

A STUDY OF THE RELATIONSHIP BETWEEN EVALUATION REPORT CHARACTERISTICS AND MEASURES OF USE IN TWO COUNTRIES

Craig Russon
Palmer College of Chiropractic
Davenport, Iowa

Abstract: This article reports the results of secondary analysis research done in order to determine if there were differences in evaluation utilization among major cultural groups. First, the evaluation utilization literature is reviewed. Second, original research and secondary analysis methods are discussed. Finally, the results of four cross-cultural canonical correlation analyses are presented. The results show that, among the American subjects, there was a strong correlation between evaluation utilization and a high-order construct called evaluation implementation. This relationship did not exist for the Mexico subjects. These results are discussed in a cross-cultural context and conclusions are drawn.

Résumé: Le présent article rapporte les résultats d'une analyse de recherche secondaire, qui a été produite dans le but de déterminer les différences existant lorsqu'une méthode d'évaluation est utilisée par différents groupes culturels. Premièrement, la littérature de la méthode d'évaluation a été révisée. Deuxièmement, les méthodes originales de recherche et les méthodes d'analyse secondaire sont discutées. Finalement, les résultats de quatre analyses de corrélation canonicales inter-culturelles sont présentés. Les résultats ont démontré que parmi les sujets américains une forte corrélation existait entre l'utilisation d'évaluation et un construit d'ordre supérieur appelé l'implantation d'évaluation. Ce lien n'existait pas en ce qui concerne les sujets mexicains. Ces résultats sont discutés dans un contexte inter-culturel, suivis d'une conclusion.

Evaluation utilization has been the subject of research for the last two decades. A recent review of the literature that has accumulated over that period of time was conducted by Cousins and Leithwood (1986). At the conclusion of their review, the authors put forward a research agenda containing, among other topics, the following suggestions for further research: Can the existence of the

hypothetical, higher-order categories—evaluation implementation and decision- or policy-setting characteristics —be verified? To what extent do factors vary in their impact on decision-makers' use of evaluation results? Which factors have the greatest impact? How do factors interact with one another to affect use? This article reports the results of secondary analysis research that was done in an effort to shed light on the above questions.

REVIEW OF THE LITERATURE

A review of the literature on evaluation utilization should logically begin with a paper presented by Carol Weiss at the 1966 meeting of the American Sociological Association. In her paper, Weiss called for the systematic study of evaluation utilization issues. She stated:

Better knowledge of what kinds of evaluation have an impact on decision-making, and under what conditions, should help to encourage more effective development of evaluation practice. (1972, p. 326)

Sometime later, Patton et al. (1977) published a classic on the subject of evaluation utilization. They interviewed decision-makers and evaluators from 20 federal health evaluation projects. Two major findings emerged from the study: first, that political and interpersonal factors influenced evaluation use; and second, that evaluation results were used by decision-makers but not always in the clear-cut and organization-shaking ways that evaluators anticipated.

The latter finding stimulated a flurry of research activity designed to identify the different types of evaluation utilization. Researchers identified three broad categories of utilization (Leviton & Hughes, 1981). The first type was instrumental or allocative use (Weiner, Rubin, & Sachse, 1977), the traditional type of use in which the audience implemented the evaluation report recommendations immediately to make programmatic changes. The second type was conceptual use (Rich, 1977). In this type of use, the audience read the evaluation report, stored the information, continued to gather data from various sources, and at some point in the future may act upon the evaluation information. The third type is symbolic or persuasive use (Braskamp & Brown, 1980; Pelz, 1978; Young & Comptois, 1979), which involved drawing on evaluation evidence in an attempt to convince others to support a political position, or to defend such a position from attack (Leviton & Hughes, 1981).

Many of the studies cited above were included in a comprehensive review of the evaluation utilization literature conducted by Cousins and Leithwood (1986). The review examined methodology, dependent variables, and independent variables from 65 empirical studies that took place from 1971 to 1985.

