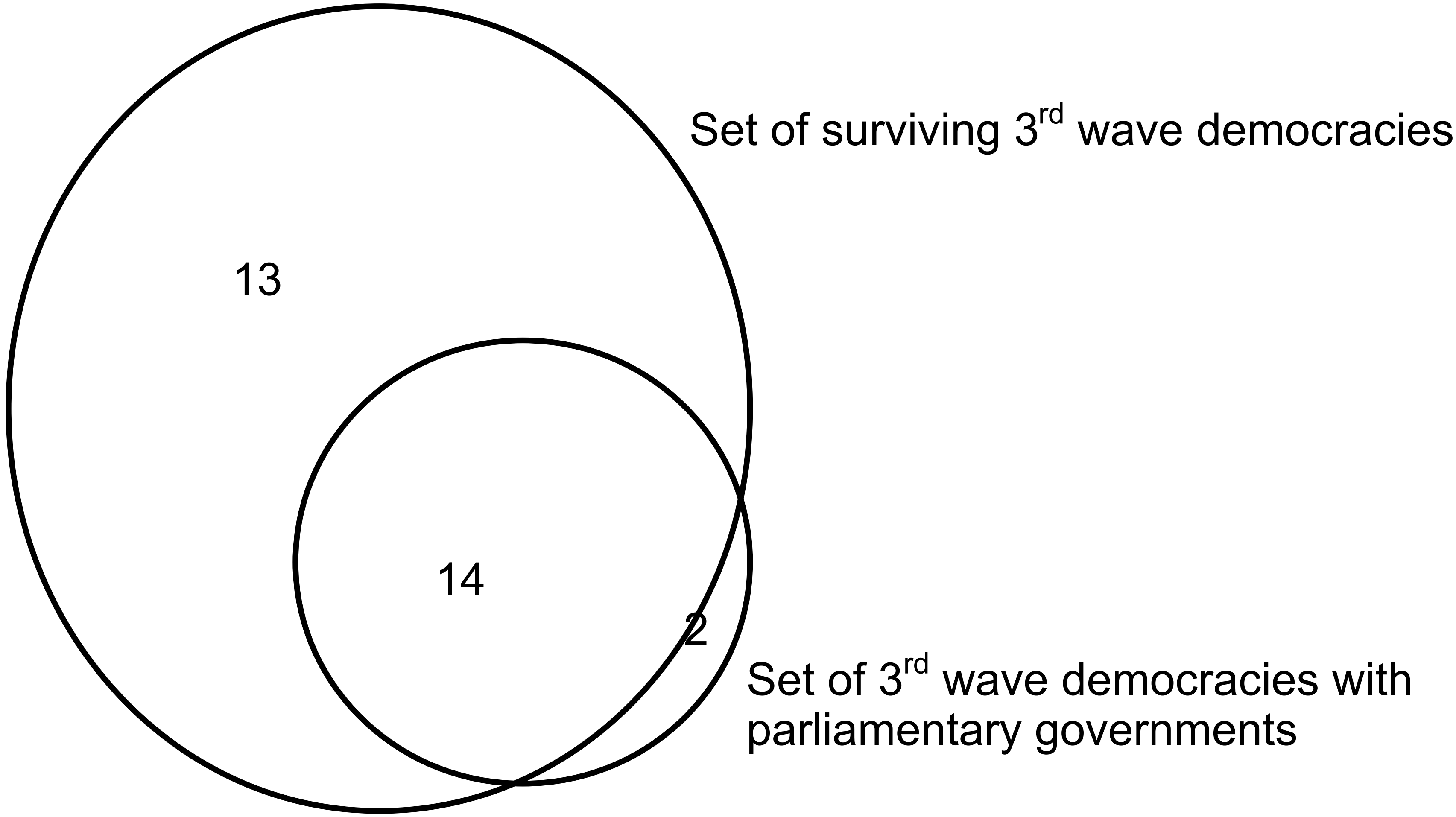
The Degree of Consistency of a Subset Relation

The consistency of a crisp subset relation is straightforward. It is simply the number of cases displaying both the cause and the outcome (cell 2) divided by the number of cases displaying the causal condition (the sum of cells 2 and 4). For example, in the following table the consistency of the evidence with the argument that having a parliamentary form of government is sufficient for the survival of a 3rd wave democracy is:

$$14/(2 + 14) = 14/16 = 0.875$$

	Presidential form	Parliamentary form
3rd wave democracy survived	13	14
3rd wave democracy collapsed	11	2

Only cases in the second column are involved in the assessment of consistency of the evidence with the subset relation. The following is a Venn diagram illustrating this high but less-than-perfect degree of consistency.



