

## DELIBERATION IN CURRICULUM CONFERENCES

MARTIN MULDER, *University of Twente*

In a recent article, Reid draws attention to the institutional context of curriculum deliberation.<sup>1</sup> He stresses that educators neglect one of Schwab's commonplaces, the milieu, or according to Reid, the institutional context of the debate on curriculum issues.<sup>2</sup> "Whatever success has been achieved, and I believe it is considerable, in raising the quality of deliberation in schools and among teachers, we must admit that educators have so far failed to address questions of the quality of curriculum deliberation involving broader communities."<sup>3</sup>

Processes of the design, approval, and implementation of a core curriculum, however, seem to move curricular decision making from the professional island (classrooms and curriculum meetings) toward public arenas (advisory bodies and consultative committees). In this post-industrial information era, we can easily see that curriculum decisions formerly the privilege of the *maître* now require consensus of the people and institutions involved. Clark argues that many actors are making critical curricular decisions, and he distinguishes 10 categories of decision makers: the public, political leaders, textbook publishers, testing authorities, the media, higher education personnel, professional organizations, central administration, teacher groups, and individual teachers.<sup>4</sup>

One premise of curriculum conferences is to meet the need for participative debate on curriculum issues at the district or national level. A curriculum conference is a specific, practical strategy for deliberative curriculum problem solving. Walker positions the curriculum conference as a strategy for agenda

---

<sup>1</sup>William A. Reid, "The Institutional Context of Curriculum Deliberation: A Research Agenda," *Journal of Curriculum and Supervision* 4 (Fall 1988): 3-16.

<sup>2</sup>Joseph J. Schwab, "The Practical: A Language for Curriculum," in *Science, Curriculum, and Liberal Education. Selected Essays*, ed. Ian Westbury and N. J. Wilkhof (Chicago: University of Chicago Press, 1978), pp. 287-321.

<sup>3</sup>William A. Reid, "The Institutional Context of Curriculum Deliberation: A Research Agenda," *Journal of Curriculum and Supervision* 4 (Fall 1988): 9.

<sup>4</sup>Richard W. Clark, "Who Decides? The Basic Policy Issue," in *Critical Issues in Curriculum*, 87th Yearbook of the National Society for the Study of Education, ed. Laurel N. Tanner (Chicago: University of Chicago Press, 1988), pp. 175-204.

building in curriculum policymaking.<sup>5</sup> In this nominal group approach, 10 to 20 participants try to agree on a curriculum within a few days.

Several studies have addressed aspects of this strategy. The first curriculum conference, on microprocessors and their implications for education, was held at the Institute of Science Education in the Federal Republic of Germany.<sup>6</sup> In 1983, curriculum conferences were held on inservice teacher training and control engineering in Switzerland. More curriculum conferences followed, in 1984 on computer science, in 1985 a second one on inservice teacher training, in 1987 on robotics, and in 1988 on personal computing. Several curriculum researchers in the Netherlands picked up this approach and adapted it to national and institutional contexts. From 1984 onwards, curriculum conferences have addressed curriculum innovation in mechanical engineering, office technology and sales education,<sup>7</sup> production engineering,<sup>8</sup> office automation,<sup>9</sup> service engineering,<sup>10</sup> and retail trade and administration.<sup>11</sup>

During the second half of the 1980s, inquiries started to analyze the processes of curriculum conferences. These inquiries resulted in several publications on process reconstructions, motives, expectations, and experiences of participants,<sup>12</sup> analysis and determination of performance requirements,<sup>13</sup> participation and decision making,<sup>14</sup> and reliability testing of an argumentation analysis and observation analysis instrument.<sup>15</sup>

Until now, deliberation remained predominantly a theoretical concept, which Walker, Connelly, Fox, Westbury, Wick, Reid, Eisner, Pereira, Roby,

<sup>5</sup>Decker F. Walker, *Fundamentals of Curriculum* (San Diego: Harcourt Brace Jovanovich, 1990).

<sup>6</sup>Karl Frey, *Curriculum-Konferenz Gebiet Mikroprozessor* (Kiel: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, 1981).

<sup>7</sup>W. J. Nijhof and Martin Mulder, *Basissvaardigheden in het beroeps onderwijs* ('s-Gravenhage: Instituut voor Onderzoek van het Onderwijs, 1986).

