

Scale Validity  
A Computer Content Analysis Approach

Donald G. McTavish  
Department of Sociology  
University of Minnesota  
Minneapolis, Minnesota 55455

E-mail: [mctavish@atlas.socsci.umn.edu](mailto:mctavish@atlas.socsci.umn.edu)  
Phone: (612) 462-5231 home  
(612) 624-4075 university  
(612) 624-7020 fax  
Preferred mailing address: 6810 Viking Blvd NE, Wyoming, MN 55092

ABSTRACT

This paper presents a computer content analysis approach to the problem of assessing the validity of Likert scales used in survey research. The wording of questions which make up a scale communicates meaning to the respondent in order to acquire relevant responses which, when combined, serve to measure some concept. Traditionally, beyond informal judgments of face validity, investigators examine responses of a sample of subjects to scale questions in order to determine if expected relationships and structure are found. The content analysis procedure proposed here directly examines the wording of scale questions to see if expected meanings, relationships and structure are found in order to assess the validity of a scale. Possibilities of this approach are illustrated using two previously analyzed sets of scales: alienation, and perception of police and crime.

Keywords: MCCA, Minnesota Contextual Content Analysis, Computer Content Analysis, Scale Validity, Likert Scale, Content Validity, Survey Questions

AUTHOR'S NOTE

The analysis reported here was partially supported by grants from Rural Sociology and the University Computer Center. I want to acknowledge the earlier work of Debbie Felt, a former graduate student who contributed substantially to the analysis and preparation of an earlier draft from which the current paper draws. I am indebted to Susan Schrader who provided helpful comments and suggestions. I also want to acknowledge Pete Cattrell who, as a graduate student at the University of Minnesota, first utilized the MCCA content analysis approach to analyze measurement items as a seminar project in the author's content analysis class. Contact the author

for information about MCCA.

## SCALE VALIDITY: A COMPUTER CONTENT ANALYSIS APPROACH

### INTRODUCTION

By traditional definition, a measure is valid to the extent that it measures the concept that the investigator intends to measure (Kerlinger 1973). Measurement involves linking abstract concepts and empirical indicants (Blalock 1982). Validity is evident when measures behave as they are supposed to behave as technical concepts in theory. Deductions from theory about expected (high or low) relationships between a proposed measure and other measures are, if confirmed, core evidence about that measure's validity.

Zeller and Carmines (1980) illustrate a variety of ways by which the validity of measures of an intended concept can be assessed when multiple-item (multiple-question) scales are used. Examples such as the 9-question "social isolation" scale and others are shown in Tables 1 and 3, below. Traditional validity (and reliability) assessment, including that proposed by Zeller and Carmines, starts with subject-response data. Sampled subjects are asked to respond to proposed scale questions and their responses are analyzed.

In addition to analysis of subject-response data, investigators utilize informed but relatively informal face or content validity judgments to generate, evaluate and revise scale items (Kerlinger 1973, p.458). Face validity is used to assess the extent to which scale items are related to the concept an investigator wishes to measure and the extent to which they have relevant meaning for respondents. For a multidimensional construct, this involves a more complex judgment of content coverage to determine how items tap different facets of a concept's defined "universe of content".

For many scales, the concept the investigator wants to measure is the obvious topic of scale items and is recognized by respondents. Validity depends on how well question wording communicates this understanding to the respondent. For other scales, the concept an investigator has in mind may be a more abstract idea underlying the structure of a set of questions where the respondent does not, and need not appreciate the concept as a whole. Nevertheless, the meanings in scale items which the respondent does understand are central to the validity of the scale. Questions relevant to alienation, for example, ask about the subject's social isolation, sense of power, feelings of meaninglessness and normlessness. Here, one might expect the structure of meaning understood by the respondent to reflect the structure of the underlying abstract concept although the subject may not know the concept.

It is well recognized that validity outcomes may be sample specific. Different subject samples may require different question wording to communicate the meaning essential to valid measurement (Dillman 1978, pp.95-7; Labaw 1980, p.11). Establishing effective comparability of meaning is a key measurement problem especially in cross-cultural survey research (Marsh 1967, pp.271-80; Osgood 1967).

This article attempts to contribute to content validity assessment of scale items by describing and illustrating a computer content analytic approach. Two scales analyzed by Zeller and Carmines (1980) using subject-response data are used to illustrate the computer content analysis approach proposed here. This procedure is intended to complement, not replace, other systematic approaches to the assessment of scale validity.

Computer content analysis permits systematic and reliable scoring of emphasized ideas and the social perspective (or social "context") that is used in phrasing scale items. Such scoring permits a more systematic, comparative analysis of meaning structures represented by scale items. Hypotheses can be tested about the way in which the scale stimulus is structured, about similarities between the content of items and definitional statements, and about the "closeness" of item wording to typical language used by a particular subject pool (and, thus, the likelihood they will appropriately understand scale items). The computer content analysis procedure may also provide leads which aid in evaluating and revising problematic scale items. First the computer content analysis approach used in the study is briefly explained. Then two illustrations are given utilizing scales previously analyzed in detail by Zeller and Carmines (1980) with subject-response data, to show how this approach may be utilized. Parallels are evident between the structure of stimulus items and the structure of subject responses.