Consistency of subset relation = 14/(2+14) = 14/16 = 0.875

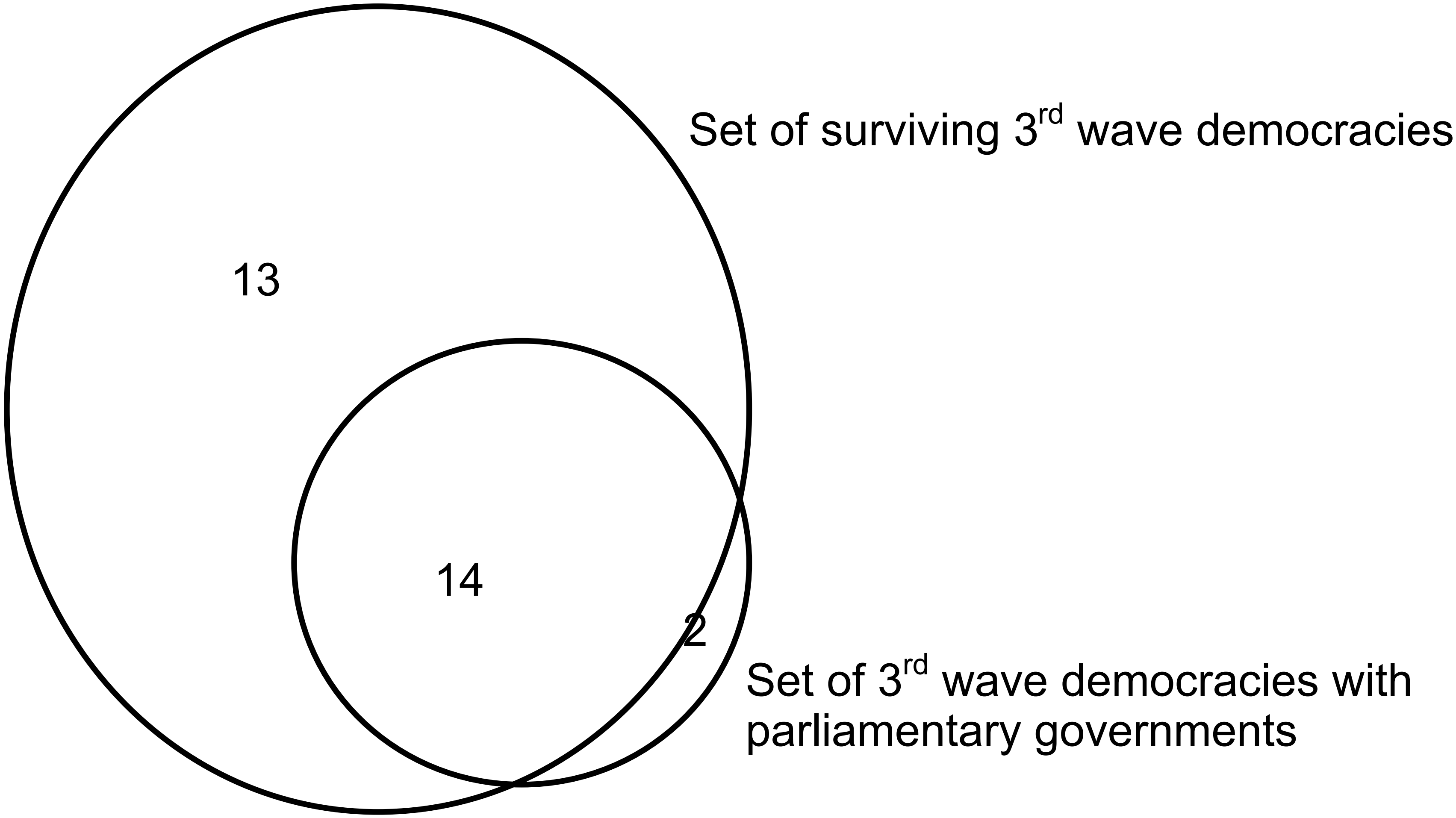
The Degree of Coverage of a Subset Relation

