

Pre-Assessment of Scale Reliability A Computer Content Analysis Approach

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

ABSTRACT:

This paper examines the relationship between certain features of the language used in scale questions and internal consistency reliability of the scale. It is hypothesized that greater inconsistency in contextual perspective, suggested by the wording of questions making up a scale, leads to lower internal consistency reliability of the scale, due to interpretative confusion this introduces for the respondent. Measurement of the context of item wording and topical emphasis patterns utilizes a computer-based contextual content analysis procedure (MCCA). This approach provides an early, systematic means for evaluating scale reliability which may be useful in developing scales for survey research. Data consist of quantitative measures of word patterns for each question from a sample of 46 Likert scales where internal consistency reliability has been reported. A range of reliabilities is represented as well as a range of substantive areas.

Pre-Assessment of Scale Reliability A Computer Content Analysis Approach

Introduction

Investigators estimating the reliability of a series of Likert scale questions generally examine subject response data using test-retest, split-half or some form of internal consistency coefficient such as Cronbach's alpha (Carmines and Zeller, 1979). Statements are initially crafted based upon an investigator's insight into a domain of conceptual meaning and informed guesses about how people will understand and react

to the phrasing of ideas in statements, both singly and as a set. Typically, in the early stages of scale development or revision, the investigator is unaided by systematic quantitative tools. Once pretest data are obtained from a sample of respondents, however, standard statistical analyses can help identify problematic items and estimate the scale's reliability.

Using computer-based content analysis procedures, the phrasing of scale questions themselves can be systematically examined during the scale development process, *prior* to field pretesting. Such an examination can provide valuable information to assist an investigator in improving the measurement qualities of scales. In prior work, content analysis has been used to examine the *content validity* of scales (McTavish, 1997; Pierce, McTavish, & Knudsen 1986; McTavish & Felt 1985; Furnham & Henderson, 1982). This paper suggests that computer content analysis procedures may also be useful to alert an investigator to *reliability* problems of a scale even prior to the expense of pretesting with samples of respondents.

Scoring Features of Question Wording.

The Social Context of a Text One of the features of text which helps in deciphering its meaning is its general orientation or the way in which it is "framed" for its audience. A text could be framed, for example, in terms of social expectations having to do with a practical, problem-solving or "business-like" approach. This would be communicated in the pattern of word choices, largely without regard for its substantive content. Communication in different social settings typically frames even similar topics in somewhat distinctive ways by the choice and use of words. Thus, a discussion of *serious crime* would be expected, in a religious context, to be framed in terms of assumptions about right and wrong. The same topic discussed in an academic setting would likely be framed in a more analytic way, more open to alternative assessments and facts about extenuating circumstances. Socialized individuals quickly read the extant social situation and its underlying perspectives, and understand the social context of what is being said. These contextual perspectives are also evident in question wording.

Differences in contextual perspective and potential confusion can be seen when strangers meet and attempt to figure out where the other is "coming from". Differences in contextual perspectives are also evident across statuses in organizations (e.g. managers and employees) and in marriage and small group conflicts. Identifying contextual perspectives is an aspect of the definition of a social situation which people use in orienting their expectations and organizing their behavior (Berger and Luckmann 1966; St Clair and Giles 1980, p. 265; Goffman, 1974).

These shared perspectives can be organized at several levels. The more general are

broad social/institutional perspectives. In general, the greater the contextual difference between actors, the more problematic sharing meanings becomes. In this paper we focus on contextual differences between questions in a scale which are evident in question wording.

Computer Content Analysis Content analysis is a procedure which can be used to systematically examine text for its various meanings (Krippendorff 1980; Weber, 1990). However in hand content analysis, coder reliability is a serious issue, especially where scoring schemes are at all complex. Computer-assisted content analysis handles coder reliability issues and provides a more sensitive and systematic scoring of textual data. Thus far, computer content analysis appears not to have been extended to the examination of the reliability of Likert scales.

