Selecting on the Dependent Variable
or
Analyzing Qualitative Data on "Multiple Instances"

Ethnographers and other qualitative researchers often have "multiple instances," meaning they have many cases of "the same thing." A researcher who conducts in-depth, open-ended interviews with 20 Olympic athletes in order to find out how they maintain their commitment has 20 instances of commitment to study. How should qualitative researchers generalize across instances?

One simple way to generalize across cases is to see if they all display the same ingredients for the outcome. These ingredients can be viewed as **causally relevant** or as **constitutive**. If they are viewed as constitutive, the outcome, in effect, is inherent or embedded in the conditions.

For example, the researcher studying how 20 Olympic athletes maintain their commitment might list some common ingredients, based on his/her field work: (1) devotion to a daily exercise routine, (2) a feeling of separateness from (or superiority to) non-athletes, (3) development of pre- or post work-out rituals, (4) food preferences and practices that make eating with others (especially non-athletes) problematic, and (5) associating primarily with others athletes. Do the 20 Olympic-level athletes share these features?

Case #	devotion to exercise	feeling of separateness	workout rituals	assoc. w/ athletes	separate food
1	yes	yes	yes	yes	yes
2	yes	no	yes	yes	yes
3	yes	yes	yes	no	no
4	yes	yes	yes	no	yes
5	yes	yes	yes	no	no
6	yes	yes	no	yes	yes
7	yes	no	no	yes	yes
8	yes	no	yes	yes	no
9	yes	no	yes	yes	yes
10	yes	yes	no	no	yes
11	yes	yes	yes	no	yes
12	yes	yes	no	yes	yes
13	yes	no	yes	yes	yes
14	yes	no	yes	yes	yes
15	yes	yes	yes	no	yes
16	yes	yes	no	yes	yes
17	yes	yes	yes	no	yes
18	yes	yes	yes	yes	yes
19	yes	yes	no	yes	yes
20	yes	no	yes	yes	no

Analyzing the Cross-Case Evidence

There are (at least) three ways to analyze this evidence.

1**. Analytic Induction**: The conditions must be shared uniformly, without exceptions. This is the technique developed by Alfred Lindesmith and used in his study *Opiate Addiction*. It is strict and rigorous. According to the table (which is hypothetical), the only thing they all share is "devotion to exercise." So in very strict terms, there is not a whole lot of agreement across the 20 cases and thus not very much to talk about.

2. **Composite Portrait**. What characteristics do “most” cases share? Notice that 13/20 have a feeling of separateness; 14/20 practice workout rituals; 13/20 associate primarily with other athletes; and 16/20 have distinctive food preferences or habits. The "composite portrait" has all five features, therefore, because all five features are more commonly present than absent across all 20 cases. This "composite portrait" can be reinforced simply by counting the number of "yes's" each case displays. Two cases have five yes's; twelve cases have four yes's; and six cases have three yes's. No case has less than 3 out of the 5 traits, which means even the "worst" fitting case has 3 of the 5 traits. It is possible to apply probabilistic criteria to these assessments, but it is rarely done.

3. **Substitutable Conditions.** It is also possible to look at the evidence in terms of "substitutability." The basic idea here is that there is more than one way to satisfy a general condition and that some of these feature may be alternate ways of satisfying the same condition. Notice, for example, that whenever there is a "no" in the "feeling of separateness" column there is a "yes" in the "associates mostly with other athletes" column. The reverse is also true: whenever there is a "no" in the "associates with athletes" column, there is a "yes" in the "feeling of separateness" column. If you think of these two things as "equivalent" or substitutable, then every Olympic athlete displays EITHER a feeling of separateness OR a preference for associating with other athletes. Notice that this same general relation exists between "workout rituals" and "separate foods." Whenever there is a "no" in one column, there is a "yes" in the other. Thus, these two features can also be seen as equivalent or substitutable. The general picture that emerges from this analysis is that the shared features are not one (devotion to exercise) but three:

1. Devotion to a daily exercise routine
2. EITHER feeling of separateness OR associates primarily with other athletes
3. EITHER performs workout rituals OR eats separate food

Substitutability and Concept Elaboration

Generally, when substitutability is identified, it goes hand-in-hand with concept development or elaboration. What do the pairs of substitutable "ingredients" represent at a more abstract level?

