## A SET-THEORETIC APPROACH TO ORGANIZATIONAL CONFIGURATIONS

Instead of using interaction effects, clustering algorithms, or deviation scores, a set-theoretic approach uses Boolean algebra to determine which combinations of organizational characteristics combine to result in the outcome in question (Boswell & Brown, 1999; Ragin, 1987, 2000). At the center of set-theoretic approaches lies the idea that relationships among different variables are often best understood in terms of *set membership*. Consider the simple case that $A$ is a member of the set $Z$ (formally: $A \subset Z$, or $A$ is a subset of $Z$). For purposes of analyzing organizational configurations, let $A$ be a firm with an efficient production system and $Z$ be the set of firms with high financial performance. Thus, the statement that firms with an efficient production system tend to exhibit high performance may be restated as such firms form a subset of high-performing firms.

At the same time, the overlap between both sets need not be absolute. For example, consider $B$, the set of firms with a high rate of product innovation. This characteristic may also result in high financial performance, thus making firms that rapidly innovate another subset of high-performing firms (formally: $B \subset Z$). Yet there may, in fact, be little overlap between the two subsets $A$ and $B$; one can easily imagine a situation where an efficient production system and a high rate of product innovation may inhibit or even preclude each other, thus making both $A$ and $B$ nonoverlapping subsets of $Z$. This may be expressed in the following logical statement:

$$A + B \rightarrow Z \qquad (1)$$

where "+" denotes the logical operator or while "$\rightarrow$" denotes the logical implication operator, as in "$A$ or $B$ implies $Z$." Both $A$ and $B$ thus present viable ways of attaining high financial performance, yet the design features involved in attaining that outcome may be quite different.

Now consider a somewhat more contingent statement: firms that exhibit an efficient production system (*A*) will be high performing *if* their environments are not heterogeneous (~C). In logical terms, this statement may be expressed as follows:

$$A \cdot {\sim}C \rightarrow Z \tag{2}$$

where " · " denotes the logical operator *and* while "~" denotes the logical *not*. In essence, the above statement presents a set-theoretic reformulation of a classic contingency hypothesis. Now let us extend the above by introducing another statement—namely, that firms with a high rate of product innovation (*B*) will be high performing if they also exhibit hierarchical control structures (*D*).[1] Combining this statement with Statement 2 from above results in the following statement:

$$A \cdot {\sim}C + B \cdot D \rightarrow Z \tag{3}$$

The Boolean statement above thus elegantly summarizes two contingency statements (or hypotheses) about the relationship of organizational characteristics, the nature of the environment, and firm performance.

To further understand a set-theoretic approach, let us now consider in more detail the nature of the set-subset relationships. Such relationships may be better understood in terms of necessity and sufficiency (Ragin, 1987), which describe the ability to generalize from a limited set of cases to larger populations. Consider again Statement 3. According to this statement, there are at least two combinations of attributes that may allow a firm to attain high performance. If, on the one hand, we take a necessary condition to denote that an outcome can be attained only if the attribute in question is present, then clearly neither of the combinations is necessary. On the other hand, if we take a sufficient condition to denote that an outcome will always be obtained if the attribute in question is present, then either of the combinations is sufficient. However, note that this finding applies only to combinations of attributes, not to individual attributes. In fact, of the individual attributes *A*, *B*, ~C, and *D*, none is either necessary or sufficient in that no attribute is present in all combinations and no attribute can by itself produce the outcome. In other words, Statement 3 denotes a situation of considerable causal complexity: four attributes combine to create the outcome, but none is by itself necessary or sufficient. Note also that such situations of causal complexity are exceedingly difficult to capture using conventional linear regression, since necessity and sufficiency are outside the focus of correlational analysis.

To analyze which different configurations of organizational characteristics may cause a certain outcome, a researcher using a set-theoretic approach first constructs a truth table that lists all possible configurations of characteristics, as well as whether these configurations lead to the outcome in question. In this regard, selection of the characteristics deemed important should be based on theoretical and substantive knowledge about their relationship with the outcome. In a second step the researcher uses Boolean logic to determine commonalities among the configurations that lead to the outcome and to generate logical statements such as those above that describe these commonalities, thus allowing for the logical reduction of statements. This reduction procedure uses the Quine-McCluskey algorithm, a common algorithm for simplifying set-theoretic statements that is implemented in software packages such as QCA (Drass & Ragin, 1992) and fs/QCA (Drass & Ragin, 1999). To illustrate how this algorithm works, consider again the relatively simple situation of causal complexity described by Statement 3. The corresponding truth table for such a situation is shown in Table 1. In this table shaded cells for characteristics indicate cells corresponding to Statement 3. Furthermore, some of the cells in the outcome column show a question mark, indicating that these combinations of conditions may show no empirical instances, a situation frequently observed in empirical research and usually referred to as a situation of limited diversity (Ragin, 1987, 2000).