#### CONTEXTUAL CONTENT ANALYSIS: A COMPUTER APPROACH

Content analysis is an established research methodology for systematic examination of textual material (Kerlinger 1973; Markoff 1975; Labaw 1980, pp.137-41). In its usual "hand" version, however, coder reliability is likely to be a severe problem, especially when the coding scheme is complex or the volume of text and cases (scale items) is large (Krippendorff 1980; Dillman 1978, p.88). Coder training and coding effort are also problems for systematic hand content analysis, undoubtedly reinforcing the use of more informal expert opinion as the methodology of choice in examining face or content validity of scales (for an exception see Furnham and Henderson's 1982, systematic, hand content analysis of four personality inventories).

Computer-aided content analysis helps solve coder reliability and text-volume problems of hand analysis. We are proposing that these procedures be applied to an analysis of the wording of survey scale items themselves; the symbolic, linguistic stimulus structure by which an investigator's concepts are effectively presented to the respondent.

Several computer content analysis procedures exist (Weitzman and Miles, 1995). From our point of view, it is critical to use a content analytic procedure which takes account of all words used in scale items. Thus, simply counting selected words or categories is not sufficient. Furthermore, we find it helpful to use a procedure which is normed against usual language usage, permitting an assessment of over and under emphasis on a profile of ideas. Finally, it is helpful to take account of the social perspectives underlying phrasing of the scale items (i.e. how the ideas they deal with are approached or framed; the underlying social assumptions by which wording is to be understood).

Computer Content Analysis. In this analysis, the Minnesota Contextual Content Analysis program (MCCA) is used (described in McTavish and Pirro, 1990, and Litkowski, 1997b). It was developed to provide an overall assessment of themes in textual data. Since most research requires a comparison between texts or texts and some standard, MCCA attempts to facilitate such comparisons, to help handle coder reliability issues, and to facilitate using the strengths of qualitative and quantitative text analysis strategies.

MCCA is a dictionary-based program which uses a relatively large number of concept categories (116) of general social science interest, such as the idea of "happy" (a category including words like "cheerful", "joyful", "attractive"). The dictionary includes categories which capture various

aspects of social structure, senses of time, social relationships, role categories, personal reactions, cognition, physical place, etc. These categories have also been used in combination to capture other concepts such as the balance of references to "we" vs. "they". Litkowski (1997a, 1997b) describes principles and desirable characteristics of such semantic categories and suggests future directions for further refinement. The dictionary includes words which account for about 90 percent of English usage. Words in a text are assigned uniquely to one category (or to the leftover list). Input text is verbatim text with punctuation. No special preparation or tagging is required other than identifying the beginning and end of a text. As in any analysis, the researcher needs to carefully gather text relevant to the research topic. Output consists of two score profiles, one that captures ideas which are emphasized and the other which captures the framing of ideas which characterizes certain different social contexts.

An idea emphasis score is developed for each category, calculated as the difference between the proportion of all words in a text that are in a given category minus the expected usage of the category's words. The expected usage of a category is based on overall norms for English word usage. This difference is divided by a measure of the expected variability in the usage of that category across different social contexts (McTavish and Pirro, 1990). The resulting score, called an E-score (idea Emphasis score) reflects the relative emphasis or de-emphasis of that idea category compared to what one would expect of general English usage. The value of this scoring is that it provides a basis for comparing relative emphasis on categories within and across texts, provides a means for detecting unusual censoring or omission of some ideas (something that is virtually impossible to do by hand), and it provides a profile of 116 emphasis scores which is able to capture with some precision, more global concepts in which one might be interested. A profile distance measure can be computed between the vector of E-scores for pairs of texts. If ideas are similarly emphasized the profile distance is small but if quite different ideas are emphasized the distance measure is larger. A matrix of these distance measures can be cluster analyzed to identify the structure of differently emphasized meanings which characterize various texts which are being compared (Kruskal and Wish, 1978).

The second score profile is called a Context emphasis score (C-score). It is computed from the profile of category use data for a text, weighted using an empirically developed series of 4 weights for each category. The weights reflect differential use of the 116 content categories found in 4 general social-institutional contexts. Weights were developed empirically in a factor analytic examination of texts from different institutional sectors of society (Cleveland, McTavish and Pirro, 1974). The four are: a) "Traditional" or an approach that is concerned with norms and appropriate behavior, the "shoulds" of social life; b) "Pragmatic", concerned with successful goal accomplishment and achievement; c) "Emotional" refers to personal criteria for enjoyment, appreciation, and personal engagement; and, d) "Analytic" or the detached, objective attention to something of interest. A zero score on all C-score dimensions is what one would expect for a broad mix of text from across the various different institutional sectors of society. Where a given context is emphasized in the text the corresponding C-score will be higher (negative C-scores indicate relative under-emphasis on a context dimension). C-scores have been useful in discriminating between social and organizational settings and measuring social distance (McTavish, Litkowski and Schrader, 1997). They represent the way in which a text is framed in terms of the four general institutional perspectives. A distance measure can be computed between texts using the profile of C-scores. This permits one to identify the structure of differences in the framing of ideas for texts that are being compared.