The second commonalities (feeling of separateness or associates primarily with other athletes) can be described at a more general level as the construction of a *social boundary* between athletes and non-athletes, either through patterns of association or through feelings related to personal identity.

The third commonality (workout rituals or separate food) could be seen as *daily practices that reinforce that identity*.

The three ingredients for maintaining commitment as an Olympic athlete, therefore, could be described as:

1. Devotion to a daily exercise routine
2. Construction of a boundary separating athletes from non-athletes
3. Everyday practices that reinforce identity as an athlete

Substitutability and QCA

It is possible to analyze substitutability using truth table analysis. More generally, it is possible to examine evidence with a constant as the dependent variable, which is an extreme case of selecting on the dependent variable. The steps are:

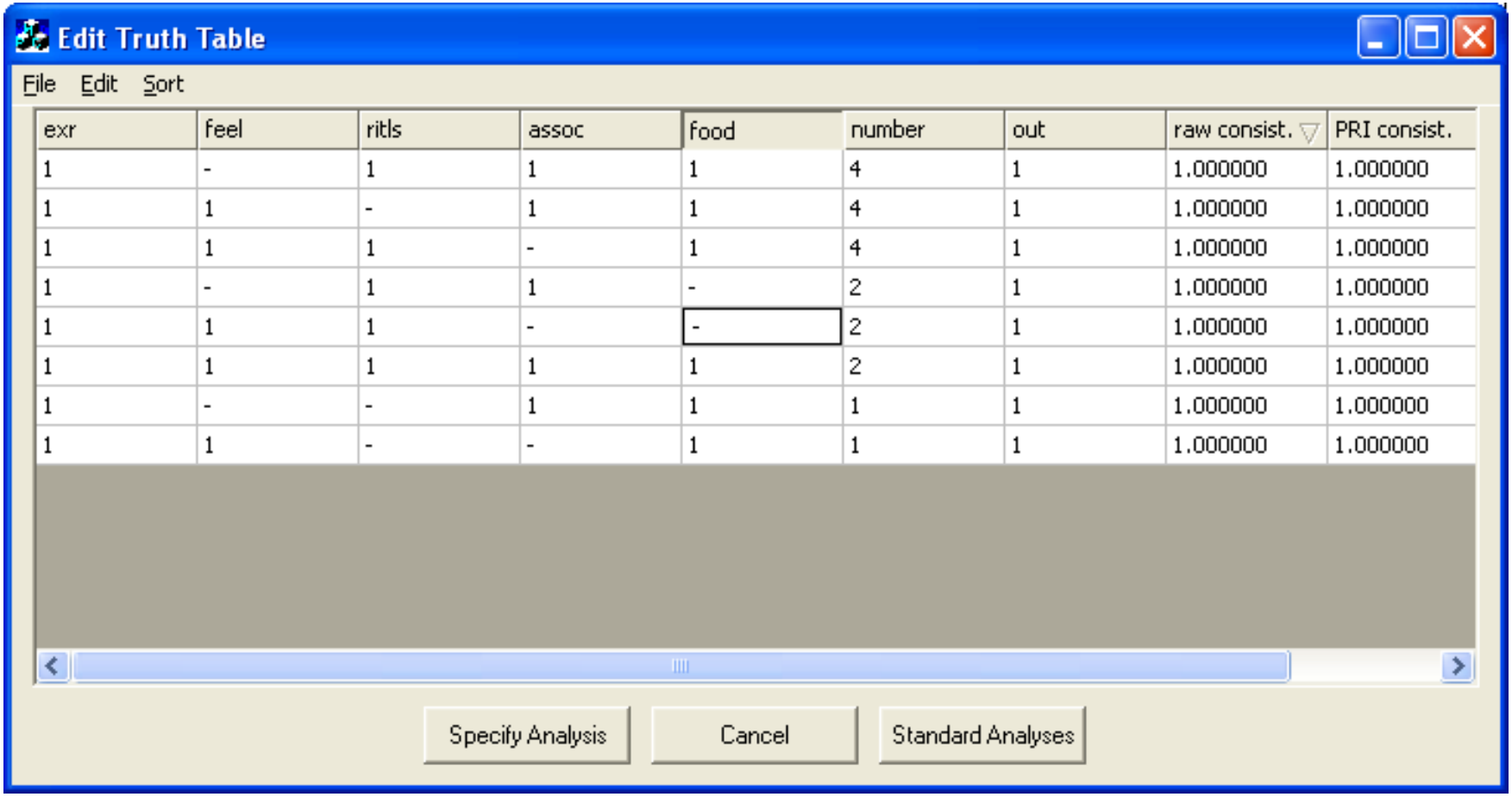
(1) Create an outcome variable that is coded "1" for every case (e.g., 1 = “maintains commitment” in the Olympic athletes example),