Once it has been established that a subset is at least roughly consistent, it is possible to assess its degree of coverage. The calculation of coverage for crisp sets is straightforward. It is simply the number of cases with both the causal condition and the outcome (again, cell 2) divided by the number of cases with the outcome (cell 1 + cell 2). In other words, coverage answers the question: What proportion of cases with the outcome has been “explained”? or How common is the cause (or causal combination) among the cases with the outcome?

$$14/(13 + 14) = 14/27 = 0.519$$

	Presidential form	Parliamentary form
3^{rd} wave democracy survived	13	14
3^{rd} wave democracy collapsed	11	2

Only cases in the first row are involved in the assessment of coverage. The previous Venn diagram also illustrates the concept of coverage. The key consideration is the proportion of the larger set covered by the interior set.



Coverage of outcome by condition = 14/(13 + 14) = 14/47 = 0.519

CONSISTENCY VERSUS COVERAGE USING CRISP SETS

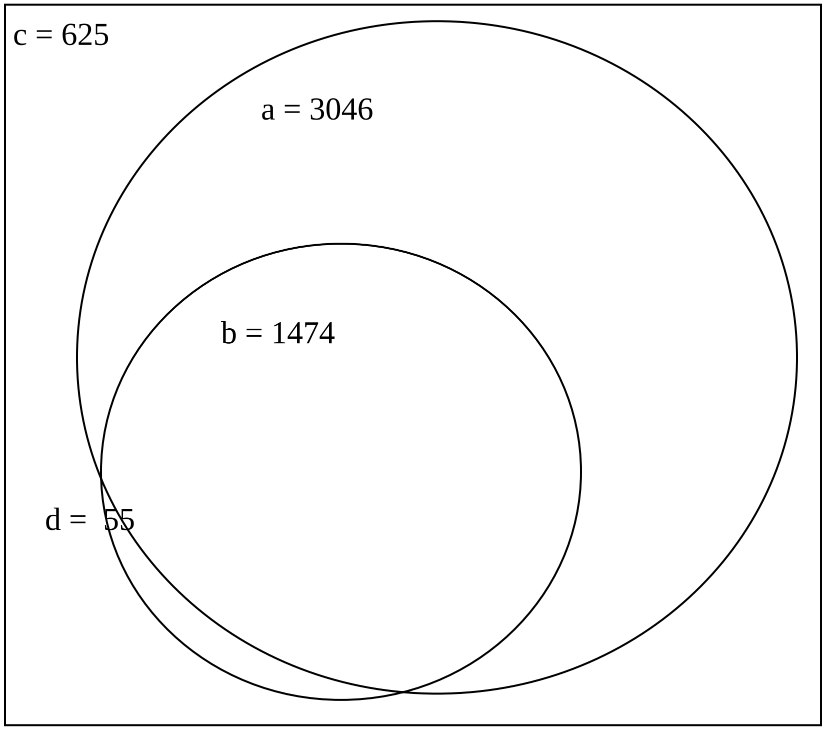
Let's start with a simple 2X2: the crosstabulation of poverty status and educational achievement

	Low/Average Educational Achievement	High Educational Achievement
Not In Poverty	a. 3046	b. 1474
In Poverty	c. 625	d. 55

Is high educational achievement sufficient for poverty avoidance? In other words, do those with high educational achievement constitute a consistent subset of those who have avoided poverty?

If so, how much of poverty avoidance is "covered" (or explained) by high educational achievement?

