On the Appropriateness of Certain Algorithms for Analyzing Needs Assessment Data: A Reaction to Cummings' (1985) Study

The following is a reaction by Earl R. Misanchuk to the article "Comparison of Three Algorithms for Analyzing Questionnaire-Type Needs Assessment Data to Establish Need Priorities" by Oliver W. Cummings that was published in Volume 8, Number 2 of JID.

Earl R. Misanchuk
The University of Saskatchewan
Saskatoon, Saskatchewan S7N 0W0

Cummings (JID, 1985, *8*(2), pgs. 11-16), using what has sometimes been called the "gap definition" of educational need—the discrepancy between an existing state of affairs and a desired or ideal state—undertook to study three different ways of analyzing educational needs assessment data. Cummings' study is gratifying to the extent that it is an indication that the profession of instructional development is maturing to the point where alternative paradigms can be pitted against one another (as well as against reality), and to the extent that it attempts to deal with some of the more thorny questions in needs assessment, using data generated in a real instructional development situation. The basic idea underlying the article is an excellent one—conducting empirical research on alternative methods, to guide the practitioner in selecting appropriate methods of analysis. There are a number of issues associated with his study, however, upon which I feel obliged to comment. Before launching into the issues, a couple of general comments about the study are in order.

One of the awkward results of a comparative study such as Cummings' is that when it is completed one still does not have a full answer to the question "What method(s) should I use?" Since the conclusions of the study are based upon correlations between ratings established by different analytical methods, all one is left knowing at the end is which methods agree most with which other methods. Unless there is some external criterion by which to judge which method is the "true" or "correct" one, there seems little point to the exercise. It could possibly be, for example, that the method which agrees least with the others is, in fact, the correct one—the one that best represents reality. Hence the determination of which analytic method to use should be based on some grounds other than comparative ones, unless one of the objects of comparison is known to be better than the others. In this case, superiority of method can only be determined through logical analysis, accounting for how many difficulties one method ameliorates, relative to the other.

The conclusions Cummings is able to reach are, at best, equivocal, and even he admits to their tentative nature, citing limitations of the data set used in the study. The equivocation is, in fact, less a function of the data used in the study than it is a function of the methodology used.

As a second general comment, it should be noted that despite Cummings' claim to the contrary, the statistic ∇_N (del) that he uses as one of the bases of this study is not, in fact, the same as the one proposed as the basis for a two-component proportionate-reduction-in-error (PRE) index of educational need (Misanchuk, 1982; 1984b), or the subsequent multi-component PRE index (Misanchuk, 1984a). Cummings postulates a quite different definition of high educational need (p. 11) than the original ∇_N (Misanchuk, 1984b, p.29). The dimensions RELEVANCE and COMPETENCE (Misanchuk, 1982; 1984b) are most certainly not the same as Cummings' DOES and SHOULD. While the dimension of COMPETENCE approximates the DOES dimension, there is only a remote relationship between RELEVANCE and SHOULD. One could point out a number of differences—SHOULD is prescriptive, while RELEVANCE is descriptive, for example—but that would skirt the main issue: The two questions used to generate the data, "What should be?" and "How relevant is the given skill to your job?" are fundamentally different. The PRE statistic was not invented to do the job Cummings attempts to have it do in his study. This is not to say that it is illegitimate for Cummings to propose the method he does, as an adaptation of ∇_N. In fact, some recent thinking on the nature of educational needs (Misanchuk, 1985) may share more common ground with Cummings' current application than with the original application. Nevertheless, it should be clear to the reader that Cummings' application is, in fact, an adaptation rather than an application of the original construct ∇_N. Because I will have occasion to refer to both the original formulation and Cummings' adaptation in this article, I propose to retain the original notation, ∇_N for the original, and use the notation ∇_C for Cummings' adaptation.

Mean Difference Analysis as an Analytic Technique

Cummings correctly identifies some of the problems associated with Mean Difference Analysis (MDA): it requires an (indefensible) assumption that the rating scales yield interval data, and it fails to take into account differences in the absolute importance level of the variable (p. 13). There is at least one other. In order to apply the MDA technique, an implicit assumption must be made that the variance of responses across all stimuli (content areas, in the case of needs assessment) is uniform, and that assumption can rarely be met in the real world. Ignoring the assumption could lead to erroneous conclusions.

To illustrate, suppose two items similar to the ones shown in Cummings'

Figure 1 had the distribution of responses for "Knowledge Level That Does Exist" shown in Table 1 below.

Table 1
Hypothetical distribution of responses to two items with same mean and different variances.