When these researchers examined the independent variables that have been used in past research, they identified two categories: characteristics of evaluation implementation and characteristics of decision- or policy-setting. The evaluation implementation category has associated with it six factors. These included:

Evaluation Quality: Characteristics of the evaluation process including sophistication of methods, rigor, types of evaluation model.

Credibility: Of the evaluator and/or the evaluation process, defined in terms of objectivity, believability, appropriateness of evaluation criteria, and so forth.

Relevance: Of the evaluation to the information needs of the decisionmaker(s) in terms of the purpose(s) of the evaluation and the organizational location of the evaluator.

Communication Quality: Clarity of reporting results to the evaluation audience(s) in terms of style, evaluator advocacy of the results, and breadth of dissemination.

Findings: Positive, negative, consistent with evaluation audience expectations, value for decisionmaking, and so forth.

Timeliness: In the dissemination of evaluation results to decisionmaker(s). (Cousins & Leithwood, 1986, p. 347)

Cousins and Leithwood incorporated the above factors and others into a preliminary framework of evaluation utilization. The framework was used for examining relationships among the factors and measures of evaluation use. They found that the relative influence of the factors varied as a function of the type of use.

METHODS

The methods used in the original research have been described in detail elsewhere; therefore, only a brief description has been given here. For further details, including a copy of the rating instrument, see Russon (1993).

In the original research, McGuire's (1985) communication matrix was the theoretical framework. This theory was used to prepare two simulated evaluation reports—one qualitative in nature and the other quantitative. A panel of evaluation and agricultural extension experts reviewed the reports using a subset of the Joint Committee standards (1981). Also, a rating instrument was prepared, and its content validity was tested using a structured Q-sort. The research materials were translated into Spanish and back-translated into English.

The research materials were pilot tested on a sample of 150 agriculture graduate students from the U.S. and 42 graduate students/college seniors from Mexico. Response rates of 42% and 100%, respectively, were obtained. Construct validity and reliability of the instrument were tested using factor analysis and Cronbach's alpha. The factors identified in the factor analysis corresponded well to the factors postulated in McGuire's matrix. The alphas, in general, were very high.

The fully developed research materials were administered to 100 mid- and high-level agricultural extension administrators from the U.S. and 100 from Mexico. Each subject rated both the qualitative and the quantitative report. Response rates of 89% and 85%, respectively, were obtained. The independent variables were culture of the subject and methodological approach. The dependent variables were a set of 32 items developed to measure eight dimensions in McGuire's theory. The data were analyzed using a doubly multivariate repeated measures MANOVA (Norusis, 1988).

The Secondary Analysis

In this study, four canonical correlations were performed: the first used U.S. subject ratings of a qualitative evaluation report; the second used Mexico subject ratings of a qualitative report; the third used U.S. subject ratings of a quantitative report; and the fourth used Mexico subject ratings of a quantitative report.

In each analysis, three items developed to measure McGuire's action dimension served as the set of dependent variables. "All Now" was an item developed to measure the respondent's intention to implement all of the simulated evaluation report recommendations immediately or, in other words, to engage in instrumental evaluation utilization. The items "Some Now and Some Later" and "All

Later” had similar interpretations. The latter two items were developed to measure the respondent’s intention to engage in conceptual evaluation utilization.

Twenty-eight items that were developed to measure the other seven dimensions in McGuire’s matrix were used as the set of covariates. It was found that many of these items could be grouped using the evaluation implementation factors (see Table 1) identified by Cousins and Leithwood (1986). All of the evaluation implementation factors identified by Cousins and Leithwood were represented except timeliness.

The data were analyzed using SPSS/PC+ V2.0 and SAS V6.07. When SPSS/PC+ 2.0 was used, alpha was set at .25 and the results were rotated using a Varimax rotation.