Reformulate the table as a Venn diagram:



Area a = Cases with low/average educational achievement, not in poverty

Area b = Cases with high educational achievement, not in poverty

Area c = Cases with low/average educational achievement, in poverty

Area d = Cases with high educational achievement, in poverty

Consistency (of subsethood) = b/(b+d) = 1474/(1474+55) = 0.958

Coverage of outcome by subset = b/(a+b) = 1474/(1474+ 3046) = 0.326

The Meaning of Coverage, with X a subset of Y

Crosstabulation of Poverty Status and Educational Achievement: Actual Frequencies

	Low/Average Educational Achievement	High Educational Achievement
Not In Poverty	a. 3046	b. 1474
In Poverty	c. 625	d. 55

Crosstabulation of Poverty Status and Educational Achievement: Altered Frequencies

	Low/Average Educational Achievement	High Educational Achievement
Not In Poverty	a. 4373	b. 147
In Poverty	c. 675	d. 5

<u>Consistency remains about the same</u>, but coverage declines dramatically once cases are shifted (arbitrarily) to the first column.

Typically, there is a trade-off between consistency and coverage; higher consistency scores usually mean lower coverage. More narrowly specified arguments may be more consistent, but apply to fewer cases. Virtually all graduates of Ivy League universities are able to avoid poverty (consistency), but they make up only a tiny fraction of people not in poverty (coverage).

"WHOLE EQUATION" CONSISTENCY

Suppose the results of a crisp-set QCA showed:

$Y \geq AB + CD$

Suppose the consistency of the two recipes considered separately reveals:

AB: 16/20 cases; consistency = 0.80
CD: 24/30 cases; consistency = 0.80

If there are cases of ABCD, which is likely, and they are all instances of the outcome (which is also likely), they may have been counted twice. Thus, the consistency of the whole equation could be less than the consistency of the individual recipes.

Suppose there are four cases of ABCD and they are all consistent; the *whole equation* consistency is:

AB(c+d): ABCD: CD(a + b):
(12 out of 16) + (4 out of 4) + (20 out of 26) = 36/46 = 0.78

"WHOLE EQUATION" COVERAGE, USING THE SAME DATA

Suppose the results of a crisp-set QCA show:

Y ≥ AB + CD

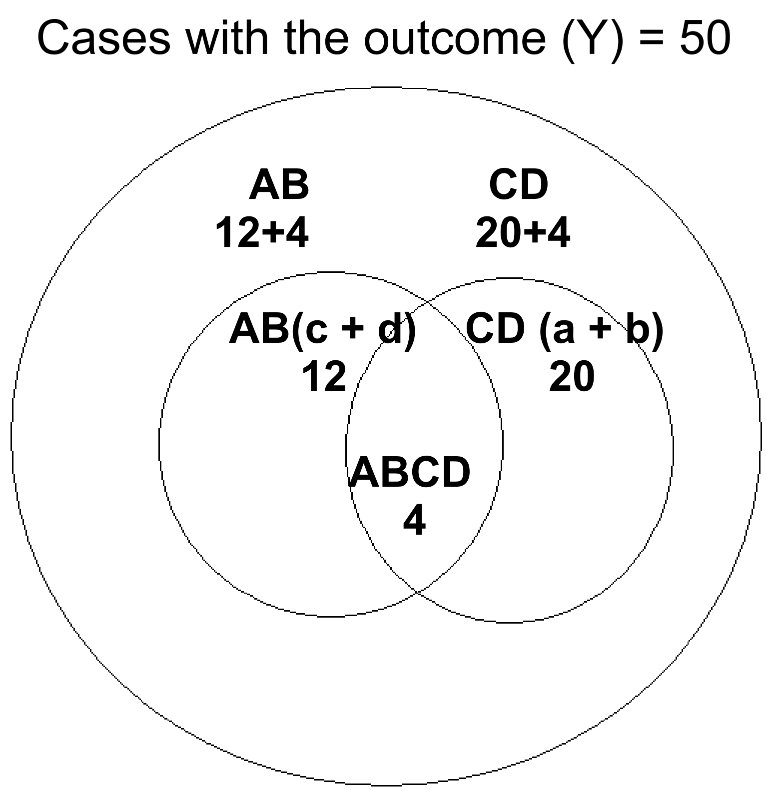
Suppose the data on raw coverage reveal:

AB: 16/50 cases; coverage = 0.32
CD: 24/50 cases; coverage = 0.48
(sum = 0.80)

Assume there are four cases of ABCD, all consistent; the *whole equation* coverage is:

AB(c+d):		**ABCD:**		**CD(a + b):**	
(12 out of 50)	+	(4 out of 50)	+	(20 out of 50)	= 36/50 = 0.72
0.24		0.08		0.40	

Partitioning Coverage ("Raw" Versus "Unique" Coverage)



AB raw coverage = 16/50 (0.32); unique coverage = 12/50 (0.24)

CD raw coverage = 24/50 (0.48); unique coverage = 20/50 (0.40)

CONSISTENCY AND COVERAGE USING FUZZY SETS

Degree to which X is a consistent subset of Y

Consistency ($\mathbf{X_i} \leq \mathbf{Y_i}$) = $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{X_i})$ *instances of X share outcome Y*

if consistent X is a consistent subset, calculate coverage:

Coverage ($\mathbf{X_i} \leq \mathbf{Y_i}$) = $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{Y_i})$ *prevalence of X among outcome Y*

Degree to which X is a consistent superset of Y

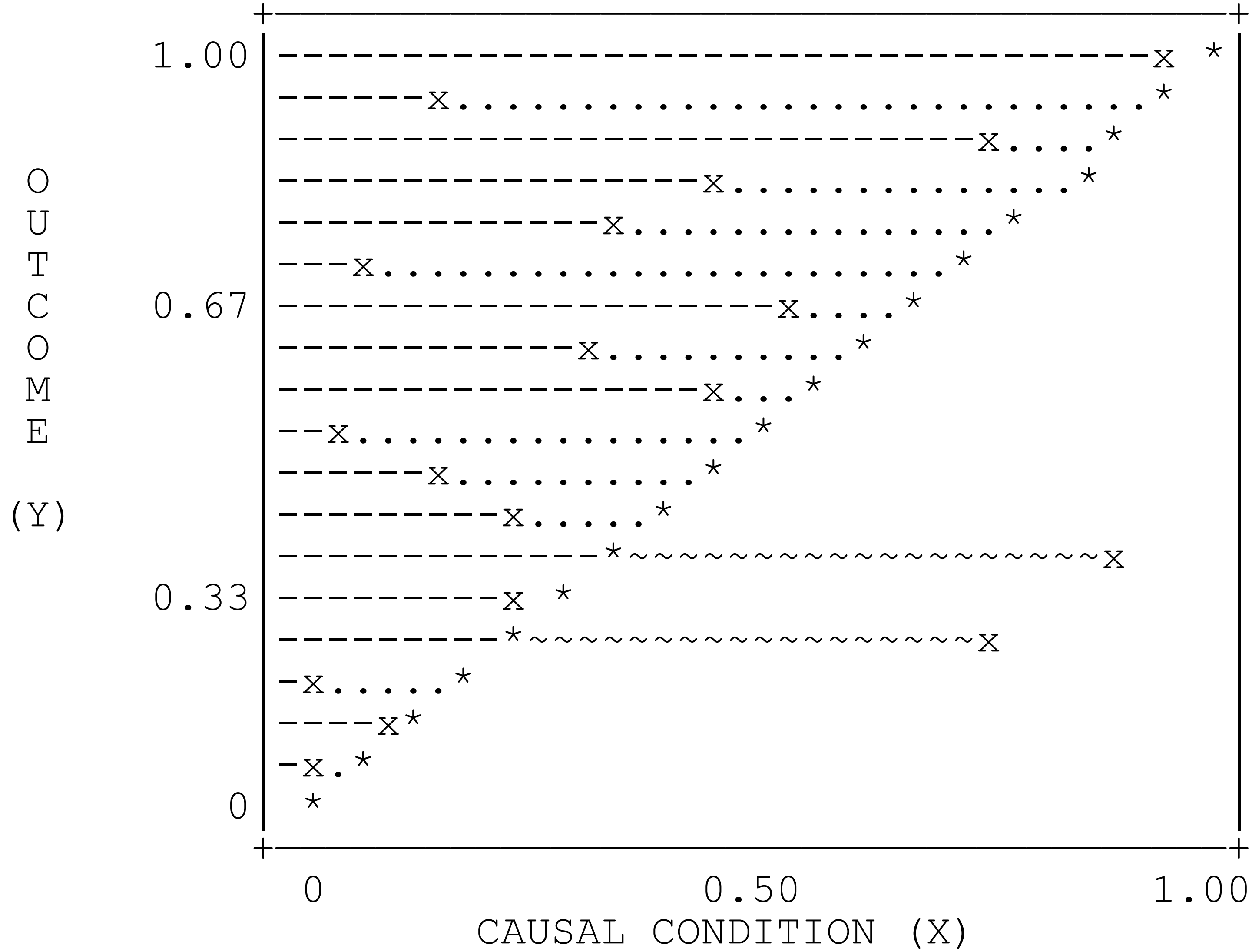
Consistency ($\mathbf{X_i} \geq \mathbf{Y_i}$) = $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{Y_i})$ *X is a shared antecedent for Y*

