GROUNDED THEORY REVIEW
An international journal

Theoretical Elaboration of Quantitative Data

Barney G. Glaser, Ph.D., Hon. Ph.D.

June 2007

Grounded Theory Review, Vol 6 (Issue #3), 1-38

The online version of this article can be found at:

https://groundedtheoryreview.org

Originally published by Sociology Press

https://sociologypress.com/

Archived by the Institute for Research and Theory Methodologies

https://www.mentoringresearchers.org/

The Grounded Theory Review (2007), vol.6, no.3

Theoretical Elaboration of Quantitative Data[1]

Barney G.Glaser, Ph.D., Hon.Ph.D.

Quantitative data is so closely associated with the current emphasis on verification that its possibilities for generating theory have been left vastly underdeveloped. However, some of our best monographs based on quantitative data indicate that they can be a very rich medium for discovering theory. In these monographs, discovery cannot be stopped, but breaks through both verifications and preconceived conceptual schemes to provide us with very interesting and important theory[i]. Yet, since the authors are still so focused on testing provisionally what they have discovered, their work is mostly written in the hedging rhetoric of verification. The result is that their statements present tests as merely "plausible suggestions." The plausibly suggested test should not be construed with our goal of the purposeful generating and suggesting of theory. The generating capacities of these sociologists and the richness of their research are, therefore, not given the fullest impetus.

Typically, discovery made through quantitative data is treated only as a byproduct of the "main work"- making accurate descriptions and verifications. When discovery forces itself on an analyst, he then writes his INDUCED hypotheses as if they had been thought up before the data were collected, so that they will seem to satisfy the logical requirements of verification.[ii] Purposeful generation of grounded theory is found usually, if at all, in short papers where a single carefully worked-out explanation of a hypothesis is offered, after an analytic wrestle between the rhetoric of tentative qualification and alternative explanations and the carefully researched, accurate data - a slight beginning for an adequate theory.

When the sociologist consciously starts out to suggest a

---

[1] Originally published as Chapter 8 in Glaser and Strauss (1967). The Discovery of Grounded Theory: Strategies for qualitative research, pp.185-220.

theory plausibly, rather than test it provisionally, then he can relax many rules for obtaining evidence and verification that would otherwise limit, stultify or squelch the generation of theory. He must give himself this freedom in the flexible use of quantitative data or he will not be able to generate theory that is adequate (as we have discussed it) in terms of sampling, saturation, integration, density of property development, and so forth. In taking this freedom he must BE clear about the rules he is relaxing (which could not be relaxed for purposes of accuracy and verifications) and he should explain his position to readers. *The freedom and flexibility that we claim for generating theory from quantitative data will lead to new strategies and styles of quantitative analysis, with their own rules yet to be discovered*. And these new styles of analyses will bring out the richness of quantitative data that is seen only implicitly while the focus remains on verification. For example, in verification studies cross-tabulations of quantitative variables continually and inadvertently lead to discoveries of new social patterns and new hypotheses but are often ignored as *not* being the purpose of the research.

In this chapter, we shall present one new strategy of quantitative analysis that facilitates the generation of theory from quantitative data. It is a variation of Lazarsfeld's elaboration analysis of survey data.[iii] In our presentation we shall indicate how, at strategic points, the rigorous rules for accuracy of evidence and verification can be relaxed in order to further the generation of theory. To be sure, there are many styles of quantitative analysis with their own rules. Our focus here is an illustration of how these numerous other styles can also be flexibly adapted to generating theory. However, we do touch on some existing general rules of quantitative analysis (*e.g.,* indexing and tests of significance); the way they are relaxed for purposes of generating theory could apply to many styles of analysis. And we shall also develop some general rules governing how to relax the usual rigor of quantitative analysis so as to facilitate the generation of theory.

The organization of this chapter is based on the successive stages of building up to theory from quantitative data. We discuss in turn the most frequent sources of data used for generating theory, how one indicates his categories and

properties with the data, how one discovers hypotheses with his conceptual indices, and how the hypotheses are then theoretically elaborated. In an appendix to this chapter, we provide examples of theoretical elaboration. For some longer examples of certain specific points, we have referred the reader to other literature.

Secondary Analysis of Quantitative Data

The sociologist whose purpose is to generate theory may, of course, collect his own survey data, but, for several reasons, he is more likely to analyze previously collected data - called secondary analysis. Surveys are usually financed for providing large-scale descriptions of current populations; and the sociologist whose interest is in theory may not wish to be involved in this part of a study, for it takes considerable time and concentration that might otherwise be used for theoretical analysis. It is easier to analyze previously collected data, for then his only responsibility is to generated theory. Sometimes, of course, after the large-scale descriptions have been accomplished, the director of the study returns to his data to engage in secondary analysis for generating a theory on an idea initially stimulated by the earlier descriptive phase.

Generating theory is a more limited, narrowly focused effort (even though the theoretical concept may be very general) than presenting the broad description of a population given by the total survey. The description may involve thousands of questionnaire items, while the theoretical analysis only requires consideration of a few hundred.[iv] Therefore, the tasks of description and analysis can conflict unless the sociologist has adequate money and time (a likelihood only for the study director and a few assistants). Theoretical analysis of quantitative data is, of course, an opportunity to be taken by many sociologists other than study directors or their assistants,[v] and so most generation of theory from quantitative data will be based on secondary analysis.

Comparative analysis requires secondary analysis when populations from several different studies are compared, such as different nations or factories. Comparative analysis of groups internal to one study does not require secondary analysis, but again it often is.

Trivial data, such as found in market surveys on consumption of products, can also have very important theoretical relevance. For example, from a study of meat consumption one can gain knowledge about the life style of social classes. Secondary analysis is a necessity in such cases because sociologists with a theoretical bent do not usually collect such data.

When using secondary analysis of quantitative data for generating theory, one point must be kept clear. Because of the heavy emphasis on accurate evidence and verification of hypotheses, the analyst usually wishes to start out with the facts as facts. One limitation of secondary analysis is the difficulty of pinning down the accuracy of findings in what is necessarily a second hand view - often without much knowledge of collection procedures and meanings of data. Also, since populations are in constant change, we have no way of knowing whether a survey accomplished some years ago for other purposes still applies meaningfully to the specific population. This problem of accuracy is not as important for generating theory about a type of social unit as it is for describing a particular social unit or verifying a hypothesis. What are relevant for theory are the general categories and properties and the general relations between them that emerge from the data. These can be applied to many current situations and locations as very relevant concepts and as hypotheses of interest to sociologists and laymen, regardless of whether the specific descriptions yielded by the data are currently accurate for the research population. Secondary analysis, then, is uniquely well suited for the generation of theory but is often severely limited for description and verification- for which it is still mostly used, with a typical preamble about "limitations."

Another limitation of secondary analysis that makes its use in description and verification questionable, but does not affect the generation of theory, is the representativeness of the population studied. Accuracy is, of course, crucial in description and verification, and the sample must therefore be carefully chosen by some form of random sampling. Secondary analysis of a random sample chosen for other reasons may introduce systematic and random biases into the secondary study, making claims to accuracy questionable. Indeed, it is often

difficult to ascertain from previously collected data what kind of sample was taken for what purpose, since records may have been destroyed, lost, misplaced or made unavailable. Many important questions concerning the sampling become unanswerable, such as how many people did not respond, how many cards were lost, and how many questionnaires were not usable. But when theory is the purpose, there are two reasons why the representativeness of the sample is not an issue. First, the direction of a relationship used to suggest a hypothesis is assumed to exist until disproved, in both biased and unbiased populations; and, second, theoretical (not statistical) sampling guides the choosing and handling of the data.