<sup>8</sup>J. M. Pieters and Martin Mulder, eds., *Productie-automatisering Een onderzoek naar curriculum en instructie in het Middelbaar Technisch Onderwijs* (Enschede: Universiteit Twente, Faculteit der Toegepaste Onderwijskunde, 1988).

<sup>9</sup>Martin Mulder and J. Van Lent, *Kantoorautomatisering Een onderzoek naar leerplanontwikkeling* (Lisse: Swets and Zeitlinger, 1988).

<sup>10</sup>Martin Mulder and J. Botke, *Het ontwerpen van een klantencursus* (Enschede: Universiteit Twente, Faculteit der Toegepaste Onderwijskunde, 1988).

<sup>11</sup>Anke Thijsen, M. De Wit, and H. Pollen, *Beroepsgerichte volwasseneneducatie ontwerpen met behulp van curriculum conferenties* (Enschede: Universiteit Twente, Faculteit der Toegepaste Onderwijskunde, 1989).

<sup>12</sup>Martin Mulder, W. J. Nijhof, and J. G. L. Remmers, "An Exploration of the Curriculum Conference" (paper presented at the annual meeting of the American Educational Research Association, Washington, DC, April 1987).

<sup>13</sup>W. J. Nijhof and Martin Mulder, "Performance Requirements Analysis and Determination, in *Developing Skills with Information Technology*, ed L. Bainbridge and S. A. Ruiz Quintanilla (London: John Wiley & Sons, 1989), pp. 131-152.

<sup>14</sup>Martin Mulder and Anke Thijsen, "Decision Making in Curriculum Conferences: A Study of Convergence of Opinion," *Journal of Curriculum Studies* 22 (July-August 1990): 343-360.

<sup>15</sup>Martin Mulder and J. te Brake, "Reliability Testing of Two Analysis Instruments for Decision Making in Curriculum Conferences," *Studies in Educational Evaluation* 16 (No. 3, 1990): 529-550.

Knitter, and others have intensively discussed.<sup>16</sup> As Harris points out, however, "this flurry of curriculum theorizing appears to have generated little 'practical' activity, or at least few published reports of such activity."<sup>17</sup> Harris pleads for the wide communication of the concept and application of deliberation in practical situations. Codifying curriculum deliberations could help us disseminate the concept and train curriculum workers—for instance, through inservice workshops or seminars and in college programs for teacher education. Thus, we might perceive curriculum deliberation as a craft, not an art.<sup>18</sup>

Perhaps the recent curriculum literature indicates a change; many papers and publications have recently appeared on case studies and other in-depth studies of deliberations in curriculum groups.<sup>19</sup> Although I tend to agree with

<sup>16</sup>Decker F. Walker, "A Study of Deliberation in Three Curriculum Projects," *Curriculum Theory Network* 7 (1971): 118-134, Decker F. Walker, "A Naturalistic Model for Curriculum Development," *School Review* 80 (November 1971): 51-65, F. Michael Connelly, "The Functions of Curriculum Development," *Interchange* 3 (No. 2/3, 1972): 161-177, Seymour A. Fox, "A Practical Image of the 'Practical,'" *Curriculum Theory Network* (Fall 1972): 45-47, Ian Westbury, "The Character of a Curriculum for a 'Practical' Curriculum," *Curriculum Theory Network* (Fall 1972): 25-36, Ian Westbury, "The Aristotelian 'Art' of Rhetoric and the 'Art' of Curriculum," *Philosophy of Education. Proceedings of the 23rd Annual Meeting of the Philosophy of Education Society* (Edwardsville, Southern Illinois University at Edwardsville, 1972), Warner Wick, "Knowledge and Action. The Theory and Practice of the 'Practical,'" *Curriculum Theory Network* (Fall 1972): 37-44, William A. Reid, *Thinking about the Curriculum. The Nature and Treatment of Curriculum Problems* (London: Routledge and Kegan Paul, 1978), William A. Reid, "Practical Reasoning and Curriculum Theory. In Search of a New Paradigm," *Curriculum Inquiry* 9 (Fall 1979): 187-207, William A. Reid, "The Deliberative Approach to the Study of the Curriculum and Its Relation to Critical Pluralism," in *Rethinking Curriculum Studies*, ed. Martin Lawn and Len Barton (New York: Halsted Press, 1981), pp. 160-190, Elliot Eisner, "No Easy Answers. Joseph Schwab's Contributions to Curriculum," *Curriculum Inquiry* 14 (Summer 1984): 201-210, Peter Pereira, "Deliberation and the Arts of Perception," *Journal of Curriculum Studies* 16 (October-December 1984): 347-366, Thomas W. Roby, "Habitus in Impeding Deliberation," *Journal of Curriculum Studies* 17 (January-March 1985): 17-35, William Knitter, "Curriculum Deliberation. Pluralism and the Practical," *Journal of Curriculum Studies* 17 (October-December 1985): 383-397.