The MCCA approach helps identify the structure of differences between texts in terms of meanings and perspectives used in the texts. More complex concepts and themes are caught in patterns of idea or contextual categories. This approach has been used in examining a wide variety of texts including poetry, interviews or focus group transcripts, media and advertising text, court and clinical records, etc. MCCA has been used in examining scales in at least two other areas (Rajecki and McTavish, 1995, and Pierce, McTavish and Knudsen, 1986).

The data for the following illustrations come from an MCCA analysis of the text of each question which is part of a Likert scale (e.g. "Most people seldom feel lonely."). The exact wording of each question constitutes the input data. These are shown in Tables 1 and 3, below. The questions are traditional Likert items with response categories strongly agree to strongly disagree. These standard response categories are omitted from the input text. MCCA procedures are uniformly applied to all words in each question in a scale. Scores generated by MCCA for each item consist of a profile of 116 idea E-scores (plus a 117th leftover category for words not in the MCCA dictionary) and a profile of 4 C-scores. All together the output data matrix is a k-item by 121 (117+4) score matrix. Scores reflect an attempt to quantitatively measure some aspects of the meaning of items and provide a systematic basis for further comparative inquiry into the similarities and differences between items. Score profiles are displayed by MCCA in a variety of ways thought to be useful in comparing texts. Raw scores are provided for use in further statistical analysis in conjunction with other variables. Traditional statistical analysis can be performed on these scores and other scores can be added to the data set (such as the scale identity of each item, or reliability and validity coefficients developed from respondent-data experience with different samples of respondents).

#### SCALE CONTENT VALIDITY: TWO EXAMPLES.

Our purpose is to illustrate an empirical, computer content analytic way to assess face validity. We will use two sets of scales from Zeller and Carmines (1980). The first consists of three scales having to do with perceptions of crime and police protection. The second is Dean's measure of 4 facets of alienation. In both cases Zeller and Carmines provide careful validity and reliability analysis using respondent data with which our content analytic results can be compared. These two approaches provide supporting but somewhat different perspectives on content validity of these scales. Certain other comparative items are added as explained below. Example 1: Perception of Crime and Police

Table 1 lists scale items used in this first example. Six scale items concern citizen evaluation of police service performance. Of these, three constitute a sub-scale called "Crime Fighting", and three define a "Police Style" sub-scale. Three additional items measure the "Perceived Probability of Crime" (PPC). Respondent data came from a study evaluating police services in 60 neighborhoods of three metropolitan areas: St. Louis, Rochester, and Tampa-

---

[Table 1 about here](#)

St. Petersburg. Some 11,000 phone interviews were conducted during 1977. Validity and reliability analysis of these data are provided by McIver, Carmines and Zeller (Zeller and Carmines, 1980, appendix). Several approaches to validity and reliability analysis confirm a three-scale solution.

Several expectations for content analytic results stem from their analysis. First, we expect

that all three scales ("Crime Fighting", "Police Style", and "PPC") will have high intra-scale similarity both contextually and in terms of ideas emphasized. We also expect that there will be inter-scale differences. Secondly, in confirming the content of what they term the "Crime Fighting" scale, they noted greater similarity between responses to that scale and items on PPC. Thus, we would expect to see this similarity as well, confirming their interpretation of the crime fighting scale's content validity. Finally, they note that the crime fighting scale is less satisfactory and in particular item "C" (see Table 1) did not seem to fit as well with other items in their analysis. This too ought to be evident in our data. Beyond these expectations, we would expect that all three scales are distinct from other concepts. For illustration, we have chosen Bradburn's (1965) "Happiness" scale item (see "K" in Table 1), to see whether perception of crime and happiness are tapping a common "personal welfare" concept. Comparisons relevant to these expectations can be approached in a number of ways.

A similar analysis can be conducted using the profile of 116 idea scores (plus the leftover category) for each item. A 9 x 9 distance matrix can be constructed which measures differences between idea profiles for each pair of scale items (and for the "added items" listed in Table 1). A cluster analysis of this matrix resulted in a three-dimensional solution and Figure 2 shows the plot for dimensions 1 and 2.

First, it is apparent that the three scales are distinct from each other in the cluster plot. The similarity in idea-emphasis between items within the police style and the crime perception scales is clear. Virtually the same ideas are similarly emphasized. The crime fighting scale, however, emphasizes a broader range of ideas. If one adds the third dimension to Figure 1, the

---

Figure 1 about here

first two scales remain tight clusters but the crime fighting scale has item "B" raised (above the page) and items "A", and "C" lowered (with "C" lowest). Item "J" is between "A" and "C" in this third dimension. The topics emphasized by Bradburn's item ("K") are quite distinct from topics addressed by the other scales. The profile of ideas raised by ex-inmates in describing their prison experiences is close to ideas expressed in the crime fighting scale. Their discussion is closest in topics emphasized to scale item "A" in Figure 1 but are raised above the page in the third dimension, as is item "B".

The specific ideas emphasized can be seen in an examination of E-scores for specific categories. Table 2 provides selected E-scores showing over and under emphasis on ideas as compared to expected English usage. For this illustration, only E-scores with an absolute value of 20.0 or larger are listed 3.