(2) Change all the 0's to dashes (according to our theoretical and substantive knowledge, if an individual has a "no" on one trait, changing it to "yes" would only bring about a stronger case for commitment), and

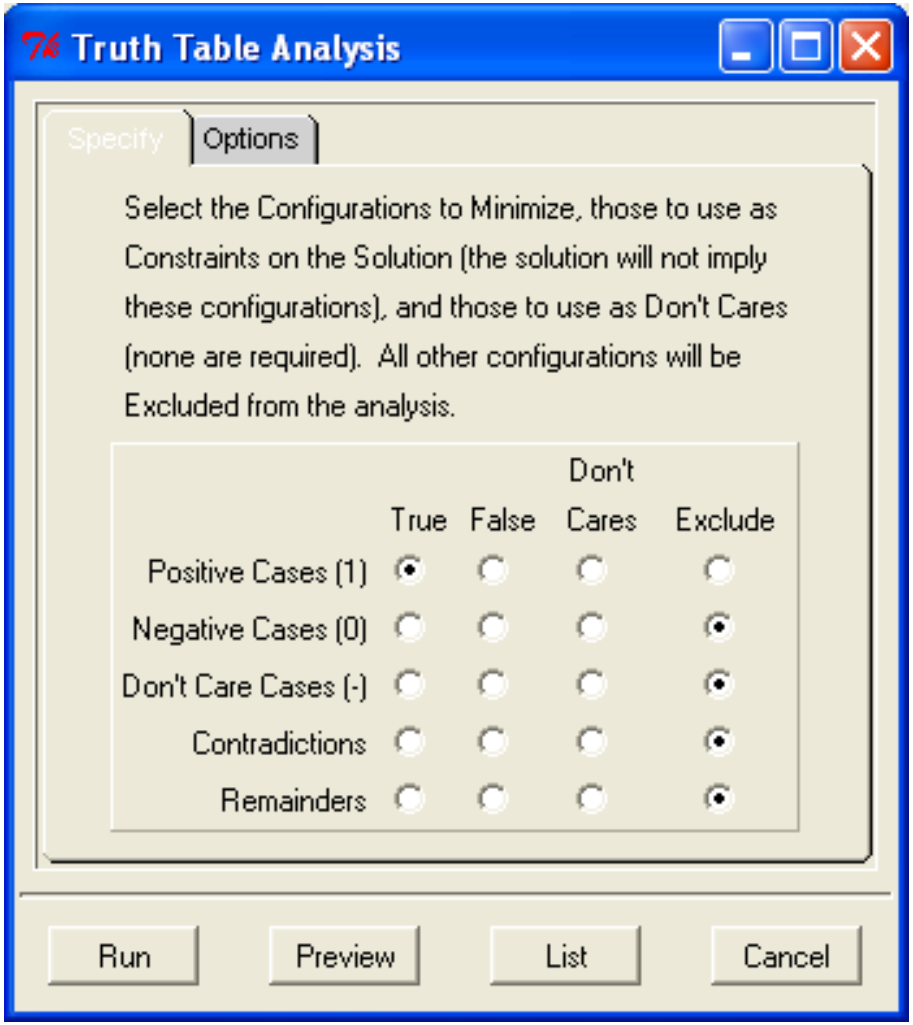
(3) Analyze the resulting truth table, setting the "remainders" to false.