If X is a consistent superset, calculate coverage:

Coverage ($\mathbf{X_i} \geq \mathbf{Y_i}$) = $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{X_i})$ *relevance of X as an antecedent*

Notice that the formulas overlap, that is, the consistency of X as a subset of Y is the same as the coverage of Y as a subset of X, while the coverage of X as a subset of Y is the same as the consistency of Y as a subset of X.

Fuzzy Scatterplot Illustrating the Concepts of Consistency and Coverage



Dashes represent min(X,Y). Dots represent the portion of Y that is not covered by X. The tilde signs represent the portion of X that is inconsistent with the subset relation.

Protocol and Interpretation

Calculation of Consistency Comes First! The first task is to determine whether there is a set-theoretic relation between the cause or causal combination (X) and the outcome (Y). If X is a subset of Y, then X MAY BE sufficient for Y. If X is a superset of Y, then X MAY BE necessary for Y. Note that if the X scores indicate membership in a *combination* of conditions, then it is unlikely for Y to be a subset of X. (Recall that when conditions are combined, membership in the combination is determined by the minimum value of the component memberships).

The set-theoretic consistency calculations are:

(a) The degree to which X is a subset of Y: $\text{sum}(\min(X_i,Y_i))/\text{sum}(X_i)$

(b) The degree to which X is a superset of Y $\text{sum}(\min(X_i,Y_i))/\text{sum}(Y_i)$

These two calculations differ only in the denominator. If X scores are consistently less than Y scores, then the value of (a) will be 1.0 and the value of (b) will be substantially less than 1.0, perhaps 0.5 or even lower. If X scores are consistently greater than Y scores, then the value of (a) will be substantially less than 1.0, perhaps 0.5 or lower, and the value of (b) will be 1.0.

The closer the value of (a) or (b) is to 1.0, the stronger the case that a set-theoretic relation exists. If they are both close to 1.0, then X and Y scores are approximately equal across cases. Of course, both (a) and (b) may be substantially less than 1.0, indicating simply that neither X nor Y is a subset of the other (not even a rough subset). In general, (a) is the key calculation when working with combinations of conditions, because it is very rare for scores in a combination of conditions to be consistently greater than scores in the outcome.

Calculation of Coverage

Once it has been established that a set-theoretic relation exists, it is reasonable to assess its coverage or relevance. If X is a subset of Y, this assessment is the same asking how important is the combination of conditions represented by X in accounting for Y: How much of Y does X cover? If X is a superset of Y, then the calculation of coverage shows the relevance of X as a necessary condition. If Y is dwarfed by X, then X is not really providing a ceiling on the expression of Y. It just happens to be something that is strongly present whenever there is any membership in Y. On the other hand, if Y is a substantial subset of X, then X is more relevant as a necessary condition--X's limiting power on Y is more apparent.