Table 1
The Grouping of Covariate Variables Using the Evaluation Implementation Factors of Cousins and Leithwood (1986)

Evaluation Implementation Factor	Covariate Variable
Evaluation Quality	Object Examination Information Sources Valid Reliable Appropriate Analysis
Credibility	Objectively Reported Relevant Important Data Scope
Relevance	Help Improve Quality Help Improve Suitability Help Improve Efficiency
Communication Quality	Presentation of Ideas No Abrupt Shifts No Redundancies Succinct
Findings	Direct Approach Recommendation 1 Recommendation 2 Recommendation 3

RESULTS

There are four subsections presented in this section. Each contains the results of the canonical correlation analysis for one of the groups named above.

U.S. Subjects/Qualitative Report

This analysis began by computing the canonical correlations and the sample canonical variate coefficient vectors. This information is presented in Table 2. The vectors contain the weights by which the scores were multiplied in order to maximize the relationship between the two sets of variables.

Table 2 shows that the canonical correlation between the two sets of variables (.82) was high. Taking the square of the canonical correlation gave the total variance explained by the function (.68). One function was all that was necessary to explain about two-thirds of the total variance.

In Table 3, the correlations between each sample canonical variate and its component variables are presented. These "structure coefficients" served as the basis for later interpretations.

The measure of instrumental use (All Now) had a correlation of .85 with its sample canonical variate. "Interesting," "Imaginative,"

Table 2
U.S. Subjects/Qualitative Report: Sample Canonical
Correlation and Canonical Variate Coefficient Vectors

Standardized Dependent Variables				Pi*	Standardized Covariate Variables						
U1	U2	U3			V1	V2	V3	V4	V5	V6	
a 1'	1.02	-.28	.59	.82	b 1'	.08	-.41	-.09	.09	.31	.41
						V7	V8	V9	V10	V11	V12
						-.22	-.14	.35	-.33	-.19	-.19
						V13	V14	V15	V16	V17	V18
						.01	-.24	.10	.25	-.41	.11
						V19	V20	V21	V22	V23	V24
						-.22	-.15	.44	.05	-.05	-.19
						V25	V26	V27	V28		
						.27	.11	-.08	.33		

“Valid,” “Reliable,” “Help Improve Quality,” “Help Improve Suitability,” “Help Improve Effectiveness,” “Help Improve Efficiency,” and “Rec’s 1-3” all had correlations of over .50 with their sample canonical variate.

The proportion of the total standardized sample variance explained by the variables in the dependent variable sample canonical variate was 34.28; that for the covariates was 20.29. From these statistics, it was inferred that the first sample canonical variate was a slightly “better” representative of its set than the second sample canonical variate was of its set.

Table 3
U.S. Subjects/Qualitative Report: Correlation between Dependent and Covariate Variables and Their Respective Canonical Variables

U1		V1	
All Now	.85	Presentation of Ideas	.13
Some Now & Some Later	.41	No Abrupt Shifts	.10
All later	.37	No Redundancies	.21
		Succinct	.41
		Direct Approach	.37
		Interesting	.60
		Novel	.57
		Imaginative	.51
		Vivid	.45
		Amount of Information	.27
		Not Overly Complex	-.30
		Not Overly Technical	-.13
		Object Examination	.36
		Information Sources	.36
		Valid	.53
		Reliable	.54
		Appropriate Analysis	.20
		Objectively Reported	.48
		Help Improve Quality	.53
		Help Improve Suitability	.53
		Help Improve Effectiveness	.59
		Help Improve Efficiency	.51
		Relevant	.48
		Important Data	.40
		Scope	.49
		Recommendation 1	.58
		Recommendation 2	.62
		Recommendation 3	.61

Lastly, multivariate and univariate tests of significance were conducted. The multivariate test of significance tested the null hypothesis that the covariance matrix equalled zero. If the null hypothesis is rejected, then univariate tests of significance are conducted to test the individual canonical correlations. Table 4 presents the results of a Hotelling's multivariate test of significance and univariate F-tests.