---

Table 2 about here

From Table 2 it is clear that Bradburn's "happiness" item taps a different content than any of the other scales, including the perceived crime scale. The only overlap is with the category "object". Zeller and Carmines found the crime fighting and perceived crime scales to be more closely correlated ( $r = -.41$ ) in subject-response data than was true of police style and perceived crime ( $r = -.25$ ). This led them to conclude that the items validly measure a concept they call "Crime Fighting". In our analysis, the average difference between the crime fighting scale and perceived crime is indeed smaller than the difference between police style and perceived crime,

both contextually (27.2 vs 101.0) and conceptually (19.0 vs 86.3). However, from Table 3, the conceptual overlap between "crime fighting" items and "perceived crime" items appears primarily in dealing with "cognition" (and the category "you" which is shared with the "police style" scale). There does not appear to be evidence here of content overlap between these scales in "crime and crime fighting" ideas, such as "sinning", a traditional norm-deviance category. It is not clear to us that "crime fighting" is the concept most likely to be measured by that scale.

The revised scale item "J" was to have covered the content of old item "C" in a way that was more in line with the crime fighting scale. Table 2 suggests that a good deal of overlap exists, especially in the emphasis on "speed". On the other hand, "J" represents a new emphasis on "self-other" and "changing", but does not address the idea category of "incompetence". With this added perspective, an improved revision might be attempted. We have found that the interplay of this analysis, the item wording itself, and a new emphasis on conceptual clarity, often leads to better ideas about item wording and what it might mean to respondents.

A discriminant function analysis of 11 items (9 police scale items, the revised item, and Bradburn's happiness item) was conducted using C-scores and E-scores as predictors of the known scale placement of the 11 items. Three significant E-score predictors permitted all 11 items to be correctly classified. They are the categories "time", "happy", and "sense". Crime fighting was low on all three categories. Police style was high only on the category "sense". Perceived crime was high only on the category "time" (although the size of the E-score, 11.3, was not large enough to be listed in Table 2). The Bradburn item was high only on category "happy".

Other scale analysis techniques are useful as well, (e.g. factor analysis) where the MCCA scores are treated as "cases" and scale items are treated as "variables". Analysis of covariance structures to test hypotheses about underlying conceptual variables is also helpful. Weber (1983) suggests this approach on more traditional content analytic data.

## Example 2: Alienation Scales

The second example of contextual content validity analysis utilizes Dean's traditional four sub-scales of alienation. Discussions by Zeller and Carmines (1980) and others note issues in the literature concerning the adequacy of the conceptual definition of alienation as well as the clarity of items measuring the sub-scales of this complex concept. They also note empirical data supportive of the four-scales which have been used to measure dimensions of alienation. As Sartori (1984) notes, concept analysis is not sufficiently pursued in our literature, and this observation seems relevant here as well. Content analysis of the scales may provide an avenue for clarifying content aspects of this concept. For example, we might ask whether there is overlap between the content of scales and definitional statements. Does there appear to be a structure of intra-scale similarity and inter-scale difference which would support the four-dimensional alienation structure? Could one sort the 36 items into the proper sub-scale from content and context scores based on their wording? Does this analysis provide any insight into problematic items and how one might strengthen them?

Our analysis builds upon subject-response validity analysis found in Zeller and Carmines (1980). Their analysis is based upon the Neal and Groat (1974) sample of 334 married Ohio women who were interviewed in 1963 and again in 1971, using all four alienation scales. The 36 alienation items are listed in Table 4, with definitional statements from work by Neal and Groat (1974). All 36 items and the 4 definitional statements were scored using MCCA.

---

Table 3 about here

The data matrices used here include C-scores and E-scores for each item, similar to those discussed in the first example. As in the first example, a set of interview data are added for comparison (see Table 3). Text used here is from research by Woodworth (1982) who conducted conversational interviews about life satisfaction with 41 older women who were residents of Ohio nursing homes.

C-scores (not shown here) indicate that the perspective taken in each of the scales is quite different. "Social Isolation" and "Powerlessness" both take a traditional approach but the first de-emphasizes pragmatics and the second de-emphasizes the emotional context dimension. "Normlessness" is pragmatic and not emotional but "Meaninglessness" is a scale that is primarily emotional, not pragmatic. Thus these data suggest that, overall, sub-scale differences are distinct in the way they are framed. There is also a parallelism between life satisfaction interviews and the Meaninglessness scale, suggesting that older women in nursing homes might easily appreciate the perspective taken by these items and presumably provide more reliable and valid responses. This might not be the case for older residents answering the Powerlessness scale which is contextually quite different from the way they talk about life satisfaction. For this group of subjects, attention to introductory instructions for the powerlessness scale may be needed. Although the four definitional statements are not elaborated, they can serve as an illustration of our approach (see Table 3).

A distance matrix between these scales, scale definitions and the life satisfaction interviews can be constructed from E-score profiles for each item. In this case a four-dimensional solution is satisfactory and a plot of the first two dimensions is shown in Figure 2. The definitions and scales show considerable difference in idea profiles although a basic structural similarity between items is also evident. Note that the life satisfaction item is close to the powerlessness and meaninglessness scales in kinds of ideas that are emphasized.