To find out whether any of the four conditions is necessary for causing the outcome, we would examine whether the condition is always present in all cases where the outcome is achieved. Clearly, this is not the case here. However, the truth table shows that there are seven

[1] For the moment, I will not consider the empirical truth of these examples but merely use them to illustrate set-theoretic relationships.

**TABLE 1**
**Truth Table for Hypothetical Combinations of Organizational Characteristics**

| | Organizational Characteristics | | | | Outcome |
|---|---|---|---|---|---|
| | *A* | *B* | *C* | *D* | *Z* |
| Configuration Number | Efficient Production System | High Rate of Product Innovation | Heterogeneous Environment | Hierarchical Control Structure | High Performance |
| 1 | Yes | Yes | Yes | Yes | Yes |
| 2 | Yes | Yes | Yes | No | No |
| 3 | Yes | Yes | No | Yes | Yes |
| 4 | Yes | Yes | No | No | Yes |
| 5 | Yes | No | Yes | Yes | No |
| 6 | Yes | No | Yes | No | No |
| 7 | Yes | No | No | Yes | Yes |
| 8 | Yes | No | No | No | Yes |
| 9 | No | Yes | Yes | Yes | Yes |
| 10 | No | Yes | Yes | No | No |
| 11 | No | Yes | No | Yes | Yes |
| 12 | No | Yes | No | No | No |
| 13 | No | No | Yes | Yes | ? |
| 14 | No | No | Yes | No | No |
| 15 | No | No | No | Yes | ? |
| 16 | No | No | No | No | ? |

different configurations of the four individual organizational characteristics sufficient for causing the outcome. The combinations are listed below:

1. $A \cdot B \cdot C \cdot D$
3. $A \cdot B \cdot {\sim}C \cdot D$
4. $A \cdot B \cdot {\sim}C \cdot {\sim}D$
7. $A \cdot {\sim}B \cdot {\sim}C \cdot D$
8. $A \cdot {\sim}B \cdot {\sim}C \cdot {\sim}D$
9. ${\sim}A \cdot B \cdot C \cdot D$
11. ${\sim}A \cdot B \cdot {\sim}C \cdot D$

While these combinations are all sufficient for causing high performance, the seven combinations can be simplified, since some combinations are logically redundant. For example, firms with an efficient production system ($A$) that are not in heterogeneous environments (${\sim}C$) may or may not have a high rate of product innovation ($B$ or ${\sim}B$) and may or may not exhibit a hierarchical control structure ($D$ or ${\sim}D$). Either way, the combination of $A$ and ${\sim}C$ will still be sufficient to cause the outcome. As a result, the seven combinations may be logically reduced and simplified using the Quine-McCluskey algorithm and simplifying assumptions (cf. Ragin, 1987, 2000). In Boolean algebra, this proceeds as follows for Combinations 3 and 4:

$$A \cdot B \cdot {\sim}C \cdot {\sim}D + A \cdot B \cdot {\sim}C \cdot D =$$

$$A \cdot B \cdot {\sim}C(D + {\sim}D) =$$

$$A \cdot B \cdot {\sim}C$$

Similarly, Combinations 7 and 8 can also be simplified:

$$A \cdot {\sim}B \cdot {\sim}C \cdot D + A \cdot {\sim}B \cdot {\sim}C \cdot {\sim}D =$$

$$A \cdot {\sim}B \cdot {\sim}C(D + {\sim}D) =$$

$$A \cdot {\sim}B \cdot {\sim}C$$

Finally, combining the results from both simplifications leads to the following:

$$A \cdot B \cdot {\sim}C + A \cdot {\sim}B \cdot {\sim}C =$$

$$A \cdot {\sim}C(B + {\sim}B) =$$

$$A \cdot {\sim}C$$

Using some very simple operations, we have thus arrived at a statement that, by itself, contains all four logical combinations involving $A$ and ${\sim}C$ (3, 4, 7, and 8) that may lead to the outcome in question. The same operations can, of course, be applied to the four logical combinations involving $B$ and $D$ (1, 3, 9, and 11) that are also sufficient for producing high performance. The result is again a simple statement that contains all combinations that may cause the outcome:

$$A \cdot {\sim}C + B \cdot D \rightarrow Z \quad (4)$$

From Crisp to Fuzzy Sets