Table 4
Analysis of Variance for All Four Canonical Correlation Analyses

CANONICAL CORRELATION	F
U.S. Subjects/Qualitative Report	
Multivariate	1.77*
Univariate	
1	1.68*
2	1.11
3	.10
Mexico Subjects/Qualitative Report	
Multivariate	1.19
U.S. Subjects/Quantitative Report	
Multivariate	1.20
Mexico Subjects/Quantitative Report	
Multivariate	1.99

* $p < .05$

Table 5
Mexico Subjects/Qualitative Report: Sample Canonical Correlation and Canonical Variate Coefficient Vectors

Standardized Dependent Variables				Pi*	Standardized Covariate Variables						
U1	U2	U3			V1	V2	V3	V4	V5	V6	
<i>a 1'</i>	.73	-.78	-.07	.69	<i>b 1'</i>	-.34	.40	.44	.38	-.76	.38
						V7	V8	V9	V10	V11	V12
						-.06	-.14	.06	-.67	.21	-.16
						V13	V14	V15	V16	V17	V18
						.07	.12	-.38	-.34	-.07	.08
						V19	V20	V21	V22	V23	V24
						-.25	.17	.03	.41	.37	.21
						V25	V26	V27	V28		
						.01	.01	.03	.27		

The results of the analysis of variance did indeed lead to rejection of the null hypothesis. Therefore, the univariate F-tests were conducted, the results of which suggest that only the first function was nonzero.

Mexico Subjects/Qualitative Report

This analysis also began by computing the canonical correlation and the sample canonical variate coefficient vectors. This information is presented in Table 5.

Table 5 shows that the canonical correlation between the two sets of variables (.69) was moderately high. The total amount of variance explained by the function was a little less than half (.48).

In Table 6, the correlations between each sample canonical variate and its component variables are presented. These “structure coefficients” served as the basis for later interpretations.

The measure of instrumental use (All Now) and the measure of conceptual use (All Later) had correlations of over .50 with their sample canonical variate. “Rec 2” and “Non-technical” had respective correlations of .43 and -.40 with their sample canonical variate.

The proportion of the total standardized sample variance explained by the variables in the dependent variable sample canonical variate was 30.49; that for the covariates was 6.23. The first sample canonical variate explained a large proportion of the variance, whereas the second sample canonical variate explained almost none.

Lastly, a Hotelling’s multivariate test of significance was conducted. The results (see Table 4) showed that the multivariate test was not significant. Therefore, univariate F-tests were not conducted.

U.S. Subjects/Quantitative Report

As before, the analysis began by computing the canonical correlation and the sample canonical variate coefficient vectors. This information is presented in Table 7.

Table 7 shows that the canonical correlation between the two sets of variables (.75) was high. The total amount of variance explained by the function was more than half (.57).

In Table 8, the correlations between each sample canonical variate and its component variables are presented. These “structure coefficients” served as the basis for later interpretations.

The measure of instrumental use (All Now) had a correlation of $-.85$ with its sample canonical variate. “Interesting,” “Novel,” “Imaginative,” “Sufficiently Examined,” “Valid,” “Reliable,” “Help Improve Quality,” “Help Improve Suitability,” “Help Improve Effectiveness,” “Help Improve Efficiency,” “Important Data,” “Scope,” and “Rec’s 1-3” all had correlations of over $-.50$ with their sample canonical variate. Note that the signs consistently flipped in this analysis. This occurrence did not affect the interpretation.

Table 6
Mexico Subjects/Qualitative Report: Correlation between Dependent and Covariate Variables and Their Respective Canonical Variates

U1		V1	
All Now	.67	Presentation of Ideas	.09
Some Now & Some Later	.13	No Abrupt Shifts	.11
All Later	.67	No Redundancies	.22
		Succinct	-.01
		Direct Approach	-.09
		Interesting	.31
		Novel	.21
		Imaginative	.13
		Vivid	.22
		Amount of Information	-.21
		Not Overly Complex	.06
		Not Overly Technical	-.40
		Object Examination	.35
		Information Sources	-.12
		Valid	-.06
		Reliable	-.09
		Appropriate Analysis	.35
		Objectively Reported	.22
		Help Improve Quality	.15
		Help Improve Suitability	.29
		Help Improve Effectiveness	.27
		Help Improve Efficiency	.38
		Relevant	.31
		Important Data	.39
		Scope	.34
		Recommendation 1	.11
		Recommendation 2	.43
		Recommendation 3	.23

The proportion of the total standardized sample variance explained by the variables in the dependent variable sample canonical variate was 25.52; that for the covariates was 27.58. Both sample canonical variates were good representatives of their respective sets.