---

Figure 2 about here

Table 4 shows major ideas emphasized by the four scales and their definitional statements. Again, the distinctiveness of scale emphases and their lack of overlap with their definitional statement is apparent. One avenue for scale revision might involve a conceptual assessment of emphasized ideas and their correspondence with a clarified conceptual definition of each scale.

---

Table 4 about here

Figure 3 is a plot of the context distance matrix for all 36 items, scale definitions and the life satisfaction interviews. KYST, a non-metric multidimensional scaling program was used to plot the distance matrices (see Kruskal, Young and Seery, 1977). A five dimensional cluster solution was most appropriate and only a two-dimensional plot of dimensions 1 and 2 is shown here as illustration. This underscores the complexity of overlap between alienation scales and items. Nevertheless, somewhat distinctive sub-scale regions can be identified for the four scales even in this two-dimensional plot.

---

Figure 3 about here

Social isolation is a large region including items 1 through 9, possibly with the exception of item 8. Powerlessness could include items 10-13 and 15. Normlessness could include items 20 through 26 and possibly 27. Finally, Meaninglessness could include items 28 through 32 (see Table 7 for items and numbers). If the regions are identified in this way, then items excluded from the regions are the set of problematic items noted in factor analysis work of Neal and Groat (1974) and Zeller and Carmines (1980). For example, items 8, 16, 19, 26, 27, 34, 35, and 36 all have low communalities (from .07 to .24) in factor analyses of the scales. Some items have loadings on more than one factor. For example item 14 is on meaninglessness and powerlessness (.30 and .68) as is item 17 (.42 and .47). Item 20 is loaded on powerlessness and normlessness (.31 and .51). Finally, item 34 is loaded on meaninglessness and social isolation (.32 and .31). Other items have low factor loadings in one or both measured years, suggesting their marginal stability as a defining part of the alienation scale construct. That is true for items 18, 26, and 33 which have low factor loadings in 1963 (.19, .19, .17), but acceptable loadings in 1971 (.50, .50, .40). Items 35 and 36 have the reversed situation with higher loadings in 1963 (.27, .31) than in 1971 (.14, .05). Items 19, 27, 34, and 35 have marginal loadings for both time one and time two data (ranging from .14 to .27 for both years).

The powerlessness and meaninglessness scale items are more clearly definers of separate factors, suggesting that these scales may have items which are more parallel in context and content emphasis than is true of other scales.

A discriminant function analysis (not shown here) was conducted using C-score and E-score data from each of the 36 items as predictors of correct scale placement. All items could be correctly placed on the basis of three scores: the pragmatic contextual score, and the cognitive and social pressure idea score categories, although a larger number of idea scores significantly contributed to the discriminations. Two functions could be identified. Function one emphasizes "structure", "merit", "achievement", "consideration" and "they". Function two emphasizes "feel", "heard", "like", "guide" and "being". The social isolation scale is relatively negative on both function 1 and 2. Normlessness is positive on both functions. Powerlessness and Meaninglessness load positively on function 2 and negatively on function 1. These diagnostic data may be helpful in examining what it is that distinguishes the emphases in these scales and thus represents, for the respondent, the effective placement of the scales.

## CONCLUSION

In this paper we propose an empirical approach to measure face or content validity. Scale item wording directs the respondent's attention, defines the topic of interest and sets the underlying contextual perspective and meanings enabling the desired subject response. A computer-based, contextual content analysis approach (MCCA) is used to score scale items, definitional statements and comparative text used in these illustrations. Two kinds of scores are created using the MCCA procedure. One is C-scores, which consist of a profile of emphasis on four dimensions called "traditional", "pragmatic", "emotional", and "analytic". The other, E-scores, are a profile of 116 normed scores showing over and under emphasis upon dictionary-based idea categories (plus a leftover category). Thus one can identify the kinds of ideas which are emphasized and also the way in which these ideas are framed in scale statements.

MCCA scores were generated for two sets of scales, police service scales and alienation scales. In each case, similarity and difference between items and scales could be seen and these paralleled the subject-response validity analyses conducted by Zeller and Carmines (1980). An

in- depth examination of specific substantive aspects of a scale would, of course, be pursued beyond that illustrated here.

In this paper computer content analysis procedures are applied to the wording of questions themselves, in this case, scale item wording. This helps reveal the structure of differences in meanings of a set of items which measure a concept. Differences between items used to measure different concepts or different facets of the same concept can be seen. Problematic items can often be identified and the analysis suggests ways in which scale items might be revised. Revisions can be tracked, graphically using multidimensional scaling of distances between items computed from E-scores and C- scores. Useful comparisons can also be made between the meanings embedded in the phrasing of scale items and a) text reflecting how different target populations handle similar concepts (perhaps leading to versions of a scale which better fit diverse populations), and b) meanings embedded in elaborated formal conceptual definitions of a concept the scale is intended to measure. While not a panacea, the approach suggested here appears to have utility in scale construction and validity assessment prior to field pre-testing, thus complementing traditional subject-response validity analysis.