<sup>17</sup>Irene B. Harris, "Communicating the Character of 'Deliberation,'" *Journal of Curriculum Studies* 18 (April-June 1986): 115-132.

<sup>18</sup>*Ibid.*, p. 130.

<sup>19</sup>Graham W. F. Orpwood, "The Reflective Deliberator: A Case Study of Curriculum Policymaking," *Journal of Curriculum Studies* 17 (July-September 1985): 293-304, Elaine Atkins, "From Competing Paradigms to Final Consensus. A Case Study of the Deliberations of a Problematic Curriculum Group" (paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1988), Lynne M. Hannay, R. Asselin, and Wayne O. Seller, "Curriculum Development as Practiced in the Real World of Public Schools. A Multi-site Study" (paper presented at the annual meeting of the American Educational Research Association, Washington, DC, April 1987), Stewart A. Bonser, and Shirley J. Grundy, "Reflective Deliberation in the Formulation of a School Curriculum Policy," *Journal of Curriculum Studies* 20 (January-February 1988): 35-45, Lynne M. Hannay and Wayne O. Seller, "The Curriculum Leadership Role in Facilitating Curriculum Deliberation" (paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1988), Maurice Holt, "Deliberation, Innovation, and Management Conflict and Compromise in a Comprehensive School" (paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1988), Maurice Holt, "Managing Curriculum Change in a Comprehensive School. Conflict, Compromise, and Deliberation," *Journal of Curriculum Studies* 22 (March-April 1990): 137-148, Kerry J. Kennedy, "Creating a Context for Curriculum Deliberation by Teachers" (paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1988).

Harris, the analysis of deliberation processes has only just begun, and our curriculum research is not so advanced that we can present cases of curriculum deliberation sufficiently well to illustrate appropriate principles for guiding practical curriculum deliberation.

### THEORETICAL FRAMEWORK

The crucial concepts are deliberation, argumentation, and interaction—concepts embedded in Schwab's theories on "the practical" and Frey's "justification" or "curriculum legitimation."<sup>20</sup> Both theories weave curriculum design into the context of the development process. These processes are predominantly situation-specific because contexts tend to vary significantly on certain variables. Moreover, curriculum design is much more than a mere technical, rational process, a normative component also characterizes decision making during curriculum design. Besides a rational analysis of decisions and solutions to problems, intuition and praxeological theories (e.g., beliefs and personal theories) also play a part in the process. These basic theoretical notions have implications for organizing the curriculum design process.

According to Schwab, we need to consider five commonplaces when designing curriculums: teacher, student, milieu, subject matter, and curriculum. We have to integrate our expertise in and experience with these commonplaces during the curriculum design process. Practical curriculum development should preferably take place through the method of deliberation. Schwab describes several characteristics of deliberation: It concerns concrete cases, multiple principles always apply, cogent characteristics are always present in cases, it is complex and arduous, and ends and means mutually determine each other. Further, he recommends how to implement deliberative curriculum design: identify the broadest range of problems and what each one requires, generate and consider the widest possible variety of solutions, rehearse each alternative using a representative variety of people who must live with the consequences of a chosen action, and choose the best alternative. Finally, Schwab identifies the categories of participants of curriculum deliberation: teachers, principals, school board and community members, students, subject-matter specialists, and psychologists.