What is more important for generating theory is the *scope* of the population, which can be increased when the analyst is less concerned about representativeness. Representativeness usually requires some purification of the original sample to obtain a clear-cut population for a smaller study; the sociologist's take for his analysis carefully stratified samples from a larger survey sample. This tactic cuts down on scope by weeding out the possible (but never proven) "contaminating" influences of some respondents. For example, one may wish to take all scientists out a national survey for study, but then, if he purifies the group by weeding out all but the PhD's, he loses the population scope that could have been afforded by keeping the scientists with the MD's, MS's, and BS's.

Concepts and Indices

In the last decade, the flexible use of concepts and their empirical indices in quantitative analysis has been advanced greatly by Lazarsfeld. A number of publications[vi] have carried his work on the "process by which concepts are translated into empirical indices." We wish to mention here only a few general points and urge the reader to study the footnoted references for the general argument and the examples.

When the discovery and generation of theory is the goal of a survey analysis, "crude" or "general duty" indices (as described in detail by Lazarsfeld) sufficed to indicate the concepts of the theory and to establish general relationships between them, which in turn become the basis for suggesting hypotheses for the emerging theory. Similar crude indices,

usually a single questionnaire item or a simple summation index of two to six items, are interchangeable when based on similar, but different indicators. "Interchangeability of indices," as Lazarsfeld demonstrates, means that we obtain the same findings in cross-tabulations with other variables when two indices of the same category are based on reasonably similar but different sets of indicators. Therefore, the analyst does not have to be certain that he has the most accurate index, judged on the basis of either precision or the best set of indicators.

Crude indices, when correlated with other variables, also yield the same relationships in direction as the more precise indices yielded by factor analysis, latent structure analysis, a Guttman scale, or elaborate scales involving dozens of items. Since for generating theory we are only looking for general relationships of direction - a positive or negative relation between concepts, and not either precise measurement of each person in the study or exact magnitudes of relationship - it is easier, faster, and considerably more economical to use the crude index. Even when crude indices result in obvious misclassification of some cases, they still yield the information necessary for generating a grounded theory.[vii]

Crude indices of categories or properties can also be based on either a single questionnaire item or a series of items summed into an index. However, for indices of the core categories, it is perhaps preferable to use two to six item summation indices, since the category will usually be based on at least two dimensions and each should be indicated by at least one item. Further, crude indices need only be dichotomized to obtain comparative groups, not cut into several groups. Whether an index is cut in two, three, or four groups, the same general relation will appear when it is cross-tabulated with another variable, provided that the cutting point is statistically established with criterion variables as a meaningful break in the data.[viii] Dichotomizing an index is financially economical and saves cases for cross-tabulation when the number of cases is small and when the analyst engages in multivariate analysis of three or more variables. Indeed, even if a trichotomous index is used, the analyst, except in cases of exceptional patterns, still ends up talking about the general positive or negative relation between two variables.

When generating theory, validation of a core index - demonstrating that the index measures the concept to a sufficient probable degree - need not be a special operation in which a theoretically relevant relation between two variables is sacrificed from the substance of the analysis itself to prove the validity of the argument, as is typically necessary in verifications.[ix] If the index "works" - that is, if it is consistently related to a whole series of variables that, when put together, yield an integrated theory - this is validation enough of a core index. Integration of the theory is, in fact, a more trustworthy validation of an index than the standard method of merely showing that an obvious relationship exists between the index and another questionnaire item, and that therefore the index must measure what it is supposed to measure.[x]

For example, the core index of "professional recognition" in *Organizational Scientists* (Glaser, 1964) could easily have been validated by showing that professional recognition is positively related to receiving promotions; but instead the whole book shows the validity of the index by the way the substantive theory on scientists' organizational careers is integrated.[xi] In fact, the theory becomes integrated around the core index of recognition because of the multiple relationships with that index, indicating that the theory works - it provides relevant explanations and consequences of organizational careers. Lazarsfeld's methods for specifying concepts and for selecting sub-sets of items to construct indices of the concepts are excellent for ensuring that categories will fit the data and will work or be relevant. This fulfillment of the two major requirements of grounded theory explains why the index becomes validated by the whole theory.

We make these statements in the service of generating theory. If the analyst wishes to describe or verify, these issues must be argued on different grounds, because his problems of precision, dichotomization, and validation of indices are different. The analyst must therefore be clear about his purpose. However, most survey analysts are *not* clear, because Lazarsfeld never has made the distinction between the purpose of generation and those of verification and/or description with accurate findings. He writes not of theory but of "empirical propositions" and "statistical relations." We see clearly how his

work on concepts and indices is valuable for generating theory through conceptual indices and general relations between them. But others who wish to discover "facts" and verify hypotheses, especially by secondary analysis, must argue for Lazarsfeld's methods on their own. Indeed, there are many sociologists who use his methods and stay on the empirical level of description or harp on their findings in the verification rhetoric, even when attempting to suggest theoretical hypotheses.

The survey analyst chooses his categories in the same manner as the researcher doing qualitative analysis. An initial scheme of concepts and hypotheses, usually applied to quantitative data in attempting verifications, is not needed. Concepts whose fit will be emergent are found in previous descriptive analyses with quantitative data, or in other quantitative or qualitative data on the same subject. Also, categories and properties emerge during the collecting and analyzing of quantitative data as readily as they do with qualitative. It must be remembered that qualitative data suggesting a category may also be used as another slice of data for the quantitative analysis.

The theoretical relevance of the concept is soon demonstrated by whether or not its index actually works in a multitude of cross-tabulations. If the index does not work, then the analyst should question the theoretical relevance of his concept before he questions the method of index formation. In quantitative analyses, it is typical to observe a non-emergent category derived from a logico-deductive theory (say, on self-image, role conflict, or status congruency), forcibly indexed - and then found to be related to nothing of theoretical relevance. The analyst then finds fault with the precision of the method of index formation, rather than with the relevance of a category derived from an undergrounded theory, since he seldom questions his faith in the logico-deductive theorist when the latter is a charismatic figure in the profession. Much survey analysis fails for this reason, but we hear failures only through friends; tact prevents citing examples.

It is possible to index any category, but while, with emergent categories, the analyst is almost sure to discover many relations between indices, "ought" categories, from

undergrounded theories are a risk. To stay on the empirical level, using no theoretical categories, is one alternative to taking the chance of directing theoretical research through logically deducted categories such an "anomie" or "authority relations." Yet people who do not trust logico-deductive theory, but wish to do theoretical work, could very safely attempt discovery of grounded theory as another alternative.

Discovering Hypotheses

In generating theory, preconceived hypotheses are not necessary for correlating or cross-tabulating two variables (called runs) with indices of core categories and properties. Indeed, the rule for generation of theory is *not* to have any pre-set or valued hypotheses, but to maintain a sensitivity to all possible theoretical relevances among the hundreds of possible runs afforded by large surveys. In contrast, necessarily preconceived hypotheses direct exactly what two variable correlations to use as tests in verificational studies. Indeed, verificational rules state that data should be collected for tests *after* the hypothesis has been formulated- though they seldom are. For generating theory the data can be collected at *any time*. As we have said, it is usually collected beforehand because most discovery and generation is a secondary analysis of data collected for other purposes, and because the hypotheses come after the analysis - they are suggested from findings, not tested with them.