## NOTES

1. Dictionary category labels attempt to characterize, overall, the words included in a category, although somewhat different labels might better capture the sub-set of words in a category which are actually found in a given text. Content analysis makes use of E-scores to direct an examination of particular words and phrases (via key-word-in-context, "KWIC", or global-search computer utilities) to refine idea usage and category labels. This would be essential where large amounts of text were analyzed, such as in analyzing open-ended interviews. A complementary approach is to focus on the profile of E-scores which captures a pattern of emphasis across idea categories (as in the cluster analysis in Figure 1), or to examine networks of co-occurrence of ideas in text (a part of MCCA but not presented here). Table 2 presents the dictionary category labels and does not pursue the examination of words and phrases in text since items are brief and are listed in Table 1.

## REFERENCES

- Blalock, H. M. (1982). *Conceptualization and measurement in the social sciences*. Beverly Hills, CA: Sage.
- Bradburn, N. M., & Caplovitz, D. (1965). *Reports on happiness*. Chicago: Aldine.
- Cleveland, C. E., D. McTavish and E. B. Pirro (1974, September 5-13). Contextual content analysis. Proceedings of the ISSC/CISS Workshop on Content Analysis in the Social Sciences, a conference sponsored by the Standing Committee on Social Science Data of the International Social Science Council, UNESCO, Centro Nazionale Universitario del Colcolo Electronico (CUNCE), Pisa, Italy.
- Dillman, D. A. (1978). *Mail and telephone surveys: the total design method*. New York: John Wiley.
- Furnham, A., & Henderson, M. (1982). "A content analysis of four personality inventories." *Journal of Clinical Psychology* 38 (4) :818-825.
- Kerlinger, F. N. (1973). *Foundations of behavioral research*, 2nd edition. New York: Holt, Rinehart and Winston.

- Krippendorff, K. (1980). *Content analysis: an introduction to its methodology*. Beverly Hills, CA: Sage.
- Kruskal, Joseph B. and Myron Wish (1978). *Multidimensional Scaling*, Sage University Paper no 11, Sage Publications, Beverly Hills, CA.
- Labaw, P. (1980). *Advanced questionnaire design*. Cambridge, MA: Abt Books.
- Litkowski, Kenneth C., (to appear, a). "Category development based on semantic principles" *Social Science Computer Review*. (electronic manuscript available at ken@clres.com).
- Litkowski, Kenneth C., (1997b). "Desiderata for Tagging with WordNet synsets or MCCA categories." 4th Meeting of the ACL Special Interest Group on the Lexicon. Washington, DC: Association for Computational Linguistics. (electronic manuscript available at ken@clres.com..
- Markoff, J., Shapiro, G., & Weitman, S.R. (1975). "Toward the integration of content analysis and general methodology." Chapter 1 in D.R. Heise (ed.), *Sociological methodology: 1975*. San Francisco, CA: Jossey-Bass.
- Marsh, R. M. (1967). *Comparative sociology: a codification of cross-societal analysis*. New York: Harcourt, Brace and World.
- McTavish, Donald G., Kenneth C. Litkowski, and Susan Schrader, (1997). "A computer content analysis approach to measuring social distance in residential organizations for older people", *Social Science Computer Review*, 15:2 (summer), p170-180.
- McTavish, Donald G. and Ellen B. Pirro (1990). "Contextual content analysis." *Quality and Quantity* 24: 245-265.
- Neal, A. G., & Groat, H.T. (1974). "Social class correlates of stability and change in levels of alienation: a longitudinal study." *Sociological Quarterly* 15:548.
- Osgood, C. E. (1967). "Cross cultural comparability in attitude measurement via multilingual semantic differentials." Chapter 13 in M. Fishbein (ed.), *Readings in Attitude Theory and Measurement*. New York: John Wiley.
- Pierce, Jon L., D. G. McTavish, and Kjell R. Knudsen (1986). The measurement of job characteristics: a content and contextual analytic look at scale validity, *Journal of Occupational Behavior*, vol 7, p299-313.
- Rajecki, D. W. and D. G. McTavish, (1995). Content analysis of attitude scale items from the theory of reasoned action and the tripartite (ABC) model, (manuscript)
- Sartori, G. (ed.). (1984). *Social science concepts: systematic analysis*. Beverly Hills, CA: Sage.
- Weber, R. P. (1983). "Measurement models for content analysis." *Quality and Quantity* 17: 127-149.
- Weitzman, Eben A. and Matthew B. Miles, (1995) *A software sourcebook: computer programs for qualitative data analysis*, Thousand Oaks, CA, Sage Publications.
- Woodworth, D. L. (1982). *A content analytic study of religious meaning and life satisfaction*. PhD dissertation, University of Minnesota.
- Zeller, R. A., & Carmines, E.G. (1980). *Measurement in the Social Sciences: The Link Between Theory and Data*. Cambridge: Cambridge University Press.
- Zeller, R. A., Neal, A.G., & Groat, H.T. (1980). "On the Reliability and Stability of Alienation Measures: A Longitudinal Analysis." *Social Forces* 58 (4) :1195-1204.

#### SOFTWARE CITED

The computer content analysis procedure used in this analysis is called the Minnesota Contextual

Content Analysis program (MCCA) version 8.3, a mainframe implementation at the University of Minnesota. MCCA is available in a micro-computer version as an option in DIMAP, a dictionary development package by Ken Litkowski. The PC version of MCCA is available through (CL Research, 20239 Lea Pond Place, Gaithersburg, MD 20879) or web site, <http://www.clres.com>. Contact the author for further information on MCCA (e-mail [mctavish@atlas.socsci.umn.edu](mailto:mctavish@atlas.socsci.umn.edu)).