Assume we have shown that X is a rough subset of Y, that is, calculation (a), shown previously, is 1.0 or close to 1.0. The calculation of coverage is simply:

$$\mathrm{sum}(\min(X_i,Y_i))/\mathrm{sum}(Y_i)$$

Notice that this is the same as calculation (b) above. It makes sense as a calculation of the coverage of Y by X ONLY IF it has been established that X is a subset of Y (or at least a rough subset).

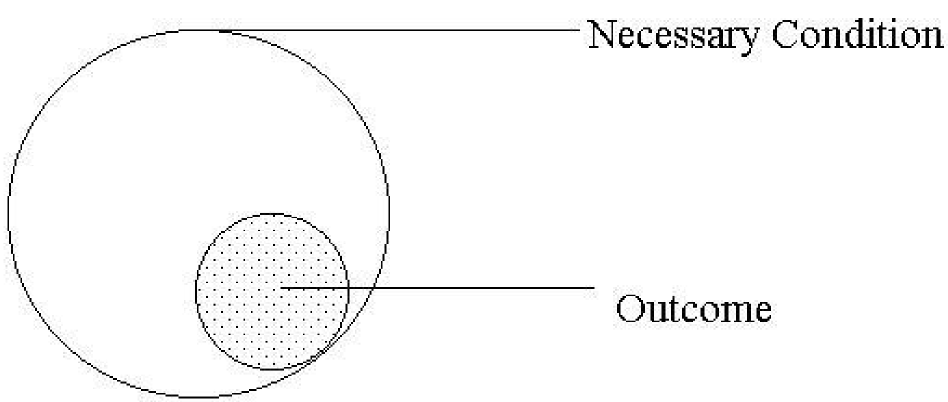
Assume we have shown that X is a rough superset Y, that is, calculation (b), shown previously, is 1.0 or close to 1.0. The calculation of coverage (i.e., the “relevance” of a necessary condition) is:

$$\mathrm{sum}(\min(X_i,Y_i))/\mathrm{sum}(X_i)$$

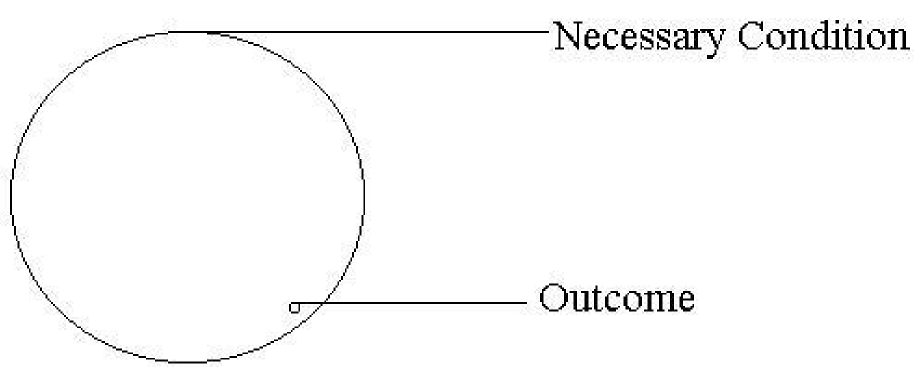
Notice that this is the same as calculation (a) above. It makes sense as a calculation of the coverage of X by Y (relevance of X as a necessary condition for Y) ONLY IF it has been established that X is a superset of Y (or at least a rough superset).

Figure 3.5: Venn diagram illustrating necessary conditions

a. Empirically Relevant Necessary Condition



b. Empirically Irrelevant Necessary Condition



Protocol for Assessing Consistency and Coverage

	Type of set-theoretic relation:	
Procedure:	*Cause X is a subset of outcome Y (instances of X share outcome Y)*	*Cause X is a superset of outcome Y (X is a shared antecedent condition)*
Step 1	Assess consistency using $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{X_i})$	Assess consistency using $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{Y_i})$
Step 2	If consistent, assess prevalence using $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{Y_i})$	If consistent, assess relevance using $\Sigma(\min(\mathbf{X_i},\mathbf{Y_i}))/\Sigma(\mathbf{X_i})$

Partitioning Coverage: Fuzzy Sets

As with crisp sets, it is possible to partition coverage to assess the degree of overlap of the causal combinations. With crisp sets, a case is either covered by a combination or its not covered. With fuzzy sets, by contrast, different combinations can cover a given case's outcome to different degrees. For example, membership in the outcome might equal 0.8; membership in causal combination A·B might be 0.6 and membership in causal combination C·D might be 0.4. Both combinations offer some coverage, though A·B's coverage is superior (i.e., closer to 0.8 without exceeding it).