Lastly, a Hotelling's multivariate test of significance was conducted. The results (see Table 4) showed that the multivariate test was not significant, and therefore univariate F-tests were not conducted.

Mexico Subjects/Quantitative Report

Once again, the analysis began by computing the canonical correlation and the sample canonical variate coefficient vectors. This information is presented in Table 9.

Table 9 shows that the canonical correlation between the two sets of variables (.67) was moderately high. The total amount of variance explained by the function was less than half (.45).

In Table 10, the correlations between each sample canonical variate and its component variables were presented. These "structure coefficients" served as the basis for later interpretations.

The measure of instrumental use (All Now) and the measure of conceptual use (All Later) had respective correlations of .67 and -.53

Table 7
U.S. Subjects/Quantitative Report: Sample Canonical Correlation and Canonical Variate Coefficient Vectors

Standardized Dependent Variables			PI*	Standardized Covariate Variables											
U1	U2	U3		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
.75															
<i>a1'</i> -1.06	.23	-.60		<i>b1'</i> -.09	-.36	.07	.10	.15	-.22	-.08	-.30	-.01	.11	.22	.24
				V13	V14	V15	V16	V17	V18	.10	.02	-.01	-.12	-.24	.22
				V19	V20	V21	V22	V23	V24	.09	-.12	-.19	-.06	.37	-.23
				V25	V26	V27	V28			-.37	.09	-.32	-.07		

with their sample canonical variate. Of all the component variables, only "Rec 3" had a correlation of over .50 with its sample canonical variate. "Rec 1," "Rec 2," and "Defensible Sources" had correlations in the mid-40s.

The proportion of the total standardized sample variance explained by the variables in the dependent variable sample canonical variate was 25.67; that for the covariates was 8.87. Once again, the first sample canonical variate explained a large proportion of the variance and the second sample canonical variate explained almost none.

Table 8
U.S. Subjects/Quantitative Report: Correlation between Dependent and Covariate Variables and Their Respective Canonical Variables

U1		V1	
All Now	-.86	Presentation of Ideas	-.41
Some Now & Some Later	-.15	No Abrupt Shifts	-.38
All Later	-.02	No Redundancies	-.45
		Succinct	-.32
		Direct Approach	-.45
		Interesting	-.66
		Novel	-.56
		Imaginative	-.60
		Vivid	-.53
		Amount of Information	-.34
		Not Overly Complex	.03
		Not Overly Technical	.08
		Object Examination	-.60
		Information Sources	-.46
		Valid	-.64
		Reliable	-.55
		Appropriate Analysis	-.48
		Objectively Reported	-.44
		Help Improve Quality	-.64
		Help Improve Suitability	-.55
		Help Improve Effectiveness	-.63
		Help Improve Efficiency	-.66
		Relevant	-.49
		Important Data	-.59
		Scope	-.65
		Recommendation 1	-.61
		Recommendation 2	-.72
		Recommendation 3	-.52

The high canonical correlation between the index of evaluation utilization and the index of evaluation implementation suggested that American audiences are more likely to engage in instrumental and conceptual utilization if the qualitative report meets a criterion for evaluation implementation that includes evaluation quality, credibility, and the like.

The interpretation of the results from the U.S. subjects/quantitative report analysis was similar to that put forward above. However, this interpretation should be viewed somewhat skeptically because the multivariate test of significance was not significant.