In this analysis, contextual perspectives as well as ideas expressed in text are scored using the Minnesota Contextual Content Analysis program, or MCCA (McTavish and Pirro, 1990; Litkowski, 1997b). MCCA was developed to provide an overall assessment of themes and the social context of textual data and to facilitate comparisons between texts.

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 "speed" (a category including words like "hasty", "quick", "rapid", and "sudden"). The dictionary includes words which account for about 90% 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. Output consists of two score profiles, one of which captures the ideas emphasized and the other captures the framing of ideas that characterize different social contexts.

An idea emphasis score (E-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 usage. This difference is divided by a measure of the expected variability in usage of that category across different social contexts (McTavish and Pirro, 1990). The resulting score measures 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, a means for detecting unusual censoring or omission of some ideas (something that is virtually impossible to do by hand), and the profile of 116 emphasis scores is able to capture 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 that characterize various texts that are being compared.

The second score profile is called a context score (C-score). It is computed from the profile of category usage data for a text, weighted using an empirically developed series of 4 weights for each category. Weights were developed empirically in a factor analytic examination of texts from different institutional sectors of society (Cleveland, McTavish and Pirro, 1974). The weights reflect differential use of the 116 content categories in selected social-institutional settings.

MCCA uses a set of four context or C-scores to locate a text in a four-dimensional "social context" space. These four contextual perspectives appear to be widely known and shared in a culture and serve to differentiate various social institutions and organizations. The *traditional* context is concerned with rules for appropriate behavior or norms which are often emphasized in such social institutions as the military, the judicial system, the family, and religion. Another perspective is *pragmatic*, which includes concerns about success and efficient accomplishment of instrumental goals. This perspective is often associated with various business ventures. A third perspective is *emotional*, which is concerned with comfort, self-actualization, and personal standards for enjoyment or appreciation. Leisure-oriented organizations emphasize this perspective. Finally, an *analytic* perspective is characterized by curiosity and study of events in a more objective or intellectualized fashion. This perspective tends to be emphasized in research and educational institutions. The four C-scores aid in identifying differences between texts in the way they frame what is being discussed.

Contextual distance between texts can be measured by a Euclidean distance between C-score profiles for pairs of texts. The larger the distance measure, the more the texts take a distinctive contextual approach. These contextual distances can also be clustered to display the structure of differences in social perspective taken by a set of texts.

In past work, religious texts generally have a stronger tradition C-score and small scores on the other three contextual dimensions. Interviews with managers, for example, are higher on the pragmatic dimension. Research reports in scholarly journals tend to be higher on the analytic dimension (and very low on the emotional dimension). Taped interviews about a person's own life situation tend to be expressed in a more personal and emotional way.

MCCA has been used in a number of studies where contextual distances or emphasized meanings were of interest. Differences based upon respondent descriptions of their organizational setting cluster administrators, staff and residents of nursing homes into separate clusters (McTavish and Felt 1984) Descriptions of work contexts serve to identify people with a particular organization. The context of phrasing of youth organization mission statements map the historic developmental structure of these

organizations (Erickson 1986). C-scores have been useful in discriminating between organizational settings and in measuring social distance between statuses (McTavish, Litkowski, & Schrader, 1997). The validity of scales has also been examined using MCCA (McTavish, 1997; Pierce, McTavish & Knudsen, 1986).

An Illustration

The data for the following illustration comes from an MCCA analysis of the text of each Likert scale item (e.g. “My parents were concerned who my friends were.”). Scale questions are input data, as shown in Table 1a, for the *Parental Acceptance Scale*, a 12-item scale measuring parental acceptance or rejection (Gage 1988). Questions are traditional Likert items with response categories strongly agree to strongly disagree. These standard response categories are omitted from the input text. The small number of words in a typical statement is not a problem here because of the very intentional nature of scale item wording and the way in which the MCCA norming procedure takes account of all words in a statement.

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. 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 coefficients developed from respondent-data experience with different samples of respondents). Measures of the C-score and E-score distance between scale items is used to predict published scale reliability scores.