The method of deliberation is a strongly context-bound and practical mode of curricular problem solving. Frey has elaborated the concept of

---

<sup>20</sup>Joseph J. Schwab, "The Practical. A Language for Curriculum," in *Science, Curriculum, and Liberal Education. Selected Essays*, ed. Ian Westbury and N. J. Wilkhof (Chicago: University of Chicago Press, 1978), pp. 287–321; Karl Frey, "Rechtfertigung von Bildungsinhalten im elementaren Diskurs: Ein Entwurf für den Bereich der didaktischen Rekonstruktion," in *Curriculumentwicklung—Begründung und Legitimation*, ed. R. Künzli (München: Kösel, 1975), pp. 103–129; Karl Frey, *Curriculum-Konferenz. Gebiet Mikroprozessor* (Kiel: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, 1981); K. Frey, *Curriculum Conference. An Approach for Curriculum Development in Groups* (Kiel: Institute for Science Education, 1982).

didactical reconstruction and developed a more detailed set of guidelines for organizing the curriculum design process. This curriculum conference method gives the concept concrete form.

The three most important concepts in this study are deliberation, argumentation, and interaction. Argumentation and interaction are important characteristics of deliberation.

- *Deliberation* is "a rather complex, fluid, transactional discipline aimed at identification of the desirable and at either attainment of the desired or at alteration of the desires."<sup>21</sup> Or according to Reid, curriculum deliberation is curriculum discussion and deliberative problem solving.<sup>22</sup> Schwab proposes several deliberation rules.

- *Argumentation* takes place when an opinion can be interpreted in a linguistic way, and one or more overtly expressed reasons are linguistically clear. Frey's curriculum conference model uses the argumentation rules of Füglistner as guidelines for justifying curriculum decisions.<sup>23</sup>

- *Interaction* is the reciprocal communication between participants during a curriculum conference. Frey's curriculum conference model uses the theory of thematic interaction and Cohn's interaction rules as a condition for improving the quality of the decision-making process.<sup>24</sup>

Curriculum deliberation here deals with the decision-making process in the curriculum design stage, although deliberation may also be the desirable mode of decision making in other curricular contexts. Nolan and Short also apply the concept of deliberation to curriculum implementation and evaluation.<sup>25</sup>

Mulder and te Brake have developed and pilot-tested prototypes of two instruments.<sup>26</sup> These instruments have categories for analyzing argumentation in typed protocols of curriculum conferences and a specific computer program for analyzing the interaction of the decision-making process during the conferences. Their study accurately describes the development of the instrument prototypes and the reliability testing with trained analysts and observers. This article presents the results achieved with the reliable components of these instruments.

<sup>21</sup>Joseph J. Schwab, "The Practical. A Language for Curriculum," in *Science, Curriculum, and Liberal Education. Selected Essays*, ed. Ian Westbury and N. J. Wilkhof (Chicago: University of Chicago Press, 1978), p. 291.

<sup>22</sup>William A. Reid, "The Institutional Context of Curriculum Deliberation: A Research Agenda," *Journal of Curriculum and Supervision* 4 (Fall 1988): 6.

<sup>23</sup>P. Füglistner, *Lehrzielberatung. Zur Reflexion Didaktischen Handelns mit Schülern* (München: Kösel, 1978).

<sup>24</sup>R. Cohn, *Von der Psychoanalyse zur Themenzentrierten Interaktion* (Stuttgart: Klett, 1985).

<sup>25</sup>James F. Nolan and Edmund C. Short, "Combining Qualitative and Quantitative Modes of Inquiry to Inform Curriculum Deliberation. A Proposed Framework," *Curriculum Perspectives* 5 (May 1985): 15-23.

<sup>26</sup>Martin Mulder and J. te Brake, "Reliability Testing of Two Analysis Instruments for Decision Making in Curriculum Conferences," *Studies in Educational Evaluation* 16 (No. 3, 1990): 529-550.

## DESIGN OF THE INQUIRY

The analysis of the argumentation and interaction processes during curriculum conferences consists of four stages:

1. *Tracing episodes.* I traced 18 episodes and used some of them for analysis.

2. *Tracing stages within episodes.* I distinguished opening, continuation, and terminating propositions.

3. *Tracing the nature of the deliberation process within episodes.* I distinguished four categories: the sender, the reaction, the message, and the receiver.

4. *Tracing the nature of argumentation and interaction.* In doing this analysis of the curriculum conferences, the study focused primarily on three variables: the types of *message* within the deliberation process, the types of *argument* made, and the nature of the *interaction*. Categories within each variable are defined as follows:

- *types of message*—argumentation, conclusions, (other) terminating propositions, and miscellaneous

- *types of argument*—political, societal, scientific, job-oriented, educational, anthropological, curriculum-technological, research-technical, and miscellaneous

- *interaction*—speaking time of senders and the nature of the messages

The research material consisted of typed protocols, six texts were selected for argumentation analysis. Further, video registrations of three sections were taken for observation analysis.