In order to saturate all possible findings for suggesting hypotheses, the analyst may take his core concepts and run them with literally every other questionnaire item in the survey that seems remotely relevant to his area of interest.[xii] At this point, the theory of the core indices starts to emerge. Clusters of items are discovered as associated with the index. Indeed, this strategy (an unbelievable "sin" in verification studies) virtually discovers theory for the analyst by providing associations to be conceptualized and analyzed.

Table I Cosmopolitan Orientation			
	Motivation to Advance Knowledge		
	High	*Low*	*Difference*
Personal contacts outside organization are very important as sources of scientific information	56%	35%	+21%
If had to, would prefer to move to a university	72%	43%	+21%
Belonging to an organization with prestige in the scientific world is of the utmost importance	40%	21%	+19%
Very strong involvement with close professional work associates	40%	26%	+14%
Very strong sense of belonging to section (principal research group)	44%	27%	+17%
Basic research, as a result of clinical program, is likely to			
-benefit	42%	56%	-14%
-suffer	40%	29%	+11%
Those who would worry about a substantial emphasis on applied as well as basic research	38%	19%	+19%
Base for each percent	(186)	(146)	
Local Orientation			
	Motivation to Advance Knowledge		
	High	*Low*	*Difference*
Having an important job in the organization is of the utmost importance	30%	12%	+18%

Association with high-level persons having important responsibilities is of the utmost or considerable importance	55%	42%	+13%
Having a very strong sense of belonging to the organization	31%	19%	+12%
Interested in a higher level job in the organization which entails stimulating or advising subordinating professionals about their work	77%	67%	+10%
Interested in a higher level job entailing administrative planning or coordination	68%	56%	+12%
Base for each	(186)	(146)	

He induces a theory simply from the general relationships he has found. He need not concern himself with theoretical explanations of what he has found in comparison with what he was supposed to find, as is done in verification studies.

One comparative strategy for generating theory from findings is to compare clusters of relationships within the context of the emerging theory. For example, in Table I we see that "motivation to advance knowledge" (a crude index) is consistently related to two clusters of items, those indicating a *cosmopolitan* orientation – toward the profession – and those indicating a *local* orientation - toward their research organization. Thus we discover and suggest theoretically that highly motivated scientists within research organizations devoted to basic research (a structural condition) possess the property of being local-cosmopolitans.xiii Table II bears out the suggested hypothesis, by showing that in their work activities, highly motivated scientists are both local and cosmopolitan oriented: as more working hours and activities are added to the work week, the highly motivated scientists spend more time on both professional and organizational activities.

Table II Motivation to Advance Knowledge			
Consecutive addition of hours per week spent on various work activities	*High*	*Low*	*Difference*
21 or more hours: own research	76%	61%	+15%
36 or more hours: plus other professional productive work	63%	48%	+14%
41 or more hours: plus non-productive professional work	69%	48%	+21%
51 or more hours: plus other organization activities for total work week	55%	48%	+17%
Base for each percent	(186)	(146)	

Consistency Indices

These two variable runs showing clusters of associations are analyzed comparatively in two ways: *within* and *between* consistency indices. A consistency index is a list of single questionnaire items which all indicate the same category, such as cosmopolitan, and all relate separately to the core index in the same constant direction. The indicators are not added together first and then related to the core index, as in summation indices. Summation indices are best for the core categories, but consistency indices are best for the categories to which a core index is to be related. This strategy allows the analyst to see how the core concept relates to each individual indicator of another category. If inconsistencies in associations between the consistency index and the core index occur for what appeared to be substantively consistent indicators, they are quickly caught and compared to the underlying meaning of the differences within the set of indicators and the emerging theory.

Table III Percentage of Researchers with High Motivation as Related to Their Previous Experience			
Background Experience	*Previous Experience %*	*No Previous Experience %*	*Difference %*
Emphasis on advance of knowledge:			
University Employment	65 (180) ♦	45 (152)	+20
Research and teaching	61 (247)	40 (58)	+19
PhD. Education	62 (164)	40 (58) † 55 (110)#	+22
Emphasis on application of knowledge:			
Medical or clinical practice	58 (244)	55 (88)	+3
Hospitals	57 (111)	55 (121)	+2
Industry	58 (78)	56 (254)	+2
Private practice or business	58 (36)	56 (296)	+2
Government agencies	48 (117)	61 (215)	-13
U.S. Public Health Service	47 (68)	58 (264)	-11
	♦ Figures in parenthesis indicate number of cases.	† Education less than doctorate. # M.D.	

For example, in Table III we see that within the consistency index of applied experience, high motivation to advance knowledge (not to apply it) is not related to previous experiences in private or group practice, hospitals or industry. [xiv] These particular applied experiences then, we theoretically suggest, neither engender nor inhibit motivation negatively related to applied experience in government agencies and the U.S. Public Health Service, or (theoretically) why do these experiences inhibit or reduce motivation to advance knowledge? We suggest that it is because these two experiences, in contrast

to the first four, imply routine service in the application of knowledge.

If all items on "experience in application of knowledge experiences" had been combined first in a summation index, and then related to motivation, these inconsistent comparisons of groups within the consistency index (from which we discovered strategic structural conditions varying the core category) would have been missed; hence, so would an important hypothesis of the theory: the effect of "routine" applications on the scientists' motivation to advance knowledge. The property of "routine application" would have been missed had the analyst simply constructed a summation index, since all the items on applied experience would have seemed internally consistent when tested - all items positively related to each other. Therefore there would have been no suspicion that correlating an applied experience index with another index was actually summing inconsistencies.

Comparisons *between* different consistency indices are also used as a strategy of comparative analysis. We saw in Table I that, since high motivation is positively associated with both a local and cosmopolitan orientation, the analyst can suggest, on the basis of this comparison between consistency indices that scientists highly motivated in research are local-cosmopolitans in a basic research organization.

These two comparative strategies - comparing within and between consistency indices associated with a summation index - occur in three or more variable associations also; but then the analyst is using traditional analytic strategies, which we discuss in the next section. Also, once a detailed analysis of the association with a consistency index is accomplished, then the consistency index can be summed and dichotomized for further analyses with three or more variables. These analyses are more complicated, requiring reduction of details and the saving of cases for cross-tabulation. For example, the first part of Table III shows motivation to advance knowledge related to a consistency index on one kind of previous experience in science - experience emphasizing advancement of knowledge. Table IV shows the summation index of previous experience in science related to two other summation indices - motivation to advance knowledge and professional recognition - for the theoretical

purpose of suggesting hypotheses bearing on the interaction between the three indices.

Table IV Relation of Recognition to Science Experience for Research Workers with High Motivation			
Science Experience	*High Recognition %*	*Low Recognition %*	*Difference %*
Full	76 (46) ♦	69 (52)	+7
Some	68 (75)	42 (99)	+26
None	44 (23)	35 (37)	+9
♦ Figures in parenthesis indicate number of cases.			