Consider the analysis of the impact of high test scores (T), high parental income (I), and college education (C) on avoiding poverty (~P) for white males. The results show:

$$T \cdot I + I \cdot C \rightarrow {\sim}P$$

Here are some coverage calculations associated with these data:

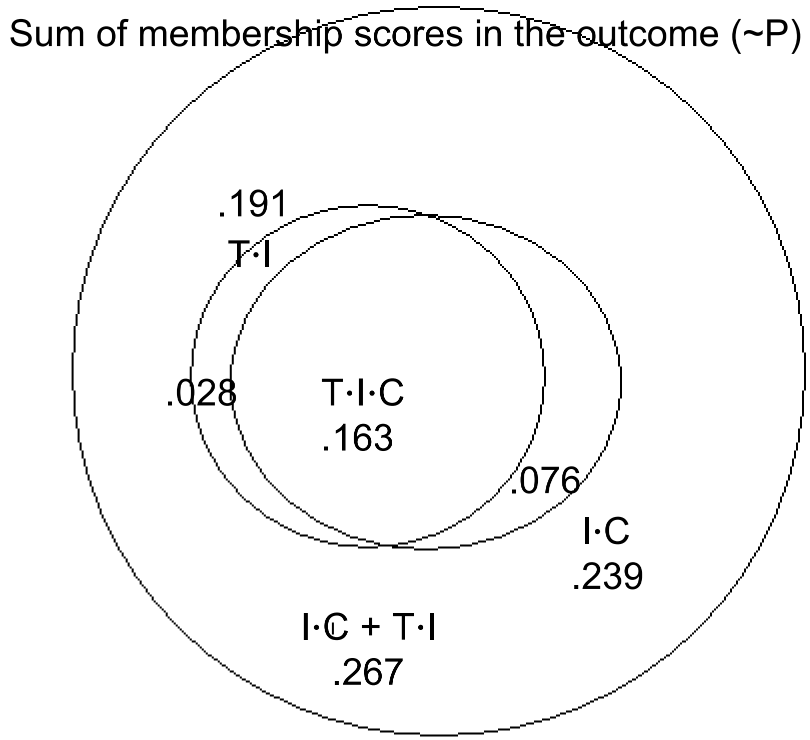
Causal Conditions	**Sum of Consistent Scores: $\Sigma(\min(X_i,Y_i))$**	**Sum of Outcome Scores: $\Sigma(Y_i)$**	**Coverage**
T·I	181.830	949.847	.191
I·C	226.792	949.847	.239
T·I + I·C	253.622	949.847	.267
T·I·C	155.000	949.847	.163

First, notes that membership in T·I + I·C is the maximum of (T·I, I·C). That is, when calculating membership in a fuzzy logic statement that includes logical *or* (+) it is necessary to take the maximum score. Thus, if your membership is 0.6 in T·I and 0.9 in I·C, then your membership in the union of these two causal recipes is 0.9.

Second, notice that on the bottom row I have included calculations for membership in the intersection of the two terms. This is included because the two paths overlap substantially. Most of the folks who are T·I also have C and most of the folks who have I·C also have T. (In other words, we live in a world of overlapping inequalities.)

Unique coverage of T·I is given by cov(T·I + I·C) - cov(I·C), which is 2.8%; unique coverage of I·C is given by cov(T·I + I·C) - cov(T·I), which is 7.6%. (These calculations are produced automatically by fsQCA.) Notice also that the coverage of T·I·C (16.3%) plus the two unique calculation (2.8% + 7.6%) is equal to the whole equation coverage (26.7%).

Illustration of Raw and Unique Coverage Using Fuzzy Sets



The numbers indicate proportions of the sum of the membership scores in the outcome that are “covered” by the different causal combinations.