For the content analysis, two independent pairs of trained undergraduate students used MARS to analyze the protocols. For the observation analysis, three independently trained undergraduate students used ÇODOM to analyze the video registrations. For reliability testing, the scores of the rater-pairs and the individual observers were combined into cross-tabs. These cross-tabs were used for further analysis of the data.

## RESULTS

### *Deliberation*

The analysis of the nature of the decision-making process took place after the first two stages were completed—the first aimed at tracing and marking the episodes during the curriculum conference and the second aimed at tracing and marking stages within episodes (opening, continuation, and terminating propositions). The third stage distinguished the sender, the reaction, the message, and the receiver—the message was most important. It encompasses argumentation, conclusions, terminating propositions (that are not conclusions), and miscellaneous messages (see Table 1).

**Table 1. Frequency Distribution of Text Fragments on the Types of Message as Classified by the Rater-Pairs**

|               | Pair 2 |      |      |       |    | Total |
|---------------|--------|------|------|-------|----|-------|
|               | Arg    | Misc | Term | Concl | NC |       |
| <i>Pair 1</i> |        |      |      |       |    |       |
| Arg           | 27     | 5    | —    | —     | 4  | 36    |
| Misc          | 3      | 1    | —    | 1     | 3  | 8     |
| Term          | —      | —    | 1    | —     | —  | 1     |
| Concl         | —      | —    | 1    | 2     | —  | 3     |
| NC            | 4      | 3    | —    | —     | —  | 7     |
| Total         | 34     | 9    | 2    | 3     | 7  | 55    |

*Note.* Arg = argumentation; Misc = miscellaneous; Term = terminating proposition; Concl = conclusion; NC = not classified.

Two rater-pairs analyzed 55 messages within episodes. The rater-pairs were sure of 87 percent (27/31) of the arguments in all messages. They were sure of 56 percent (31/55) of all messages within the episodes.

#### *Argumentation*

The fourth analysis stage traced the nature of argumentation and interaction using two analysis instruments (MARS and CODOM). This section describes the reliable data collected with MARS. The following section presents the reliable data collected with CODOM.

The rater-pairs analyzed six text fragments from the typed protocols. They selected these fragments by length and complexity. The interrater reliability on the type of argument during the message varied considerably by text (see Table 2). To test the reliability of the data, the rater-pairs used Cohen's kappa, with a reliability criterion of .50. This criterion appears reasonable.<sup>27</sup> The average kappa value of the six texts together was .549. But the data in Table 2 show that the kappa values varied from .000 to .750.

Table 3 gives the distribution of the rater-pairs' scores for the same variable, the type of argument:

- *political*—propositions related to educational policymaking; some descriptors: organization, structure, goals, finance, and flows in education
- *societal*—propositions related to social-economic aspects of society, among other things about the labor market in general; some descriptors: small and medium-sized companies (where most graduates of the educational subsystem studies find a place) and socioeconomic trends

<sup>27</sup>Ibid.

**Table 2. Interrater Reliability (in kappas) on the Types of Arguments during the Message**

|                | Kappa | No. of arguments |
|----------------|-------|------------------|
| Text 2         | .636* | 4                |
| Text 7         | .750* | 6                |
| Text 5         | .000  | 3                |
| Text 9         | .710* | 12               |
| Text 6         | .170  | 7                |
| Text 11        | .000  | 3                |
| Texts together | .549* | 35               |

Note. Reliability criterion = .50.

\* = significant.

- *scientific*—propositions borrowed from the content and structure of scientific disciplines, some descriptors: production processes and technology
- *job-oriented*—propositions related to job profiles or job practices; some descriptors: tasks and position level
- *educational*—propositions related to the organization of education; some descriptors: educational content, also about media, subjects, and practicals
- *anthropological*—propositions related to a certain philosophy of the meaning of reality, humankind, and the world, some descriptors: educational conception and societal or political philosophy