Test of Significance

Statistical tests of significance of an association between variables are not necessary when the discovered associations between indices are used for suggesting hypotheses. Selvin[xv] has argued that this rule should be relaxed for all survey analysis, but he can take this stand only because he has not made the distinction between the generating and the verifying or describing purposes of research. He questions whether these tests are appropriate with survey data, since the statistical assumptions necessary to use them cannot be met with such data and also are ineptly applied according to general sociological theory. His critics, however, seem to be more concerned with keeping the tests of significance to ascertain accuracy of evidence used for verification and description.[xvi] We wish to stay clear of this controversy because we are making an argument concerned only with these tests in relation to the generation of theory.

Testing the statistical significance of an association between indices presents a strong barrier to the generation of theory while doing nothing to help it, since the resulting accuracy (if one can actually trust the test) is not crucial. These tests direct attention away from theoretically interesting relationships that are not of sufficient magnitude to be

statistically significant. The analyst usually does not think of the associations as a grounded foundation for a hypothesis, although weak associations may be highly theoretically relevant. Also, the test, not the relationship, may be weak.

Believing that he has no findings relevant for generating theory, the analyst also usually neglects to ask what the partial relationships look like under several conditions. It is easy to forget that partials may be statistically significant even if the general relationship is not because the partials can cancel themselves out. "Cancelling out" means that the relationships may be positive under one condition and negative under another; so that when combined the partial relationships cancel themselves out to result in a weak general association. However, it is theoretically very relevant and interesting to be able to say how conditions minimize, maximize, or cancel out a relationship. Also, even if partials are weak, the theoretical relevance of a weak relationship between two indices may be the *weakness* itself.

Believing in tests of significance can also dissuade one from trusting consistent but weak relationships within and between consistency indices. Yet consistency validates the merit of relationships when it comes to the plausible reasoning required in a credible theoretical analysis.[xvii] And, as just noted, whether the level of the relationship is zero, weak, or strong; it may, if relevant, be grist for the theory.

A belief in tests of significance can also, in the process, direct one's attention away from theoretical relevance of content toward confusing statistical significance with theoretical significance, and a statistical method labelled "analysis" with theoretical analysis. Merely being statistically significant does not mean that a relationship is or should be of theoretical relevance. Such relevance depends on the meaning of the association as it relates to the theory. Also, the statistical analysis methods (for example, "factor analysis", or "analysis of variance") are not theoretical analyses. They are merely techniques for arriving at a type of fact. It is still up to the analyst to discover and analyze the theoretical relevances of these facts. In sum, the basic criterion for generating theory is theoretical relevance, and the analyst should sample his quantitative findings on this basis.

In place of making tests of significance, the sociologist can establish working rules to fit his particular situation. For example, two rules for establishing an acceptable percentage-difference level are not to consider any relationship of, say, less than 10 per cent difference; or any relationship in which three people's changing their minds or being misclassified would change the percentage to below an established level. These levels change with the number of cases used, smaller numbers of cases requiring a higher percentage-difference level. Selvin has also developed an internal replication procedure for establishing the possibility that a relationship exists.[xviii]

Standing by the rules that he may have initially established for his research is pertinent only to the beginning phases of generating theory. When the analyst has achieved theoretical relevance with his data, consistency arises in percentage-difference levels as well as in content, and he will readily learn to understand when and why a lower difference is relevant as well as a higher one. The absence of a relationship becomes just as important as an increase above the consistent percentage level, for any degree of association (or lack of it) may be part of the theory. For example, in Table III the relationship of motivation to previous experience varies at consistent percentage-difference levels - positive (20 per cent) to zero (2 per cent) to negative (-12 per cent) - thus theoretically indicating that these levels are engendered by experiences emphasizing basic research, unaffected by those experiences emphasizing applied research, and inhibited by experiences involving routine service in applied research. In Table V, a consistent percentage-difference level of 10 to 16 per cent shows in comparative relief the theoretical relevance of the stronger and weaker relationship as conditions varying the effect of recognition on satisfaction with organizational personnel.[xix]

Table V Effect of Recognition on Scientists' Satisfaction with Diverse Organizational Personnel			
	Organizational Position of Scientists		
	Junior	*Senior*	*Supervisor*
Assisting Personnel Very and fairly satisfied	+10%	+5%	+11%
Scientific personnel Very satisfied	+16%	+5%	+22%
Leadership Very satisfied	+28% -11	+26% +11	+12% +16
Institute director Very competent Fairly competent	+10% +7	+7% +2	+28% _____

Liberties in Presentation of Data

When quantitative data are reported in verificational and descriptive studies, typically each association is given in table form with a technically exact discussion of it, and then the finding is qualified by tentative statements and alternative explanations or interpretations. This style of presentation need not be used in generating theory, nor, in fact, could it be used. The multitude of relationships on which grounded theory is based is so large that this style applied to each relationship would make the report of the theory unreadable - too long, cumbersome, and slow-moving - to colleagues and quite inaccessible to laymen. It is particularly important that both colleagues and laymen readily understand the theory, [xx] since quantitative data are usually not as interesting to read as qualitative, and do not carry the reader along as easily. Therefore, the analyst must take some liberties both in presenting tables and in making statements about them. Needless to say, the liberties in presentation should not in any way change the data upon which the theory is based; it is just that for generating theory not all data must be presented and stated in exact detail. Since the possibilities are great, each analyst must decide on various liberties according to his particular directions of effort.

Let us consider here a few general liberties of presentation. Unlike Tables I through IV, Table V is a table of percentage-difference. The proportions that were compared to arrive at the differences are left out, since they were not necessary for the theoretical analysis. If it is necessary to know about a particular set of proportions, they should be mentioned in text. However, the focus of analysis in this table was on comparing percentage-differences for indicating direction and magnitude of many relationships: that is, differences in satisfaction with organizational personnel accounted for by the high and low recognition achieved by scientists at different stages of their organizational careers. Both the direction and magnitude of these relationships were important for the analysis; if only direction of relationship had been important, the table could have been further simplified by leaving out numbers and using only plus and minus signs. These flexible renditions of quantitative evidence are in the service of generating theory. No information is lost, distorted, or purposively concealed. It is just that only enough information is presented to show, in the simplest possible manner, the grounded basis of the emerging theory. Verification requires a more detailed rendition of the data - showing all N's, sub-N's and compared high and low percentages - so that the reader can verify the verification for himself.

Because of the overabundance of separate associations necessary in generating theory (literally hundreds, in contrast to the few necessary in verificational studies), another general liberty may be taken in presenting tables, particularly two variable tables. Unless a whole configuration of consistency indices are shown together in a table for visual comparisons, it is enough to state in the written text two variable associations in their direction and (if necessary) magnitude; presenting a table would be repetitious. When theoretically necessary, proportions and N's can be provided in a footnote.

While verificational studies require exactitude, statements about associations can be more flexibly written when theory is the goal. For example, "more successful investigators have satisfactory research facilities provided to them as a reward by the organization" is a statement that assumes the reader understands that three liberties have been taken with this

reporting of a two-variable table. First, the "successful" investigators have been *compared* with less successful investigators - the statement is comparative. Second, "more" means proportionately more - the comparison is relative, not absolute. And third, that the organization provides these research facilities as rewards to the successful investigators is a theoretical inference from the finding that they simply have more satisfactory research facilities than the less successful investigators. Such a hypothesis is more readable than the precise, literal statement: "A higher proportion of those scientists with high professional recognition than those scientists with low professional recognition have satisfactory research facilities. We tentatively suggest that these facilities are provided as rewards to the more successful scientists by the organization."

