we conjecture that our bounds are likely to be more precise than those of C. R. Rao. For example, for $\mu = 80$, our bound is $m \leq 25$ and the Rao bound is $m \leq 31$.

**Acknowledgements.** The authors would like to thank the referee for making suggestions which resulted in the improvement of this paper.

Thanks are also due to Gregory A. Kriegsmann and Daljit S. Ahluwalia for providing facilities during the preparation of this paper.

## References

[1] R.C. Bose and K.A. Bush, Orthogonal arrays of strength two and three, *Ann. Math. Statist.* **23** (1952), 508–524.

[2] K.A. Bush, Orthogonal arrays of index unity, *Ann. Math. Statist.* **23** (1952), 426–434.

[3] M.C. Chakrabarti, A note on skewness and kurtosis, *Bull. Calcutta Math. Soc.* **38** (1946), 133–136.

[4] C.S. Cheng, Orthogonal arrays with variable number of symbols, *Ann. Statist.* **8** (1980), 447–453.

[5] C.S. Cheng, Optimality of some weighing and $2^n$ fractional factorial designs, *Ann. Statist.* **8** (1980), 436–446.

[6] D.V. Chopra, On the maximum number of constraints for some orthogonal arrays of strength four, *Congressus Numerantium* **XV** (1975), 113–123.

[7] D.V. Chopra, Further investigations on balanced arrays, *J. Computational Statistics and Data Analysis* **12** (1991), 231–237.

[8] D.V. Chopra and R. Dios, On two levels balanced arrays of strength six, *J. Combin. Math. Combin. Comput.* **24** (1997), 249–253.

[9] M. Lakshmanamurti, On the upper bound of $\sum_{i=1}^{n} x_i^m$ subject to the conditions $\sum_{i=1}^{n} x_i = 0$ and $\sum_{i=1}^{n} x_i^2 = n$, *Math. Student* **18** (1950), 111–116.

[10] D.S. Mitrinovic, *Analytic Inequalities*, Springer-Verlag, New York, 1970.

[11] C.R. Rao, Hypercubes of strength 'd' leading to confounded designs in factorial experiments, *Bull. Calcutta Math. Soc.* **38** (1946), 67–78.

[12] C.R. Rao, Factorial experiments derivable from combinatorial arrangements of arrays, *J. Roy. Statist. Soc. Suppl.* **9** (1947), 128–139.

[13] C.R. Rao, Some combinatorial problems of arrays and applications to design of experiments, *A Survey of Combinatorial Theory* (ed. J.N. Srivastava, et. al.), North-Holland Publishing Co., 1973, 349–359.

[14] G.M. Saha, R. Mukerjee and S. Kageyama, Bounds on the number of constraints for balanced arrays of strength $t$, *J. Statist. Plann. Inf.* **18** (1988), 255–265.

[15] E. Seiden, On the problem of construction of orthogonal arrays, *Ann. Math. Statist.* **25** (1954), 151–156.

[16] E. Seiden and R. Zemach, On orthogonal arrays, *Ann. Math. Statist.* **27** (1966), 1355–1370.

[17] J.N. Srivastava and D.V. Chopra, Balanced arrays and orthogonal arrays, *A Survey of Combinatorial Theory* (ed. J.N. Srivastava, et. al.), North-Holland Publishing Co., 1973, 349–359.

[18] S. Yamamoto, Y. Fujii and M. Mitsuoka, Three-symbol orthogonal arrays of strength 2 and index 2 having maximal constraints: computational Study, *Jr. Comb. Inf. and Syst. Sci.* **18** (1993), 209–215.

[19] S. Yamamoto, M. Kuwada and R. Yuan, On the maximum number of constraints for s-symbol balanced arrays of strength $t$, *Commun. Statist. Theory Meth.* 14 (1985), 2447–2456.