	Almost None 1	2	3	4	Extensive 5
Item					
1	0	1	28	1	0
2	14	1	0	1	14

Although both items have the same mean, the variances are vastly different. The information contained in the statistical concept of variance is clearly important to the instructional developer. Given a constant pattern of responses, for both items, to the corresponding "Knowledge Level That Should Exist" dimension, the data illustrated above would indicate that while relatively little need exists for education or training on content Item 1, there is a substantial pool of individuals for whom Item 2 is a high need (and an equally substantial pool of individuals for whom Item 2 has almost no associated need). For the second situation, it would behoove the instructional developer not only to plan for the reduction of the educational need for that subgroup that indicated high need, but also to provide alternative plans for the other sub-group, members of which would undoubtedly be uninterested in the instruction provided the first group. Schwier (1982), one of the first to formalize the application of needs assessment data to the process of instructional design, is a source of additional examples. (The problem of variance has been recognized for some time. Misanchuk and Scissons (1978) attempted to account for the item variance by converting item means to z-scores, using the pooled inter-item variance, and employing arbitrary cutting points, but were less than successful in accounting for the problem.)

The criticisms of the MDA technique identified above are sufficiently serious that it ought not to be used, no matter how simple it is to calculate or how easy Cummings claims it is to ". . . describe to an audience that is not well versed in or is even skeptical of statistical analysis" (p. 13). If, in fact, it should not be used, then comparisons involving MDA are spurious.

Complexity and Communication

In enumerating the advantages and disadvantages of the various methods used in the study, Cummings states that the ∇_C statistic is "conceptually and computationally complex and therefore more difficult to "sell" to some audiences" (p. 13). The theory of relativity, the concept of entropy, Maxwell's equations, and factor analysis are all either conceptually or computationally complex or both. Should we advocate that they not be applied for those reasons? It is true that there is a principle in science that of two equally potent theories, the simpler one is the preferred. However, to apply that principle first requires that the competing theories be equally good at explaining reality.

That ∇_C is more conceptually and computationally complex than MDA is undeniable. Whether or not it is actually conceptually and computationally more complex than Cummings' suggested weighted needs index (WNI) is a matter I will return to later. But focussing on those criteria is missing the point. The appropriate question to be asking is "Does ∇_C do a better job of describing the real world than any competing analytic technique?" Whether the underlying concepts are immediately and intuitively understood by either clients or other laypersons—or indeed even by professionals—is of lesser concern.

Cummings notes that

> . . . the needs assessment analyst must sometimes communicate with an audience that is neither well-versed in nor impressed with statistics. This audience may be skeptical of assumptions about "monotonically increasing marginal probability" or "unit linear progression," but might embrace the logic which suggests that small discrepancies for important skills may need to be addressed, whereas relatively larger discrepancies for less important skills may not require attention (p. 12).

He appears to be implying that the WNI he describes subsequently does a better job of addressing the latter part of the quoted statement than does the ∇_C statistic. In fact, everything in the latter part of his statement is true for ∇_C. What Cummings neglects to mention is that the WNI must make some assumptions, too. The very pattern of "assigned values" in his Figure 3 makes inherent assumptions that are not explained anywhere, a point to which I will return later. Cummings seems to be confusing the necessity for making assumptions that underly an analytical technique with the necessity for telling a lay audience what those assumptions are—in gory detail.

In the practice of any profession, there is a language (or jargon) that professionals use to communicate with one another, and a second language they use to communicate with lay people. It is the professional's responsibility to act as translator, communicating as much or as little as is necessary to satisfy the needs of the clients. If doctors and lawyers feed their clients technical terminology, their clients demand explanations in plain language, or seek the counsel of others. So it can, and should, be with needs assessment.

The del statistic can quite simply be introduced to a lay audience or readership as an index of educational need. The professional should expect to exercise professional judgement in selecting the appropriate statistic to describe the phenomenon. Thus, the professional should select the appropriate statistic, based on his or her understanding of the underlying theory as it coincides with the situation at hand, then present the results and conclusions (not necessarily the underlying theory!) to the client in language the client understands. By way of example, pollsters have attempted to solve this kind of a problem by explaining probabilities associated with their findings in terms such as "correct to within two percentage points nineteen times out of twenty," rather than referring to standard errors of measurement and alpha levels. Instructional developers should be able to find ways of explaining the del statistic that are easily understandable to almost any kind of an audience—after all, their professional function is to help learners understand complex material.