Table 1b shows the four context scores for each of the 12 scale items given in Table 1a, and an overall profile for the scale as an aggregate. The social context expressed by item phrasing is emotional and traditional (positive scores on a scale with a maximum of 25.00) and pragmatic and analytic perspectives are de-emphasized (negative scores on a scale extending down to -25.00). Table 1c provides Euclidean distances between all pairs of scores and for this set of items the maximum distance is 37.4 and the average distance is 9.8. The pattern in the distance matrix clearly shows two items (items 7 and 8, indicated by **) which are relatively distant from other items in the scale. Table 1d shows the distance between pairs of items in terms of their profile of emphasis on idea categories. Here the maximum distance is 150.0 and the average distance is 112.7. Items are more consistently spaced in terms of the topics raised, suggesting a consistent sampling strategy over a domain of content.

Table 1
ILLUSTRATION OF CONTEXT SCORES FOR ONE SCALE
1a) The Parental Acceptance Scale¹

1. My parents totally ignored me.
2. My parents encouraged me to bring my friends home and tried to make things pleasant for me.
3. My parents treat me harshly.
4. My parents viewed me as a burden.
5. My parents talked to me in a warm and affectionate way.
6. My parents were concerned who my friends were.
7. My parents made me feel wanted and needed.
8. My parents forgot important events that I thought they should remember.
9. My parents tried to help me when I was scared or upset.
10. My parents compared me unfavorably to other children no matter what I did.
11. My parents let me know I was not wanted.
12. My parents treated me gently and with kindness.

1b) Context Scores for Parental Acceptance Scale Items

Item	Traditional	Pragmatic	Emotional	Analytic
1	9.62	-10.10	15.38	-14.90
2	7.63	-11.96	17.37	-13.04
3	11.20	-13.04	13.80	-11.96
4	7.32	-14.10	17.68	-10.90
5	10.98	-14.69	14.02	-10.31
6	12.24	-15.82	12.76	-9.18
*7	9.95	-3.26	15.05	-21.74
*8	15.01	-25.00	3.83	6.16
9	9.26	-12.70	15.74	-12.30
10	8.95	-11.57	16.05	-13.43
11	8.14	-12.62	16.86	-12.38
12	11.73	-13.53	13.27	-11.47
Overall	9.99	-13.36	14.44	-11.07

¹ The Parental Acceptance Scale is presented in Rohner (1986). The present scale (Gage 1988) is a shortened version with a split-half reliability that is relatively good (.90 based on a large sample of married adults in the Twin Cities). Although it is clear which items are relatively different, contextually, this scale shows much less contextual distance compared to other scales which were examined.

1c) Contextual Distances (Euclidean) Between Scale Items

Items	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	3.9	4.7	6.5	6.8	8.9	9.7	28.8	3.7	2.3	4.1	5.7
(2)		5.3	3.1	6.1	8.5	12.7	27.9	2.5	1.9	1.2	6.2
(3)			5.7	2.4	4.2	13.9	24.2	2.8	3.8	4.4	1.0
(4)				5.2	7.4	15.8	25.7	3.4	4.3	2.4	6.3
(5)					2.4	16.2	22.3	3.7	5.3	5.0	2.0
(6)						18.1	20.2	6.1	7.6	7.4	3.3
(7)**							37.4	13.4	11.8	13.5	14.7
(8)**								25.8	27.4	26.7	23.3
(9)									1.7	1.6	3.7
(10)										1.9	4.8
(11)											5.2

Maximum = 37.4, Average = 9.8

1d) Idea Emphasis Score Distances Between Scale Item Profiles

Items	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	141	80	114	146	125	125	150	133	138	111	140
(2)		141	129	105	129	129	143	101	115	141	125
(3)			114	146	125	125	117	100	123	111	100
(4)				109	125	75	150	117	138	121	120
(5)					127	127	150	114	133	146	109
(6)						125	150	117	123	106	140
(7)							150	117	138	106	120
(8)								117	120	150	117
(9)									108	100	93
(10)										123	129
(11)											140

Maximum = 150.0, Average = 112.7

It was hypothesized that greater contextual distance between scale items would introduce uncertainty and confusion for the respondent about the meaning of scale items and that this would be expressed as unreliability of a sample in response to the scale. Since scales are designed to include different items covering facets of meaning of the concept being measured, E-score distances should show little relationship to reliability.