**Table 3. Frequency Distribution of Both Rater-Pairs' Scores for the Types of Argument (1-9) during the Message**

| Argument                        | Argument |   |   |   |   |   |   |   |   | Total |
|---------------------------------|----------|---|---|---|---|---|---|---|---|-------|
|                                 | 1        | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |       |
| 1. Political                    | 2        | — | — | — | 1 | — | — | — | — | 3     |
| 2. Societal                     | 2        | — | — | — | — | — | — | — | — | 2     |
| 3. Scientific                   | —        | 1 | 8 | — | — | — | — | — | — | 9     |
| 4. Job-oriented                 | —        | — | — | 2 | — | — | — | — | — | 2     |
| 5. Educational                  | —        | 1 | — | 1 | 3 | — | — | — | 1 | 6     |
| 6. Anthropological              | —        | — | — | — | — | — | — | — | — | —     |
| 7. Curriculum-<br>technological | —        | — | — | — | — | — | — | — | — | —     |
| 8. Research-<br>technical       | —        | — | — | — | — | — | — | — | — | —     |
| 9. Miscellaneous                | 4        | 2 | — | 1 | 1 | — | — | — | 5 | 13    |
| Total                           | 8        | 4 | 8 | 4 | 5 | — | — | — | 6 | 35    |

Note. Interrater reliability, kappa = .549.



- *curriculum-technological*—propositions about the curricular problem-solving process itself (designing, developing, implementing, and evaluating); some descriptors: design methodology and developmental conditions
- *research-technical*—propositions about the preparatory study carried out and the propositions related to the results of the information material; some descriptors: information that conclusions are based on, how conclusions are drawn, how the data are compiled
- *miscellaneous*—all propositions that do not fit into the other categories

If we consider the rater-pairs' scores that do not disagree—the data on the diagonal in Table 3—20 arguments appeared unambiguous. Of these 20 arguments, 8 appeared scientific, 3 educational, 2 political, 2 job-oriented, and 5 miscellaneous. Unambiguous social, anthropological, curriculum-technological, and research-technical arguments did not occur in these texts.

### Interaction

This section presents the results of the conference participants' interaction. We considered two variables: the speaking time of the participants, or senders, which indicates their verbal dominance, and the nature of the messages the participants used during deliberation. Table 4 gives the interobserver reliability data for both variables. The kappa values for the senders varied from .61 to .91. All these values are significant at the criterion value of .50. For the nature of the messages, the kappa values varied from .10 to .79. Section C showed the least interobserver reliability.

**Table 4. Interobserver Reliability on the Senders and the Nature of the Messages (in kappas) by Conference Sections and by Observer-Pairs**

|                               | Observer-pair |      |      |
|-------------------------------|---------------|------|------|
|                               | 1-2           | 1-3  | 2-3  |
| <i>Senders</i>                |               |      |      |
| Section A                     | .67*          | .77* | .81* |
| Section B                     | .81*          | .82* | .91* |
| Section C                     | .61*          | .73* | .62* |
| <i>Nature of the messages</i> |               |      |      |
| Section A                     | .52*          | .52* | .44  |
| Section B                     | .63*          | .68* | .79* |
| Section C                     | .10           | .22  | .36  |

Note. Reliability criterion = .50.

\* = significant.

*Speaking time of senders.* To determine the speaking time of the conference participants, the three observers coded the speaker every 10 seconds. This interval is relatively large, but the observers had to code other variables simultaneously, which made the 10-second interval the best alternative. The three sections of the conference videotapes consisted of 25, 23, and 21 minutes of deliberation. The observers independently registered their scores on small hand computers specially programmed for this purpose.

Then I combined the data into cross-tabs for the different pair combinations: Observer 1–Observer 2, Observer 1–Observer 3, and Observer 2–Observer 3.

Thus, each section had three cross-tabs: A, B, and C. The row and column totals of these cross-tabs form Table 5. Because I crossed each observer by two others, two rows (for Observer 1) or column totals (for Observer 3), or one row and one column total (for Observer 2) were available. I entered these row and column totals into Table 5. Minor differences between the row and column totals created the differences between the within-observer scores. To eliminate these differences, I summed all the observers' scores and computed the senders' average speaking time. Then, I computed each sender's percentage of speaking time to compare the speaking time of participants during the three conference sections.