If there is concern that members of the profession may also have difficulty in following the development of the concepts underlying the statistic, then that is a matter for both preparatory and continuing professional education. The advent of new tools in a profession invariably necessitates the re-focussing and the fine-tuning of professional skills. (The impact of digital electronics on all fields of engineering might be an illustrative, albeit profoundly more powerful, example.) Furthermore, there is probably a point beyond which a practioner need not venture in the world of theory. For example, one might not completely understand exactly how the statistical probabilities of the values of chi-square found in tables are determined, but may be able to make valid use of the statistic nevertheless. To condone the use of a demonstrably inferior theoretical construct on the basis that it is easy to understand would be professionally inappropriate.

Returning to the point that Cummings' WNI is conceptually and computationally simpler than ∇_C, one wonders whether the simplicity may be more apparent than real. The derivation of ∇_N published (Misanchuk, 1982, 1984b) was complicated by the fact that is was necessary to show how that statistic related to, and differed from, the more general statistic from which it was adapted—Hildebrand, Laing, and Rosenthal's (1977a, 1977b) del. ∇_N may furthermore appear conceptually complex because the assumptions are explained. As pointed out earlier, WNI makes some assumptions too; Cummings, however, does not identify them, nor explain why they were made. For example, in his Figure 3, the suggested distribution of cell values presumably has some rationale, but the reader is left guessing as to what that rationale might be. Why, for example, is cell (2,4) assigned a value of 2, and not 3, or some other value? On what basis is the change in cell values equal to 1, going from cell (1,5) to (the adjacent) cell (2,5), but equal to 3 in going from cell (1,5) to (the also adjacent) cell (2,4)?

Cummings does make one assumption explicit—that " . . . the relative ratings of SHOULD and DOES show no need for training development (from an organization point of view) if . . . the SHOULD rating fails to meet a test of 'moderate to high' in conjunction with 'moderate to low' DOES" (p. 13). He assumes, in effect, that any data outside the upper right hand corner of the matrix in his Figure 3 are relatively unimportant, and can be disregarded.

Cummings' position on this matter is similar to that taken by LeSage (1980), in his development of the dichotomized-additive coefficient. Misanchuk (1980) previously showed why that position is untenable, but another example will illustrate the point. Suppose two items differed in their distribution of responses as shown in Figures 1(a) and 1(b). In 1(a), according to Cummings' suggested procedure, no need is evidenced. Yet obviously, some need exists—not much, perhaps, but some. Certainly, more need is suggested by 1(a) than by 1(b). By essentially throwing away data, the distinction between 1(a) and 1(b) would be lost to the instructional developer. To be sure, the argument is almost academic, however, since both items exhibit relatively low need. On the other hand, consider the information contain-

1(a)

DOES \ SHOULD	1	2	3	4	5
1					
2					
3		2	25		
4		1	2		
5					

1(b)

DOES \ SHOULD	1	2	3	4	5
1					
2					
3		5	15		
4		5	5		
5					

1(c)

DOES \ SHOULD	1	2	3	4	5
1					
2					
3					
4	2	1			
5	25	2			

Figure 1. Hypothetical distributions, equally regarded as showing no need according to the Weighted Needs Index.

ed in Figure 1(c), as compared with that in 1(b). Using the same argument as presented against MDA through the example in Table 1, not capitalizing on the information available, and regarding the distributions of Figure 1(b) and 1(c) equivalent, could have significant negative instructional consequences.

As an aside, it is perfectly possible to postulate and apply the equivalent of the set of weights Cummings recommends, under the PRE approach. Cummings makes the point that the ∇_C approach ". . . does not treat responses above or below any given threshold differently" (p. 13). In fact, the substitution of different error weights could accomplish just that (subject, of course, to the concerns expressed in the last paragraph). The weights postulated by Misanchuk (1982, 1984b) are not sacred, although they do have an explicated logic underlying them. Suggesting a different set of weights simply requires a clearly explicated, persuasive argument as to why the new weights better reflect reality than those originally suggested.

What Cummings appears to be attempting to grapple with is the question of adequacy. Most approaches to needs analysis, including ∇_N, eschew the question. At best, they are capable of producing a relative ranking of needs, without any consideration for whether or not the learners are already good enough at certain tasks to forgo further training. Different analytic methods approach the problem in different ways—the typical discrepancy approach examines the dimensions SHOULD and DOES, while the PRE approach examines the dimensions RELEVANCE and COMPETENCE—but the questions of a criterion of adequacy is largely ignored. Recently, however, Misanchuk (1985) posed for discussion an approach that involves examining the dimensions RELEVANCE and COMPETENCE in those cases where the learner is below the criterion of adequacy, and the dimensions COMPETENCE and DESIRE in those cases where the learner is at or beyond that criterion. In summary, it is heartening to see Cummings attempt to address the too-long-ignored question of adequacy, but the solution he poses—ignoring everything but the upper-right-hand corner of the data matrix—falls short of the mark.