These three liberties in writing can also sometimes be taken when rendering three-variable tables, and the table need not be put in text. But more often, as noted in the next section, three-variable tables have complex purposes - for example, an interaction table showing the joint effects of two variables on a third (example 4 below). A table and some explicit reporting of it are required for the theoretical inference to be easily understood as being based on evidence.

Theoretical Elaboration

The previous section presented the first step in our style of theoretical analysis of quantitative data: saturating core indices with all possible two-variable runs; discovering relationships among the runs with theoretically relevant consistency indices, summation indices and single questionnaire items; then analyzing the findings with theoretical inferences. The next step, which cannot be neglected, is *elaboration analysis* - to make three or more variable analyses in order to saturate categories further by developing their properties and thereby achieving a denser theory. Thus, the discovery of relationships among indices provides the analyst with beginning suggestions for a theory, plus a theoretical direction and focus for its elaboration.

By "elaboration" we mean that the two-variable associations, which are the basis of theoretical hypotheses,

must have their structural conditions specified; their causes and consequences sought, with possible spurious factors checked for; and their intervening variables (delineating processes between the variables) discovered. Although this, of course, is Lazarsfeld's elaboration analysis,[xxi] we shall contribute something new to his method for our own purpose of generating theory. The next several paragraphs assume an understanding of elaboration analysis (which can easily be gained by a study of Hyman's rendition of it[xxii]). The notions on consistency analysis discussed in the previous section are subsumed in elaboration analysis.

Lazarsfeld has provided three ways of ordering the variables in an elaboration analysis: (1) temporal, (2) structural level of complexity, and (3) conceptual generality. Temporal ordering is simply the time sequence of the variables involved. Structural level of complexity is an ordering in terms of the encompassing structural levels that characterize the unit of analysis under study. For example, a nurse can be characterized by the ward she works on, the hospital she works in, the city in which the hospital is located, and the nation where the city is. Conceptual generality is an ordering by degree of abstractness of the variables. For example, a nurse says all patients should be bathed every day, which is specific opinion derived from a broader attitude of obeying all hospital rules, which attitude in turn derives from a basic value in medicine that nurses should obey hospital rules.

Lazarsfeld's elaboration analysis is seldom used in research except for the prime task of specifying the conditions of a finding; for this task, one need not understand or expressly use his formula. The reason for this lack of use is simple: the only type of ordering of variables that Lazarsfeld has actually worked out is temporal ordering - the other two types have only been suggested.[xxiii] Since survey data is typically cross-sectional in time, analysts are hard put to establish clear-cut, factual time orders in which colleagues will have confidence, because of the emphasis on accurate facts in verification and description.[xxiv] Usually there is too much temporal interrelation among cross-sectional survey variables - over time, either one could, and probably does, result in the other. Thus, elaboration analysis is often stopped in its tracks before it has a chance to

prove its usefulness. And the analyst who does not give it a chance stifles, rather than stimulates, his theoretical imagination. He has been taught not to let his imagination range on data that he cannot himself believe completely accurate, much less argue for their credibility with his colleagues. He has been taught to be sceptical of such strategies for survey data to the point of keeping an empty head about data felt unreliable.

Elaboration analysis is stimulating because the findings it produces fit the thought patterns of sociological theory. With it, the analyst can show interpretations, processes, conditions, causes, spurious factors, and consequences with actual data - not an interpretation of the data. The analyst can literally speak through elaboration tables. He need only infer from his indices the conceptual level of his talk since the tables provide the theoretical arrangement of the variables. But if temporal ordering is believed impossible in most cases, how can we allow theory to emerge from elaboration tables?

Theoretical Ordering

The theory can emerge from these tables if, at first, the analyst decides that his purpose is to generate theory, for then the accuracy of temporal ordering that would be required for verification and description is no longer crucial. He must then proceed to order his variables *theoretically:* a new principle of ordering. Lazarsfeld comes close to suggesting this principle with his "substantive" orderings by structural complexity and conceptual generality, for these are two specific examples of the general principle of theoretical ordering. But Lazarsfeld misses developing a general theoretical ordering principle because he does not consider their underlying similarity, nor how and why they can be used for the generation of theory. He missed this consideration because he is involved exclusively in establishing facts for description and verification. He never comes close to understanding that temporal sequence can be handled theoretically as well as factually.

Theoretical ordering of variables occurs by two strategies: (1) running all possible three-variable associations with each theoretically relevant two-variable association; and (2) running particular tables to fill in gaps or to answer questions, which

emerge as the theory develops, by arranging elaboration tables according to the dictates of the theory. From the findings in both strategies there emerge theoretical orderings of variables already integrated with core categories and hypotheses. The analyst then infers or suggests them as his theory.

Theoretical ordering of variables by all possible three-variable associations on core two-variable relationships is done by comparing the partial association percentage differences to the percentage difference of the original relationship. When the partials vary above and below the original relationship, then the analyst discovers conditions that minimize and maximize his core relationship. From these findings he generates theory stating "under what conditions" a phenomenon exists. Some of these conditions are antecedent to the original association and may be suggested as partial causes; others, which occur at the same time, may be called contingencies. When the partials are equal to the original relationship, then a particular condition does not vary the relationship. The analyst either regards it as theoretically relevant or ignores the finding.

When both partials are less than the original relationship (they never completely disappear), then the analyst must theoretically suggest whether the third variable is (1) an intervening variable, or (2) an antecedent variable. An antecedent variable that reduces partials may have several theoretical meanings. The original relationship may be spurious; that is, both original variables are the consequences of the third variable. This finding may be theoretically very relevant. For instance, "the more fire engines that come to a fire, the greater the damage" is a spurious relationship, with both factors accounted for by size of the fire. The antecedent variable may also suggest a process in which the third variable leads to one of the original variables, which in turn leads to the other. This inference can be tested with the second strategy of theoretical ordering, which is to answer to the question "Is this a process?" by rearranging the table to fit, thus testing for the theoretically assumed ordering of an intervening variable. If the inference proves correct, the analyst has found a value-added process - without the first variable the other two variables do not occur in process.[xxv] Thus the analyst can actively check on his theory as it emerges by testing assumed

theoretical orders. (This will be illustrated shortly.) Third, the antecedent variables always occur together and therefore are truly, not spuriously, associated, but they do not occur without discovered cause, which the analyst might wish to call a necessary condition. Thus fire engines and fires are truly associated, but are not found together unless someone has put in the alarm.

The first strategy of theoretical ordering is based on emergence: the data provides possible orders for the analyst. He need only induce theory about what he has found. This can be difficult when he has to overcome current training in quantitative analysis. He must remember that he is only looking for plausible orders among variables to suggest a theory. He is not looking for the "facts" of a description or verification. *He must think developmentally by remembering that only the data is static or cross-sectional - not his mind!* Although the data may admit of no temporal sequence, his creative imagination can consider any ordering principle for the related variables, and this principle becomes his ingenious suggestion. With imagination and ingenuity he can theoretically order his variables by time, structural complexity, conceptual generality, or in any other theoretical manner. His job is to suggest a theory based both on the *theoretically relevant order* or elaboration relationships and on the *content* of the variables he employs. He cannot think methodologically or statistically with symbols such as t factors or x leads to y; he must think theoretically about the content of his indicated categories and infer why the order of their possible relationships may be as he found them. In short, he must free himself from the exact rules of elaboration ordering as applied to descriptive and verificational studies, so he can be flexible in an imaginative, post hoc theoretical analysis of what he has discovered from the four elaboration possibilities: antecedent or current conditions (PA and PI), antecedent or intervening variables (MA and MI).