## BIOGRAPHICAL SKETCH

Don McTavish is a professor in the Department of Sociology at the University of Minnesota - Twin Cities where he teaches courses on computer content analysis methodology and research. He can be contacted at the Department of Sociology, University of Minnesota, Minneapolis, MN 55455 (E-mail: [mctavish@atlas.socsci.umn.edu](mailto:mctavish@atlas.socsci.umn.edu)).

Table 1: SCALE ITEMS FOR CITIZEN PERCEPTION OF POLICE AND CRIME  
"Crime Fighting" Scale:

A. How would you rate the overall quality of police services in your neighborhood? Remember, we mean the two or three blocks right around your home. Are they outstanding, good, adequate, inadequate, or very poor?

B. Do you think that your police department tries to provide the kind of services that people in your neighborhood want? (Responses: Yes/No)

C. When the police are called to your neighborhood, in your opinion, do they arrive very rapidly, quickly enough, slowly, or very slowly? (Also coded: Not at all) "Police Style" Scale:

D. Policemen in your neighborhood are basically honest. Do you agree or disagree? Do you feel strongly about this?

E. The police in your neighborhood are generally courteous. Do you agree or disagree? Do you feel strongly about this?

F. The police in your neighborhood treat all citizens equally according to the law. Do you agree or disagree? Do you feel strongly about this?

Perceived Probability of Crime:

G. How likely do you think it is that your home will be burglarized in the next year? Do you think it is very likely, somewhat likely, or not at all likely? H. How about vandalism; how likely do you think it is that your home will be vandalized in the next year? Do you think it is very likely, somewhat likely, or not at all likely?

I. How likely do you think it is that you will be robbed by someone with a weapon in your neighborhood in the next year? Do you think it is very likely, somewhat likely, or not at all likely?

Added Items:

J. (Revised item C, above) Do you think that the police try to react quickly when they are called to your neighborhood?

K. Taking all things together, how would you say things are these days? Very happy, Pretty happy, Not too happy. (Bradburn Happiness Scale, used here as a contrasting concept.)

L. Open-ended, conversational interviews with 16 ex-inmates (combined data) about their prison experience (25,226 words).

Figure 1: CLUSTER PLOT OF INTER-ITEM IDEA PROFILE DISTANCES FOR PERCEPTION OF POLICE AND CRIME ITEMS, WITH CONTRASTS USING THE BRADBURN ITEM AND EX-INMATE INTERVIEWS

Table 2: SELECTED E-SCORES CHARACTERIZING PERCEPTION OF POLICE AND CRIME SCALES AND THE BRADBURN HAPPINESS ITEM

Idea Category	Original Items A-B-C	Crime Fighting- Item J	Revised Style D-E-F	Police Crime	Perceived G-H-I	Bradburn Happiness K
Incompetence	75.7					
Depressed	31.1					
Positive moods	26.9					
Good	38.9		30.7			
Speed	111.3	255.1				
Cognition	25.7	68.5		68.5		
You	22.4	25.3	32.3	20.7		
Community	24.5	27.6	22.8		24.5	
Future	24.1	64.1			56.5	
Changing		93.3				
They		24.5				
Self-Other		131.7	48.1			
Processing things			30.0			
Sense			23.6			
Sinning				25.5		
Object				25.3	56.1	
Happy					488.5	
Time					22.9	
Feeling					21.3	

Table 3: ALIENATION SCALE ITEMS

Social Isolation:

1. Real friends are as easy to find as ever.

2. Sometimes I feel all alone in the world.
3. People are just naturally considerate and helpful.
4. Most people are not really sincere in their relations with others.
5. Most married people in our country lead trapped (frustrated) lives.
6. There are few dependable ties between people any more.
7. The world we live in is basically a friendly place.
8. Most people seldom feel lonely.
9. The way things are now, a person has to look out pretty much for himself.

Powerlessness:

10. People like me can change the course of world events if we make ourselves heard.
11. I think each of us can do a great deal to improve world opinion of the United States.
12. There's very little that persons like myself can do to improve world opinion of the United States.
13. The average citizen can have an influence on government decisions
14. This world is run by the few people in power, and there is not much the little guy can do about it.
15. It is only wishful thinking to believe that one can really influence what happens in society at large.
16. A lasting world peace can be achieved by those of us who work toward it.
17. More and more, I feel helpless in the fact of what's happening in the world today.
18. There's very little we can do to keep prices from going higher.
19. Wars between countries seem inevitable despite the efforts of men to prevent them.

Normlessness:

20. In getting a good paying job, it's necessary to exaggerate one's abilities (or personal merits).
21. In getting a job promotion, some degree of "apple polishing" is required.
22. In order to get elected to public office, a candidate must make promises he does not intend to keep.
23. Having "pull" is more important than ability in getting a government job.
24. Success in business can easily be achieved without taking advantage of gullible people.
25. Those running our government must hush up many things that go on behind the scenes if they wish to stay in office.
26. In order to have a good income, salesmen must use high pressure salesmanship.
27. Those elected to public office have to serve special interests (e.g. big business or labor) as well as the public's interest.