Let's return now to the point of computational complexity. Without having collected empirical evidence on the matter, I would venture to guess that for most people, the difference in computational complexity between the WNI and ∇_C would not be an issue. Both formulas require the summing of products across two dimensions. The elimination of the subtraction contained in the ∇_C statistic is trivial. With computers becoming increasingly commonplace in the professional's toolbox—including Cummings' own (p. 14)—it is difficult to see why the issue of computational complexity is raised at all.

Similarity and Difference

Examination of the two formulas leads to the last point I wish to make: the WNI is really very similar to ∇_C. As Cummings notes, the acceptance of the assumptions associated with ∇_C—the set of error weights and marginal values proposed by Misanchuk (1982, 1984b)—causes the denominator of ∇_C to reduce to a constant. Linear transformations such as division by a constant or subtraction from a constant do not affect the relative value of a statistic in any salient way, although they do change the actual value. Cummings posits a set of "cell values"—which are functionally equivalent to error weights—and he proposes using cell frequencies rather than proportions—but also divides them by N, the total number of respondents, which of course ultimately yields cell proportions. The changes that Cummings suggests in moving from ∇_C to the WNI—especially the elimination from consideration of two-thirds of the cells in which data could potentially accrue—are non-productive at best and counter-productive at worst. Finally, since ∇_N—and by analogy ∇_C—are adaptations of a statistic which is quite well developed and understood (Hildebrand, Laing, & Rosenthal, 1977a, b), there are already available ancillary statistics (standard error, tests of significance, etc.) that could prove useful in the analysis of educational and training needs. Analogs for the WNI presumably do not yet exist.

Conclusion

In mis-applying ∇_N to the SHOULD-DOES situation, Cummings may very well have inadvertently discovered a legitimate application of PRE analysis. In using that mis-application, ∇_C as a "straw man" for comparison with the WNI, however, he focussed on the wrong features of the statistic. If it could be shown, through germane, rational argument—including the explication of assumptions—that either the WNI or ∇_C is appropriate for the analysis of discrepancy needs analysis data, then a real contribution would have been made. (The appropriateness of the discrepancy conceptualization of educational and training needs, as opposed to the needs components conceptualization, is a separate matter entirely.) Focussing on characteristics of the statistic that are of secondary importance and ignoring implicit underlying assumpions in the name of convenience and simplicity of communication, however, do little to further the advance of needs analysis. Until persuasive, germane arguments are forthcoming, there seems little reason to discard ∇_N in favor of either the WNI or ∇_C.

References

Cummings, O. W. (1985). Comparison of three algorithms for analyzing questionnaire-type needs assessment data to establish need priorities. *Journal of Instructional Development, 8(2), 11-16.*

Hildebrand, D. K., Laing, J. D., & Rosenthal, H. (1977a). *Analysis of ordinal data.* Sage university paper series on quantitative applications in the social sciences, no. 8. London: Sage.

Hildebrand, D. K., Laing, J. D., & Rosenthal, H. (1977b). *Prediction analysis of cross classifications.* New York: Wiley.

LeSage, E. C., Jr. (1980). A quantitative approach to educational needs assessment. *Canadian Journal of University Continuing Education, 6*(2), 6-13.

Misanchuk, E. R (1980). A methodological note on quantitative approaches to needs assessment. *Canadian Journal of University Continuing Education, 7*(1), 31-33.

Misanchuk, E. R. (1982, May). *The analysis of multi-component training needs data.* Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Dallas, TX. (ERIC Document Reproduction Service No. ED 274 409)

Misanchuk, E. R. (1984a, January). *The analysis of multi-component educational and training needs data: The multivarate case.* Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Dallas, TX. (ERIC Document Reproduction Service No. ED 247 409)

Misanchuk, E. R. (1984b). The analysis of multi-component educational and training needs data. *Journal of Instructional Development, 7(1), 28-33.*

Misanchuk, E. R. (1985). *Toward a re-conceptualization of educational needs and wants.* Manuscript submitted for publication.

Misanchuk, E. R., & Scissons, E. H. (1978). *The Saskatoon business training needs identification study: Final report.* Saskatoon, SK: University of Saskatchewan, Extension Division.

Schwier, R. A. (1982, April). *Applying multi-component training needs information.* Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Dallas, TX.