**Table 5. Average Mean Frequencies (in percentages) of the Rate Senders Speak during the Conference Sections and the Corresponding Mean Time (in minutes) of Senders' Verbal Interactions**

|           | Section |    |    | $M_{ABC}$ | Time |
|-----------|---------|----|----|-----------|------|
|           | A       | B  | C  |           |      |
| Silence   | 1       | —  | 2  | 1         | 0.7  |
| Confusion | 6       | 5  | 6  | 6         | 4.1  |
| Sender 1  | 25      | 20 | 22 | 22        | 15.2 |
| Sender 2  | 3       | —  | —  | 1         | 0.7  |
| Sender 3  | 1       | 1  | —  | 1         | 0.7  |
| Sender 4  | 10      | 12 | —  | 11        | 7.6  |
| Sender 5  | 5       | —  | —  | 2         | 1.4  |
| Sender 6  | 1       | 2  | 6  | 3         | 2.1  |
| Sender 7  | 5       | 6  | —  | 4         | 2.8  |
| Sender 8  | —       | —  | —  | —         | —    |
| Sender 9  | 2       | —  | —  | 1         | 0.7  |
| Sender 10 | 2       | 5  | —  | 2         | 1.4  |
| Sender 11 | 18      | 10 | 9  | 12        | 8.3  |
| Sender 12 | 6       | 4  | —  | 3         | 2.1  |
| Sender 13 | 1       | —  | 6  | 2         | 1.4  |
| Sender 14 | 3       | 10 | 7  | 7         | 4.8  |
| Sender 15 | —       | —  | 2  | 1         | 0.7  |
| Sender 16 | 9       | 8  | 16 | 11        | 7.6  |
| Sender 17 | 2       | 1  | 2  | 2         | 2.1  |
| Sender 18 | —       | 17 | 16 | 11        | 7.6  |
| Sender 19 | 4       | —  | 6  | 3         | 2.1  |

Few periods of silence fell during the sections of the curriculum conference (varying from 0 to 2 percent). Confusion existed more frequently, varying between the sections from 5 to 6 percent. To further compare the results, I entered the means into Table 5. This table also includes the average means.

Sender 1 was the chair (an assistant university professor in curriculum technology). He spoke most frequently during the sections, from 20 to 25 percent of the time available. Sender 2 was the initiator (an assistant university professor in instructional design), Sender 3 was a study supervisor (a university professor in curriculum technology), and Sender 4 was a subject-matter expert (a university professor in the content area). Senders 2 and 3 used relatively little speaking time, both 1 percent. Sender 4, however, used 11 percent. The actual participants showed a variation in average percentages of speaking time of 0 percent (Sender 8) to 12 percent (Sender 11). The mean average of the group's percentage of speaking time was 4.3 percent, and the standard deviation was 4.0 percent.

*Nature of the messages.* To determine the nature of the senders' messages, the three observers coded each message, again every 10 seconds. They used the same research material and scoring procedure as for the analysis of the senders' speaking time.

The observers distinguished nine message categories:

- *starting discussion*—questions and remarks intended to keep the discussion going
- *giving an introduction*—at the moment of coding, the context of the sender's message not yet clear
- *asking information*—requests for information only
- *giving information*—senders' reaction to a request for information
- *asking opinions*—explicit requests for opinions, judgments, views, and arguments
- *giving opinions*—all messages that follow after starting discussion, giving information, asking opinions, checking consensus, and drawing a conclusion, as long as these messages do not take the form of questions
- *checking consensus*—questions that test whether everybody agrees with a certain statement
- *drawing a conclusion*—a conclusion formulated or a summary given
- *miscellaneous*—messages difficult to classify

Here, I also combined the data with cross-tabs for the pair combinations: Observer 1–Observer 2, Observer 1–Observer 3, and Observer 2–Observer 3. Thus, each section has three cross-tabs: A, B, and C.

I summed the observers' scores on the senders and computed the mean frequencies of the message categories. Then I computed the percentages of the scores to compare the nature of the messages during the three sections

(see Table 6). Apparently, the participants most frequently gave opinions. The observers scored this category in about 82 percent of the cases. The other categories showed far lower values, varying from .08 percent (checking consensus) to 6.39 percent (giving information). The rest of the categories had the following rank order: asking information (4.64 percent) drawing a conclusion (2.93 percent), asking opinions (1.81 percent), giving an introduction (1.37 percent), miscellaneous (.64 percent), and starting discussion (.58 percent).

Some sections showed considerable variation in the categories. Apart from the giving-opinion category, where the greatest difference between the largest and the smallest mean percentage was 13.6 percent, the drawing-conclusions, asking-information, and giving-information categories showed large section differences between the maximum and minimum mean percentages: respectively, 8.78 percent, 7.86 percent, and 6.62 percent.