In generating theory as it emerges, the analyst first discovers two-variable relationships; second, he discovers their elaboration. Then he moves into a third stage, in which he starts generating possible further elaboration of two-variable relationships within the previous elaboration, using the second

strategy of arranging variables to test theoretical orderings. He looks through his data to find indicators for the concepts he thinks are related in theoretical ways to his emerging theory. Then he arranges his elaboration tables to test if they bear out his hypotheses (for suggestion, not verification), or to discover what actually happens. At this stage of the analysis, he is theoretically sampling his data as directed by his emerging theory and he is actively directing his further runs accordingly; much as the field researcher directs his final work towards theoretically sampling data on hypotheses for filling gaps and answering the remaining questions in order to saturate categories. And much as the field worker at this stage moves quickly between situations, achieving greater relevance with smaller amounts of data, the quantitative analyst may literally camp in the IMB machine room, having successive tables run to continually check his hypotheses as he thinks them through and theoretically samples his data for them[xxvi]. He knows what his data should look like in various runs, and the runs set him straight. By this time the analyst has looked at hundreds of directions provided by the first two stages of his research. Consistency and elaboration analyses join together to provide him a grounded basis for his theory. (The appendix to this chapter gives examples.)

Conclusions

The point of this chapter has been to illustrate the careful relaxation of rules surrounding quantitative analysis, a relaxation for generating theory. The styles of quantitative analysis are multitudinous, so our discussions here include but few illustrations pertaining to the rich veins in quantitative data that can be mined when analysts relax their rigor.

One topic that we have not yet dealt with in this chapter bears mention: comparative analysis within and between surveys. To be sure, the discovery of relationships and their elaboration are all based on comparative analysis of subgroups that are readily found in the same body of data. However, sociologists have yet to explore the many possibilities for generating theory by the active creation of diverse comparison subgroups within a survey (besides core index and typologies), and by the active search for comparison subgroups on other surveys. The various survey-data libraries scattered around the

nation now facilitate comparisons between surveys.

We can suggest a few general rules for beginning this kind of exploration. The analyst can use *similar* groups for comparison between surveys; they do not have to be identical. For example, "working class: may be indicated by residential area in one study, income in another (remember that crude indications are sufficient and interchangeable).[xxvii] Further, the analyst should search for ways of comparing quickly and easily the multiple comparison groups within many different, particularly large, surveys, since one or two surveys can easily run thin on data, and what is needed for a dense, adequate theory is a great amount of data. Also, multiple comparisons should be sought and flexibly done with qualitative data on other relevant groups.

In making these multiple comparisons, the analyst should constantly focus on generating and generalizing a theory not on the comparison of differences to verify or account for a fact. Generating from differences is not easy to manage with quantitative data, since sociologists are trained to verify, and verification from differences comes very easily with quantitative data. Verifying and accounting for facts by differences are subsumed in the process of generating theory; they are not the product of quantitative research for this purpose.

Appendix to Chapter VIII: Examples of Theoretical Elaboration

Following are several examples of theoretical ordering of elaboration tables, which tell the analyst if it is possible to suggest a theoretical statement. We focus primarily on the second strategy of theoretically arranging tables to discover possible orderings for hypotheses.

1. *The discovery and generation of a performance-reward process*. In a study of organizational scientists, the analyst discovered that scientists' motivation to advance knowledge was positively associated with professional recognition for doing so. This finding suggested the theoretical inference that recognition from others maintains motivation.[xxviii] The analyst

then elaborated this relationship by suggesting the following theoretical ordering: if recognition (which indicates previous performance) maintains motivation, then motivation should result in high quality performance in research and this, in turn, should result in more professional recognition. This ordering could then be suggested as a circular, snowballing, reward process for performance within science. The problem then became to order the elaboration tables to test if theoretically (not factually) this process was grounded.

In Table VI, the magnitude of association between recognition and performance is diminished when the intervening effect of motivation is removed. Therefore, high motivation tends to be a link between receiving recognition and accomplishing further high quality research performance, tentatively demonstrating the performance-reward process as a grounded basis for a theory of this process. As a social pattern, this circular process will continue if the performance measured here results in new recognition. [xxix]

TABLE VI

	Recognition		
	Average	*Less*	*Difference*
High performance	56% (144)	44% (188)	+12%
Proportion with high performance and:			
High Motivation	60% (96)	53% (90)	+7%
Low Motivation	46% (48)	37% (98)	+9%

At this point the analyst suggested that, besides research performance, it was also possible to predict behaviour associated with research on the basis of intensity of motivation. This assertion was borne out by one indicator of research behaviour: the amount of time in a typical work week that the scientist puts into his own research activities. Fifteen per cent more of the highly motivated investigators worked 21 hours a week or longer on personal research. Furthermore, 11 per cent more of those who worked 21 or more hours a week on their own research had a high quality performance score. (Note the

discovery of two additional associations.)

TABLE VII			
	Motivation		
	High	*Low*	*Difference*
High Performance	57% (186)	38% (146)	+19%
Proportion with high performance who put:			
21 or more hours per week into their own research	60% (142)	43% (89)	+17%
Less than 21 hours per week into own research	48% (44)	35% (57)	+13%

Next, in theoretically ordering motivation, personal research time, and performance (Table VII), it can be suggested that the highly motivated investigators will tend to put more time into their own research work, and that this time in turn will tend to result in higher quality performance. The magnitude of association between motivation and performance is diminished when the intervening effect of personal research time is removed. This finding then adds a subsidiary link to the circular performance-reward process (diagrammed below).

Performance-Reward-Process in Science

→Recognition→Motivation→Time in Own Research→ ↓

↑Performance←←←←←←←←←←←←←←←←←←←←←←←←

This theory is based on one possible one-time sequence. The reverse time is also possible: some investigators may have developed a high degree of motivation because they put in more than 21 hours per week. Hard work could generate interest. Therefore, we may have another time sequence in the performance process - longer hours in research leading to a high motivation, resulting in high performance. However, this cannot be suggested because the data leave it ungrounded. In comparing proportions downward in Table VII, among those with high motivation 12 per cent more of those who worked 21 or more hours a week on their own research had a high performance score. Among those with low motivation, 8 per cent more who worked 21 hours or more a week on personal research had a high performance score. The original relation

between time and performance, is a condition that creates a slightly stronger relation between the two. This is, of course, the time sequence originally assumed, which shows it to be the only theoretically grounded sequence.

This example indicates the discovery of two-variable relationships and their theoretical elaboration in order to generate a processual theory. The theory is suggested, not tested because obviously the temporal ordering is theoretical, not factual; the data were collected on one day, except for the performance index, for which data were collected three months after the survey. However, *even theoretical ordering provides checks on itself*; even when the two elaboration tables were rearranged, the order of the process did not change.