Data

Using standard references which present and evaluate social science scales (Miller

1977; Robinson and Shaver 1973; Mangen and Peterson 1982a; Shaw and Wright 1967; Mangen and Peterson 1982b), some 46 modest sized (3 to 25 item) Likert-type scales were selected. Scales which were included were those for which Cronbach alpha or split half reliabilities were available (see Appendix for a list of scales). An attempt was made to include scales with both high and low reliabilities but, as one might suspect, relatively few scales are published where reliabilities are consistently very low. Thus, one immediate limitation in this study is the relatively modest range of reliabilities (alphas ranging from .58 to .97; split half reliabilities ranging from .58 to .92). The overall sample provides two parallel tests, one for 15 scales (127 total scale items) with alphas and 31 scales (483 total scale items) with split half reliabilities (a total of 610 questions). In most cases the reliability coefficients on substantial test samples were relatively consistent and where several reliabilities were reported these were averaged for the purposes of this examination.

All questions from a scale were entered as separate text segments on a computer file for content analysis. The Minnesota Contextual Content Analysis (MCCA) program was utilized to generate two score profiles for each item in a scale as shown in the illustration, above (Tables 1c and 1d).

For purposes of this initial investigation, Euclidean distances computed between C-score profiles for scale items are used to measure the extent to which a scale is contextually consistent or inconsistent in the perspective suggested by the phrasing of scale items as a set. E-score distances were also computed between all pairs of items to measure differences in ideas emphasized among scale items. In both cases, a maximum distance for the set of items in a scale and the average distance between all pairs of scale items in a scale were used as independent variables to predict scale reliability scores. Since the number of items in a scale is also related to scale reliability, this was incorporated into the analysis as a control variable.

The data set for this analysis, then, consists of 46 cases or scales. The dependent variable is one of two reliability measures (15 of which have Cronbach alpha measures of reliability and 31 have split half reliability measures) for each scale. Other variables consist of the number of scale items, the two C-score distances between items in a scale (i.e. average and maximum distance across all items in a scale), and the two E-score profile distances between scale items.

Findings

As Table 2a indicates, there is a correlation between split-half reliability and the four measures of distance between items within a scale. The greater the distance, the lower the scale reliability. As predicted, these correlations are significant at the .05 level for both C-score distances (maximum and average distance) predicting split-half

reliabilities. This was not the case for E-score distances.

Table 2
CORRELATION OF RELIABILITIES AND MCCA SCORES

		Original Pearson Correlation	Partial Correlation Controlling No. of Items
a) Reported Split-Half Reliability for 31 scales, with:			
Number of Scale Items		.33 (p=.04)	--
Contextual Distances:	Maximum	-.38 (p=.02)	-.65 (p=.00)
	Mean	-.40 (p=.01)	-.50 (p=.00)
Idea Profile Distances:	Maximum	-.27 (p=.07)	-.57 (p=.00)
	Mean	-.05 (p=.23)	-.23 (p=.11)
b) Reported Cronbach's Alpha Reliability for 15 scales, with:			
Number of Scale Items		.11 (p=.34)	--
Contextual Distances:	Maximum	-.32 (p=.12)	-.43 (p=.06)
	Mean	-.33 (p=.12)	-.36 (p=.10)
Idea Profile Distances:	Maximum	-.10 (p=.36)	-.20 (p=.25)
	Mean	-.14 (p=.30)	-.22 (p=.23)

Since scale reliability is positively and significantly correlated with the number of items in a scale, partial correlations, controlling on number of items, were computed. Partial correlations were stronger and in the hypothesized direction. Partial correlations for C-score distances remained significant but, contrary to expectation, the partial correlation between maximum E-score distance and split-half reliability also became statistically significant. The average idea-profile distance remained non-significant, however.