How to Set Remainders to "False"

After coding the truth table spreadsheet in the manner just described (1 and 2 on the previous slide), click the "Specify Analysis" button below the truth table spreadsheet:



The following dialogue box will open:



Make sure that “Positive Cases (1)” is clicked “True” and all the others are clicked “Exclude”. Then click “Run”. The solution will be reported in the output window.

Here's the data set:

case	exr	feel	ritls	assoc	food	outcome
1	1	1	1	1	1	1
2	1	–	1	1	1	1
3	1	1	1	–	–	1
4	1	1	1	–	1	1
5	1	1	1	–	–	1
6	1	1	–	1	1	1
7	1	–	–	1	1	1
8	1	–	1	1	–	1
9	1	–	1	1	1	1
10	1	1	–	–	1	1
11	1	1	1	–	1	1
12	1	1	–	1	1	1
13	1	–	1	1	1	1
14	1	–	1	1	1	1
15	1	1	1	–	1	1
16	1	1	–	1	1	1
17	1	1	1	–	1	1
18	1	1	1	1	1	1
19	1	1	–	1	1	1
20	1	–	1	1	–	1

Here are the results:

EXERCISE FEEL RITUALS +
EXERCISE RITUALS ASSOC +
EXERCISE FEEL FOOD +
EXERCISE ASSOC FOOD

which can be factored to show the three conditions:

EXERCISE * (FEEL + ASSOC) * (FOOD + RITUALS)

Fuzzy Set Analysis of "Only Positive Instances"

Is it also possible to do a fuzzy-set analysis when all cases have full membership in the outcome. For illustration, let's take the same data set and fuzzify the scores. The 1's have been converted to numbers between .6 and 1.0; the 0's have been converted to numbers between 0 and .4. The outcome is coded 1 for each case.

Here's the data set:

case	exr	feel	ritls	assoc	food	out
1	.6	.6	.9	.7	.6	1
2	.7	.1	.8	.8	.7	1
3	.8	.7	.7	.2	.3	1
4	.9	.9	.6	.4	.8	1
4	.6	.8	.7	.1	.4	1
6	.7	.6	.1	.8	.9	1
7	.8	.2	.3	.9	.9	1
8	.9	.4	.8	.6	.1	1
9	.6	.3	.9	.7	.8	1
10	.7	.9	.4	.2	.7	1
11	.8	.8	.6	.3	.7	1
12	.9	.7	.1	.8	.6	1
13	.6	.2	.7	.9	.6	1

14	.7	.1	.8	.6	.9	1
15	.8	.9	.9	.3	.9	1
16	.9	.8	.3	.7	.6	1
17	.6	.7	.8	.4	.7	1
18	.7	.6	.6	.8	.8	1
19	.8	.7	.2	.9	.9	1
20	.9	.1	.7	.6	.4	1

Here is the truth table spreadsheet, using the fuzzy-set/truth-table algorithm:

exr	feel	ritls	assoc	food	number	out	consist
1	0	1	1	1	4	1	1
1	1	0	1	1	4	1	1
1	1	1	0	1	4	1	1
1	0	1	1	0	2	1	1
1	1	1	0	0	2	1	1
1	1	1	1	1	2	1	1
1	0	0	1	1	1	1	1
1	1	0	0	1	1	1	1

This truth table spreadsheet can be recoded to resemble the pattern in the crisp analysis:

exr	feel	ritls	assoc	food	number	out	consist
1	-	1	1	1	4	1	1
1	1	-	1	1	4	1	1
1	1	1	-	1	4	1	1
1	-	1	1	-	2	1	1
1	1	1	-	-	2	1	1
1	1	1	1	1	2	1	1
1	-	-	1	1	1	1	1
1	1	-	-	1	1	1	1

And here are the results using the same “Specify Analysis” procedure used with crisp sets. The fuzzy set results are identical to the crisp-set results:

	raw coverage	unique coverage	consistency
EXR*FEEL*RITLS+	0.384500	0.025500	1.000000
EXR*RITLS*ASSOC+	0.390000	0.011000	1.000000
EXR*FEEL*FOOD+	0.459500	0.030000	1.000000
EXR*ASSOC*FOOD	0.489500	0.040000	1.000000

solution coverage: 0.670000
solution consistency: 1.000000