2. *Structural complexity process.* In the same study of organizational scientists, the following consequences of two different promotion systems in the organization were discovered. [xxx] The "recommended" system (in which initial consideration for a scientist's promotion was based on a supervisor's recommendation) resulted in more discrepancies between rank and actual responsibilities and in more unsatisfactory evaluations of the system than did the "routine" system (in which initial consideration for promotion was based on periodic reviews). Theoretically, it seemed that a process was involved, whereby the relative frequency of perceived discrepancies resulting from each promotion system was a reason for the relative number of unsatisfactory evaluations of each system. The analyst then arranged an elaboration table to test for this theoretical order (Table VIII), and the findings supported it - the partial associations (22 and 25 per cent) were less than the original associations (29 per cent), showing that discrepancies were an intervening variable between systems and evaluations. This theoretical process was supported by another consistency finding that among scientists in the "recommended" system there was considerably less satisfaction (29 per cent) with chances for a promotion.

TABLE VIII			
	Promotion System		
	Recommend	*Routine*	*Difference*
Evaluate promotion process as unsatisfactory	58% (184)	29% (145)	+29%
Proportion who evaluate promotion process as unsatisfactory and who observed discrepancies:			
Frequently	83% (59)	61% (28)	+22%
Occasionally	45% (125)	21% (117)	+24%

Here the theoretical ordering of variables is based on structural contexts at different levels, and assumes that the more encompassing level has a greater effect on the lesser level rather than vice versa. Thus "promotion systems" is a contextual unit that causes discrepancies in rank and responsibilities among personnel; while "discrepancies" is a property of the system that provides a structural condition affecting the way scientists evaluate their systems' promotion procedures. Thus, mixed into this structural level process are contextual properties of individuals or structural conditions under which they have a career (promotion procedures and characteristic discrepancies in rank and responsibilities) and for a system (evaluations); properties of a system (procedures, discrepancies, and dissatisfied individuals); properties of individuals (evaluations), and so forth - depending on how the analyst wishes to render and focus his theory. In short, even within this simple structural process, as found in one elaboration table, the analyst can find much grist for sociological theory.

3. *Theoretically rearranging on a table to test for alternative career processes.* The question arose about how those scientists who planned to move to relieve the pressure of a currently unsuccessful career have made this decision.[xxxi] They may (1) decide to leave the organization, and then choose the goal they plan to work for - perhaps still basic research (by going to a university) or perhaps a change of practice or applied research (by going to either a private, industrial or governmental research organization); or (2) decide to change

the goal of their work from basic research to another goal, such as applied research or "practice," and this change would necessitate leaving their organization as soon as possible.

Table IX is arranged to test for the sequence of factors in the first process: "plans to move" is tested as an intervening variable, coming between degree of recognition and preference for a preferred affiliation in a university, if the move is made. Since the original relation is nil, we discover that this theoretical elaboration test for an intervening variable is a test if the non-existent original relationship was actually a cancelling-out of a strong positive relationship (between recognition and preference under the condition of planning to move soon) and a strong negative relationship (between recognition and preference when planning to stay on in the organization). Thus this table corrects our theoretical ordering by yielding a finding that suggests that unsuccessful scientists who plan to move (11 per cent in Table X) have not yet planned to go on with either basic research or applied research or practice. They are still just planning to move because of a poor career, and they have not decided where or for what purpose.

TABLE IX

	Recognition		
	High	*Low*	*Difference*
Prefer to move to university	62% (144)	63% (188)	-1%
Proportion who prefer and who plan to:			
Move soon	66% (12)	69% (36)	-3%
Stay for time being or permanently	58% (130)	57% (152)	+1%

Table X is arranged to test the second-mentioned process in making plans to move. Preference for the university or for other organizations is tested as intervening in the decision to move as soon as possible made by those who lack recognition. Again, planning to move because of low recognition is *not* a result of planning to change work goals- both partials are not less than the original relationship of 11 per cent. What this table tells us is that the scientists' plans to move as soon as possible materialize (15 per cent) under this condition of a

certain preference for moving to a university where their research goals would be the same. On the other hand, plans to move soon hardly materialize (7 per cent), if at all when the scientists prefer an organization offering them another work goal.

TABLE X			
	Recognition		
	High	*Low*	*Difference*
Plan to move as soon as possible	8% (144)	19% (188)	-11%
Proportion who plan to move as soon as possible and who prefer to move to:			
University	10% (84)	25% (111)	-15%
Other organization	7% (60)	14% (77)	-7%

Thus, theoretical arrangements of elaboration tables, while not necessarily bearing out our theoretical guesses, discover for us *what is going on* (in, say, the decision to leave an organization because of a failing career). They fill gaps in the total theory of organizational careers and answer our specific questions.

4. *Specifying joint effects of conditions.* Seldom are both partial associations less than the original association; the most frequent finding specifies antecedent or contingent conditions that minimize and maximize relationships. These findings yield perhaps the most frequent of theoretical statements: the varying conditions under which a phenomenon exists. As we have said, the specification of conditions may apply to a single index, but as an elaboration procedure it applies to two or more variable relationships. Antecedent conditions (such as previous research experience, Table III) may, if the theory warrants, be suggested as partial causes. Conditions occurring at roughly the same time are called contingencies, denoting whether a relationship is contingent on a condition that makes it more or less pronounced. Further, for his theory, that sociologist may choose to reverse the temporal order of his specifications of conditions to obtain statements on the varying consequences of diverse aspects of a condition (types, dimensions, or degrees of

the condition). Thus, this type of elaboration table yields findings that suggest several ways to generate a theory.

TABLE XI Percentage Who Are Very Satisfied With Job Security

	Felt Recognition		
	High	*Low*	*Difference*
Organizational Position:			
Junior Investigator	67% (57)	43% (84)	+24%
Senior Investigator	70% (40)	58% (60)	+12%
Supervisor	73% (47)	73% (44)	___

Joint effects is another theoretically interesting way of talking about the specification on conditions. In Table XI we see the joint effects of scientists' organizational position and degree of professional recognition on their satisfaction with the security of their job in the organization. A standard means for rendering this table is to say that when we hold organizational position constant, professional recognition only makes for job security in the investigator position. But "holding constant" is a notion used in verification of theory, when the analyst is trying to reduce the contaminating effects of any strategic variable not in focus with his variable of interest.

To view the table in terms of joint effects of two conditions on a third lends itself better to generating theory, since no variable is assumed a constant; all are actively analyzed as part of what is going on. For example, in Table XI we see that as a scientist's organizational position advances (or for the theory, as his career advances), professional recognition becomes less important for job security (the percentage differences decrease). Another joint effect for theoretical inference is that, as the scientist's career advances, he becomes more secure in the organization through seniority, and less dependant upon his degree of professional recognition for this security (under "low recognition" security percentages increase with position). Or the analyst might say that a scientist with professional recognition to his credit tends to have a secure job no matter what his organizational position. (See percentages under "high recognition.") Thus, statements of joint effects tell

us how conditions interact together to affect a third variable - and this is theoretically rich and relevant information.

TABLE XII Proportion of Junior Investigators Who Are Very Satisfied With Security Of Job

	Recognition		
	High	*Low*	*Difference*
Promotion System			
Recommend	63% (30)	37% (51)	+26%
Routine	69% (26)	50% (32)	+19%

Two other ways of making inferences about this table are in terms of "differential impact" and "differential sensitivity."[xxxii] For Table XI the analyst can say that position has a *differential impact* on the relationship between recognition and security. In Table XII, we see the differential impact of promotion systems on junior scientists' satisfaction with job security under different conditions of professional recognition. These, again, are forms of contextual and conditional comparative analyses. Referring again to Table XI, the analyst can say that the security of the scientists with low recognition is very sensitive to organizational position, while the security of scientists with high recognition is insensitive to organizational position - thus indicating the *differential sensitivity* of the successful and unsuccessful in their job security.