A similar pattern of correlations resulted from an analysis of the 15 scales for which Cronbach alpha reliability coefficients were available (see Table 2b). In line with the initial hypothesis, correlations were stronger between the alpha reliability measure and contextual (C-score) distances and weaker for the relationship between reliabilities and the E-score distances. However, none of these correlations were statistically significant, possibly reflecting a much smaller sample size.

An Illustration of Scale Revisions

If contextual distance among scale items predicts lowered reliability, the next step would appear to be item revision to reduce this distance. The process can be illustrated using the *Parental Acceptance Scale* from Table 1. Items 7 and 8 are problematic

because they are at the greatest contextual distance from each other and from other items in the scale. The wording can be refined so that the questions express something of the same ideas but from a perspective that is closer to that expressed in the rest of the scale items (see context scores in Table 1b). Table 3 presents two revisions of item 7 (e.g. 7a and 7b) and four revisions of item 8 (e.g. 8a through 8d). In each case the C-scores for the revised statement are given. Items 7b and either 8b or 8d appear to be improvements in terms of contextual distances. Note that item 8a is a positive revision

Table 3
CONTEXT SCORES FOR REVISED SCALE ITEMS

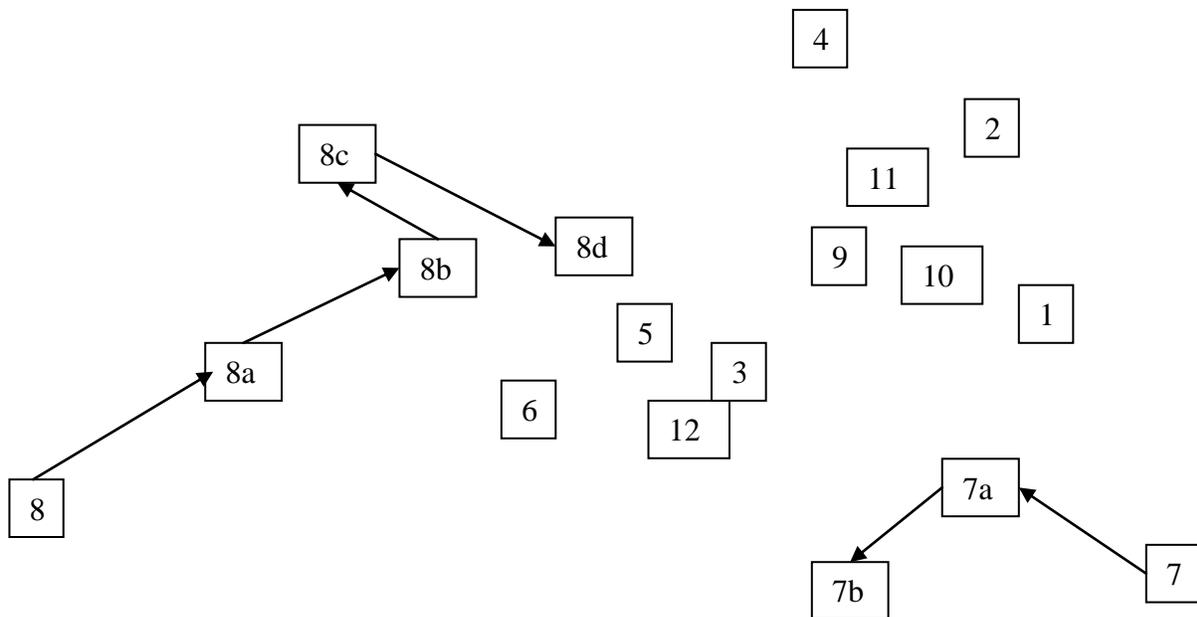
	Traditional	Pragmatic	Emotional	Analytic
<i>Revisions of Problematic Item #7:</i>				
<i>Original Item Wording.</i>				
7) "My parents made me feel wanted and needed."	9.95	- 3.26	15.05	-21.74
<i>Revised Statements.</i>				
7a) "My parents want and need me."	12.20	- 9.68	12.80	-15.32
7b) "I am wanted and needed by my parents."	13.13	-10.33	11.87	-14.67
<i>Revisions of Problematic Item #8:</i>				
<i>Original Item Wording.</i>				
8) "My parents forgot important events that I thought they should remember."	15.01	-25.00	3.83	6.16
<i>Revised Statements.</i>				
8a) "My parents remember events that I think they should remember."	13.41	-22.15	11.59	-2.85
8b) "My parents do not appreciate things that are important in my life."	10.45	-18.84	14.55	-6.16
8c) "My parents don't recall important things that happened to me in my life."	8.69	-20.44	16.31	-4.56
8d) "My parents do not appreciate important things that happened to me."	10.01	-16.54	14.99	-8.46