Meaninglessness:

28. It's hard to sleep nights when you think about recurrent crises in the world and what would happen if they exploded.

Table 4: Alienation Scale Items, continued:

29. The tensions in the world today make one wonder whether he will be around in a few years or not.
30. The international situation is so complex that it just confuses a person to think about it.
31. The only thing one can be sure of today is that he can be sure of nothing.
32. Current political events have taken an unpredictable and destructive course.
33. In spite of what some people say, the lot of the average man is getting worse, not better.
34. Most people live lives of quiet desperation.
35. With so many religions around, one really doesn't know which one to believe.
36. One should live for today and let tomorrow take care of itself.

Conceptual Definitions

Definitions are those of Neal and Groat (1974, pp.550-2; see also Zeller and Carmines 1980, p.22).

1. Social Isolation - "the view of social relations as impersonal, non-rewarding, and unfriendly."
2. Powerlessness - "a low expectancy for control over the outcomes of events (in political and economic spheres)."
3. Normlessness - "a high expectancy that socially unapproved behavior is necessary in goal attainment--the necessity of either coercion or deception in achieving socially desired political or economic goals."
4. Meaninglessness - "the perception of broader social and political events as overwhelmingly complex, without purpose, and lacking in predictability."

Woodworth Ohio Nursing Home Interviews

Comparative data are aggregated (133,065 words) from conversational interviews on life satisfaction with 41 older women residents of Ohio nursing homes, in research on religious beliefs and meaning conducted by Woodworth (1982).

Figure 2: CLUSTER PLOT OF IDEA EMPHASIS DISTANCES FOR ALIENATION SCALES, DEFINITIONS AND LIFE SATISFACTION INTERVIEWS

Table 4: SELECTED E-SCORES CHARACTERIZING ALIENATION SCALES, DEFINITIONS AND LIFE SATISFACTION INTERVIEWS

Soc Isolatn Idea Category	-----A l i e n a t i o n S c a l e s-----				Life Interviews
	Powerless Item.. Def.	Normless Item.. Def.	Meaningless Item.. Def.	Satisfaction Item.. Def.	
Eternal	46.2				
Good	34.8		18.9		23.4
Authority	31.8				
Community	28.3				
Expression	28.3	48.4			
Depressed	28.3				
Qualities	17.9				
Sensation	83.9				
Detached Roles	83.4				
Not	28.9				
Equivocation	46.8	-11.0		45.0	
Social Pressure		31.4	12.4		
Prohibit		42.7			
Error		22.9			
Status		30.5			
Control			58.9		
A-an			12.8		
Spatial			48.5	25.0	
Business		39.3		21.0	
Tradition		59.0		11.3	51.8
Object			-14.0	-14.0	10.3

Processing Thgs	16.5		
Ought	22.2		
Happy	20.1		
Submission	450.0		
Obligation	108.0		
Sin	104.0		
Activity	28.1		
If	24.5		
Strive Nouns	43.7	34.3	
Cognition		29.7	13.9
Disastrous		29.8	
Bad		20.0	
Bungling Things		19.1	170.0
Goodness Ideals			27.5
Academic			20.7

Figure 3: CLUSTER PLOT OF CONTEXT DISTANCES BETWEEN ALIENATION SCALE ITEMS, DEFINITIONS AND LIFE SATISFACTION INTERVIEWS

Source: Zeller and Carmines (1980) "Appendix." 162-186. Note: Item identification letters correspond to labels on tables and figures elsewhere in this paper.

Labels are those in Table 3. Distances (not shown here) are based on item profile differences across the 117 idea categories. The optimal clustering solution using KYST-2 (Kruskal and Wish, 1978) is 3-dimensional (stress = .044). Here the two-dimensional plot of dimensions 1 and 2 is shown.

Only E-scores showing substantial emphasis (20.0 or larger) are listed in this illustration. Heuristic category labels are given. The MCCA dictionary used here includes 116 idea categories (word groupings) which collectively include the most frequently used English words (about 90 percent). Words with more than one meaning are contextually disambiguated by MCCA.

Note: Item identification numbers correspond to labels on tables and figures elsewhere in this paper.

Only the largest E-scores (generally 20.0 or larger) are listed in this illustration. Heuristic category labels are given. The MCCA dictionary used here includes 116 idea categories (word groupings) Distances (not shown here) are based on context scores in Table 5. The optimal clustering solution using KYST-2 (Kruskal and Wish, 1978) is three dimensional (stress = .001). Here the two-dimensional plot of dimensions 1 and 2 is shown. Labels are those in Table 3.

Only the largest E-scores (generally 20.0 or larger) are listed in this illustration. Heuristic category labels are given. The MCCA dictionary used here includes 116 idea categories (word groupings) which collectively include the most frequently used English words (about 90 percent). Words with more than one meaning are contextually disambiguated by MCCA. See note 1.

Labels are those in Table 3. Distances (not shown here) are based on profile differences across the 117 idea categories. The optimal clustering solution using KYST-2 (Kruskal and Wish, 1978) is four dimensional (stress = .006). Here the two-dimensional plot of dimensions 1 and 2 is shown.