Finally, the analyst can generate minimal and maximal *configurating* conditions (a useful theoretical model) for his theory from a joint-effects table like Table XI. To be at the beginning stages of a career without recognition is to feel comparatively little satisfaction with job security. Maximum security comes at the peak of one's career in the organization, because of tenure. Though it took professional recognition to achieve this position, recognition is no longer a condition for job security.

We could suggest more ways to generate theoretical statements from joint-effects tables, as well as from the first three illustrations of elaboration tables. However, we wish only to conclude from these brief illustrations that *if quantitative*

data is handled systematically by theoretical ordering of variables in elaboration tables, the analyst will indeed find rich terrain for discovering and generating theory. We hope by our slight but purposeful loosening of the rules, via our principle of theoretical ordering, that elaboration analysis will be used more than heretofore. Its richness for research has not yet been tapped because of difficulties in using it on cross-sectional survey data to produce accurate facts for description and verification.

---

[i] For examples see James Coleman, "Research Chronicle: The Adolescent Society," and Seymour Martin Lipset, "The Biography of a Research Project: Union Democracy," in Philip Hammond (Ed.), *Sociologists at Work* (New York Basic Books, 1964).

[ii] This way of presenting one's work in a publication on research is not chicanery, but an established form in many circles of science. See Bernard Barber and Renee C. Fox, "The Case of the Floppy-eared Rabbits: An Instance of Serendipity Gained and Serendipity Lost," *American Journal of Sociology*, 64 (1958), pp. 128-29.

[iii] Paul F. Lazarsfeld, "Interpretation of Statistical Relations as a Research Operation," in Lazarsfeld and Rosenberg (Eds.), *The Language of Social Research* (Glencoe, Ill.: Free Press, 1955).

[iv] For example compare the theoretical analysis in Barney G. Glaser, *Organizational Scientists: Their Professional Careers* (Indianapolis: Bobbs-Merrill, 1964) to the description from the same study using over 100 different IBM card decks and comprising four volumes. *Human Relations in a Research Organization*, Volumes I and II (1953) and *Interpersonal Factors in Research*, Parts I and II (1957) (Ann Arbor, Mich.: Institute for Social Research). For another example see Hanan C. Selvin, *The Effects of Leadership* (New York: Free Press Glencoe, 1960).

[v] See Barney G. Glaser, "The Use of Secondary Analysis by the Independent Researcher," *The American Behavioral Scientist* (1963), pp. 11-14.

[vi] Paul F. Lazarsfeld, "Problems in Methodology," in R. Merton, L. Broom and L. Cottrell (Eds.), Sociology Today (New York: Basic Books, 1959), pp.47-67; "Evidence and Inference in Social Research," Daedalus, LXXXXVII (1958), pp.100-109; and with Wagner Thielens, The Academic Mind (Glencoe, Ill.: Free Press, 1958), pp. 402-407.

[vii] It is at this point, Lazarsfeld suggests, that technicians, who perhaps have no generative powers, take flight into precision by blaming their crude methods and trying to refine their indices instead of thinking about what they have found.

[viii] In constructing a summation index, the analyst first obtains one more group than the number of indicators he is using: four indicators lead to five groups. Before combining these groups he should cross-tabulate the five groups

with a criterion variable- he knows the relationship exists- to find out between which groups the direction of the relationship changes. He then combines all those groups positively related to the criterion variable and all those negatively. He cannot just dichotomize the index where he pleases, because he may reduce its discriminating power by combining positive and negative degrees.

[ix] For an example see Lazarsfeld and Theilens, *op. cit.,* pp. 89-90.

[x] This is a specific case of Zetterberg's rule that the total integration of a theory tends to make any one of its parts "highly plausible." See Hans L. Zetterberg, On Theory and Verification in Sociology (Totowa, N.J.: Bedminster Press, 1963), Chapter 6.

[xi] Glaser, *op. cit.*

[xii] If the analyst has enough time and money, he can run the index open (use all groups) and then dichotomize them at the breaking point for each item. This will yield more diverse information on each relationship and make the index more sensitive. This strategy is an alternative to dichotomizing on a criterion variable, but is cumbersome; and once the analyst is sure his break in the index is the most sensitive one, it may seem a waste of time for the yield of information.

[xiii] For the theoretical discussion of Tables I and II, see Glaser, *Organizational Scientists: Their Professional Careers, op. cit.,* Chapter 2.

[xiv] For the theoretical discussion of Tables III and IV, see Barney G. Glaser, "Differential Association and the Institutional Motivation of Scientists," *Administrative Science Quarterly*, 10 (1965), pp. 81-97.

[xv] Hanan Selvin, "A Critique of Tests of Significance in Survey Research," *American Sociological Review*, 22 (1957), pp. 519-27; "Statistical Significance and Sociological Theory" (July, 1960) (mimeographed, University of California, Berkley)

[xvi] Robert McGinnis, "Randomization and Inference in Sociological Research," *American Sociological Review*, 23 (1958), pp. 408-14; Leslie Kish, "Some Statistical Problems in Research Design," *American Sociological Review*, 24 (1959), pp. 328-38; and critical comments by David Gold and James Beshers in *American Sociological Review*, 23 (1958), pp. 85 and 199

[xvii] That consistency validates is a basic pattern of plausible inference. See G, Polya, *Patterns of Plausible Inference* (Princeton, N.J.: Princeton University Press, 1954), Vol. II, Chapter XII

[xviii] Selvin, *op. cit.*

[xix] For theoretical discussion of Table V see Glaser, *Organizational Scientists: Their Professional Careers, op. cit.,* Chapter 6

[xx] In contrast, verifications usually only require the understanding of a small group of colleagues working in the area.

[xxi] See references in footnotes 3 and 6

[xxii] Herbert H. Hyman, *Survey Designs and Analysis* (Glencoe, Ill.: Free Press, 1955), Chapter VII

[xxiii] Herbert H. Hyman, Survey Design and Analysis (Glencoe, Ill.: Free Press, 1955), Chapter VII

[xxiv] The evaluator of an article for one journal remarked on an elaboration tale, "More generally the whole argument about establishment vs. persistence (or stability) of the relationships suffers because the author really has no other time trend data- that is necessarily implied in statements about persistence or stability." The paper was rejected because temporal order was not an incontrovertible fact.

[xxv] See for a discussion of this type of process Neil Smelser, *Theory of Collective Behavior* (New York: Free Press of Glencoe, 1963), Chapter I.

[xxvi] This is a frequent activity among some survey analysts; see Coleman's discussion of continually having tables run as he thinks them through, *op. cit.,* pp. 203-204

[xxvii] See Herbert H. Hyman, *Political Socialization* (Glencoe, Ill.: Free Press, 1959), for examples of combining similar categories for comparative analysis

[xxviii] Glaser, *Organizational Scientists: Their Professional Careers, op. cit.,* Chapter III

[xxix] See *ibid*, p. 32

[xxx] *Ibid.,* Chapter III

[xxxi] *Ibid.,* Chapter VIII

[xxxii] *Ibid.,* Chapter IV