of the old item 8 which was stated negatively. While the context of the item is relatively similar to the original item, there are important differences. The revision, 8a, is expressed from a more emotional and less analytic perspective. This revision is more than merely a change from negative to positive but also a change in the context of the statement.

Using the KYST cluster analysis procedure (Kruskal and Wish, 1978), contextual distances between the 12 original parent acceptance scale items plus the 6 revisions are shown in the two-dimensional plot in Figure 1. This displays the contextual discrepancy of the original items 7 and 8 (also evident in the distance matrix in Table 1c) and shows how the revisions move problematic items 7 and 8 closer to the perspective expressed by the bulk of the scale items. It is predicted that a revised 12-item scale substituting the revision of items 7 and 8 which are contextually closer to the rest of the scale will result in improved reliability based on subject-response data.

Figure 1

CLUSTER PLOT OF CONTEXTUAL DISTANCES BETWEEN PARENTAL ACCEPTANCE SCALE ITEMS: Original Items and Revised Items



Conclusions

Work with contextual content analysis of text suggests that the greater the contextual distance within a text, the more difficulties there are in understanding what that text means. Furthermore, people have difficulty understanding each other (and finding a basis for communication) when they start out with a strongly different contextual

approach to a topic. Applied to the problem of forming reliable Likert scales, the prediction was that the greater the contextual distance in expression of different items within a scale, the lower the reliability of that scale. The hypothesized reason for this decreased reliability is the probability of greater confusion in interpreting the meaning of scale items which greater intrascale contextual distance creates for the respondent.

An analysis of this limited sample of 46 Likert scales (a total of 610 scale items) suggests that in general the hypotheses are confirmed even when the length of the scale is controlled. The exception is the lack of statistically significant results for the small subset of 15 scales for which Cronbach alpha coefficients were available, and the statistically significant partial correlation of maximum E-score profile distance and split-half reliability, controlling for number of items.

Attempts to increase reliability of a scale might focus on those divergent items contributing most to intra-scale contextual distance. Repeated MCCA analysis which included items with successively refined wording could be used to systematically reduce intra-scale contextual distance. This would appear to help in the process of scale development to improve the reliability of Likert scales.

This initial study focused upon intrascale statement-phrasing differences but other contextual distances might well have an impact upon scale reliability. Contextual distance could be computed between items in a scale as a whole and perspectives taken by different sub-cultural respondent groups. Context distance computed between scale wording and open-ended text on the same topic generated by a target population might help refine scale item wording to more adequately apply to specialized cultural sub-groups and consequently improve the reliability of a scale as a whole for these groups.

Further research on the consequences of social context differences for scale reliability should be pursued using larger samples of scales with a broader range of reliabilities.

Notes

I am indebted to Susan Schrader for editorial assistance and to Carol Pogue for helping to identify scales, enter these for content analysis, and summarize initial results. The analysis reported here was partially supported by grants from Rural Sociology and the University Computer Center.

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, 9208 Gue Road, Damascus, MD 20872) or web site, <http://www.clres.com>.

For information on MCCA, write: Donald G. McTavish, Department of Sociology, University of Minnesota, Minneapolis, MN 55455. E-mail address: mctavish@atlas.socsci.umn.edu.