Table 10
Mexico Subjects/Quantitative Report: Correlation between Dependent and Covariate Variables and Their Respective Canonical Variables

U1		V1	
All Now	.67	Presentation of Ideas	.19
Some Now & Some Later	-.53	No Abrupt Shifts	.18
All Later	.18	No Redundancies	.02
		Succinct	.26
		Direct Approach	.36
		Interesting	.23
		Novel	.01
		Imaginative	.34
		Vivid	.35
		Amount of Information	-.21
		Not Overly Complex	.17
		Not Overly Technical	.24
		Object Examination	.20
		Information Sources	.46
		Valid	.25
		Reliable	.36
		Appropriate Analysis	.30
		Objectively Reported	-.10
		Help Improve Quality	.30
		Help Improve Suitability	.24
		Help Improve Effectiveness	.27
		Help Improve Efficiency	.27
		Relevant	.14
		Important Data	.39
		Scope	.35
		Recommendation 1	.46
		Recommendation 2	.45
		Recommendation 3	.52

The results from the Mexico analyses did not yield any straightforward interpretations. Examination of the structural coefficients from both analyses led the researcher to conclude that the sets of dependent variables were both indices of evaluation utilization. However, in both analyses, the proportion of total standardized sample variance explained by the sets of covariate variables was very low. This suggested that, in both analyses, the second sample canonical variate was not an adequate index of anything. This finding raises the question: If evaluation utilization by the Mexico subjects did not correlate with evaluation implementation factors as it did for the American subjects, with what would it correlate?

CONCLUSION

The following conclusions were drawn from this study: (1) It supports the high-order category of evaluation implementation hypothesized by Cousins and Leithwood (1986); (2) one factor does not appear to have a greater role than another in the makeup of the evaluation implementation construct; (3) evaluation use in the U.S. was highly correlated with the evaluation implementation construct; and (4) there were cultural differences in the factors that influence evaluation utilization. However, more research is required to describe these differences.

ACKNOWLEDGEMENTS

The author gratefully acknowledges Laura Koehly and Charles Henderson for their comments on a draft of this article.

REFERENCES

- Braskamp, L., & Brown, R. (1980). *Utilization of evaluative information*. San Francisco: Jossey-Bass.
- Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Joint Committee on Standards for Educational Evaluation. (1981). *Standards for evaluations of educational programs, projects, and materials*. New York: McGraw-Hill.

- Leviton, L., & Hughes, E. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5(4), 525–548.
- McGuire, W. (1985). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (pp. 233–346). Reading, MA: Addison-Wesley.
- Norusis, M. (1988). *SPSS/PC+ advanced statistics V2.0*. Chicago: SPSS.
- Patton, M., Grimes, P., Guthrie, K., Brennan, N., Grench, B., & Blyth, D. (1977). In search of impact: An analysis of utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policymaking*. Lexington, MA: Heath.
- Pelz, D. C. (1978). Some expanded perspectives on use of social science in public policy. In J. Yinger & S. Cutler (Eds.), *Major social issues: A multidisciplinary view*. New York: Free Press.
- Rich, R. (1977). Uses of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C. Weiss (Ed.), *Using social research in public policy making*. Lexington, MA: Heath.
- Russon, C. (1993). Communication of evaluation results: A cross-cultural perspective (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1993). *Dissertation Abstracts International*, 54-05, 1778A.
- Weiner, S., Rubin, D., & Sachse, T. (1977). *Pathology in institutional structures for evaluation and a possible cure*. Stanford, CA: Stanford Evaluation Consortium.
- Weiss, C. (1972). Utilization of evaluation: Toward comparative study. In C. Weiss (Ed.) *Evaluating action programs: Readings in social action and education*. Boston: Allyn & Bacon. (Reprint of a paper presented at the meeting of the American Sociological Association, Miami Beach, FL, September 1966).
- Young, C., & Comptois, J. (1979). Increasing congressional utilization of evaluation. In F. Zweig (Ed.), *Evaluation in legislation*. Beverly Hills: Sage.