References

- Berger, P.L., & Luckmann, T. (1966). The Social Construction of Reality. Garden City, NY: Doubleday.
- Carmines, Edward G. and Richard A. Zeller (1979). Reliability and validity assessment, Sage University Paper 17, Newbury Park, Sage Publications
- 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: Wiley.
- Erickson, J. (1986). "The Elements of Character: Mission Content of American Youth Organizations since 1880." Paper presented to the annual meeting of the Midwest Sociological Society, Des Moines, March.
- Furnham, A., & Henderson, M. (1982). "A Content Analysis of Four Personality Inventories." Journal of Clinical Psychology 38: 818-825.
- Gage, M.G. (1988). Family Careers and Companion Animal Experience: A Study of Anticipatory Socialization. Minneapolis, MN: CEN/SHARE, University of Minnesota.
- Gage, M. G., & Magnuson-Martinson, S. (1988). "Intergenerational Continuity of Attitudes and Values about Dogs." Anthro Zoos 1: 232-239.
- Goffman, Erving (1978). Frame analysis, New York; Harper and Row.
- Kelle, Udo, editor (1995). Computer-aided qualitative data analysis: theory, methods and practice, Thousand Oaks, Sage Publications.
- 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 11, Beverly Hills, Sage Publications.
- Labaw, P. (1980). Advanced questionnaire design. Cambridge, Abt Books
- Litkowski, Kenneth C., (to appear, 1997a). "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).
- Mangen, D. J., & Peterson, W.A. (eds). (1982a). Research Instruments in Social Gerontology, Volume 1: Clinical and Social Psychology. Minneapolis, MN: University of Minnesota Press.
- _____. (1982b). Research Instruments in Social Gerontology, Volume 2: Social Roles and Social Participation. Minneapolis, MN: University of Minnesota Press.
- McDonald, C., & Weidetke, B. (1979). "Testing Marriage Climate." MA theses, University of Iowa, Iowa City, IA.
- McTavish, D. G.. (to appear, 1997). "Scale Validity: A Computer Content Analysis Approach." Social Science Computing Review.
- 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, D. G., & Felt, D. (1984). "Nursing Home Environments: Managed Contexts for Older People." Sociology of Rural Life 7: 1 (Fall).
- McTavish, D. G., & Pirro, E. B. (1990). "Contextual Content Analysis." Quality and Quantity 24: 245-265.
- Miller, D. C. (1977). Handbook of Research Design and Social Measurement, 3rd Edition. New York: Longman.
- Morris, C. (1946). Signs, Language and Behavior. New York: Prentice-Hall.
- Pierce, J. L., McTavish, D.G., & Knudsen, K.J. (1986). "The Measurement of Job Characteristics: A Content and Contextual Analytic Look At Scale Validity." Journal of Occupational Behavior 7: 299-313.
- Rohner, R. P. (1986). Handbook for the Study of Parental Acceptance and Rejection. Storrs, CT: University of Connecticut Center for the Study of Parental Acceptance and Rejection.
- Rosenberg, M. (1965). Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press.
- Robinson, J. P., & Shaver, P.R. (1973). Measures of Social Psychological Attitudes. Ann Arbor, MI: Institute for Social Research.
- St. Clair, R. N., & Giles, H. (eds.) (1980). The Social and Psychological Contexts of Language. Hillsdale, NJ: Lawrence Erlbaum.
- Shaw, M. E., & Wright, J.M. (1967). Scales for the Measurement of Attitudes. New York: McGraw-Hill.
- Weber, R. P. (1990). Basic Content Analysis, 2nd Edition. Beverly Hills, CA: Sage.
- Weitzman, Eben A. and Matthew B. Miles, (1995) Computer Programs for Qualitative Data Analysis: A software sourcebook. Thousand Oaks, Sage Publications.
- Wood, M. (1980). "Alternatives and Options in Computer Content Analysis." Social Science Research 9: 273-286.