**Table 6. Average Mean Frequencies (in percentages) of the Nature of the Messages during the Conference Sections**

|                        | Section |       |       | $M_{ABC}$ |
|------------------------|---------|-------|-------|-----------|
|                        | A       | B     | C     |           |
| Starting discussion    | 0.53    | 1.04  | 0.18  | 0.58      |
| Giving an introduction | 0.76    | 2.15  | 1.20  | 1.37      |
| Asking information     | 9.70    | 2.39  | 1.84  | 4.64      |
| Giving information     | 10.45   | 3.83  | 4.88  | 6.39      |
| Asking opinions        | 2.27    | 1.04  | 2.12  | 1.81      |
| Giving opinions        | 75.91   | 79.17 | 89.51 | 81.53     |
| Checking consensus     | —       | 0.24  | —     | 0.08      |
| Drawing a conclusion   | —       | 8.78  | —     | 2.93      |
| Miscellaneous          | 0.38    | 1.36  | 0.18  | 0.64      |
| Total                  | 100     | 100   | 100   | 100       |

## DISCUSSION

Argumentation occurred frequently during this curriculum conference. However, the participants reached only two conclusions during the episodes analyzed. We can interpret this astonishing finding in at least two ways: (1) the issues discussed proved to be so difficult that the participants had to devote considerable time to exchanging opinions and arguments, (2) the participants actually drew more conclusions, but they remained implicit. This latter case

would make the argumentation and observation analysis instruments partially invalid, but more analysis of the deliberation process is needed for clarification.

At this stage, I believe two important conclusions based on a thorough, profound discussion may be more valuable than many less-important conclusions. The research material does not yet allow for analyses to weight the importance of the decisions.

The kappa variations found in the argumentation analysis of .000 to .750 implies a partial error in the data set. I could have eliminated the data on Texts 5, 6, and 11, leaving 22 arguments for further analysis. I could have followed the same approach for the observation analysis: If I had taken only the corresponding scores of all observers, a data set with 928 instead of 1,100 scores (84 percent) would have remained.

In this inquiry, I chose another approach. I took a cross-tab of the rater-pairs on the type-of-argument variable. Then, I determined the fraction of the corresponding scores of the total number of scores. This method, however, resulted in the same conclusions about the relative frequency of the types of arguments. A check of the alternative approach for analyzing the nature of the messages showed essentially the same results, but the tendency in that data justified even stronger conclusions.

Scientific arguments dominated the deliberation process. Educational, political, and job-oriented arguments occurred less frequently. Unambiguous societal, anthropological, curriculum-technological, and research-technical arguments were absent. This result poses some questions. What happens when we compare Schwab's commonplaces with this distribution of argument types? We would expect teacher-, student-, milieu-, subject-matter-, and curriculum-oriented arguments, but one category dominates: subject matter. When the argumentation process becomes skewed to only one or a few types of arguments, the chair needs to stimulate other types of arguments. Training curriculum conference chairs to meet this complex task seems important. On the basis of this study, which implies a more valid operation of the theoretical construct, I recommend a tighter relationship between Schwab's commonplaces and the research instruments.

The intention of CODOM was to determine *coalition* and *dominance*, but the instrument works more in the direction of verbal interaction. It determined the relative frequency of instances when participants spoke during the conference and the nature of their messages. A further combination of data should show the relative dominance and possible coalitions in the group of participants. First, combine the scores on the fraction of the verbal interaction participants use (as an indicator of *verbal dominance*) with the participants' opinions on the curriculum design specifications before, during, and after the conference. Then, reconstruct and analyze the reaction patterns to determine whether certain participants mostly approved or disapproved each other's statements. As an analogy with sociograms in sociological research, we could construct "argograms" for the group of participants.

The chair used a large amount of time for speaking himself. His mean speaking time over the three sections studied was 22 percent. Is this amount of time desirable? That question is difficult to answer, but his speaking time seems high here. The participants themselves need sufficient time to speak; here, they had 71 percent of the time available. If we deduct the speaking time of the initiator, the supervisor, and the subject-matter expert, who together used 13 percent of the speaking time, the group participants had left only 58 percent of the time.

The percentages of the senders' verbal interactions during Sections A, B, and C showed high correlations, varying from .6382 to .7381, which implies that we could probably carry out further verbal interaction analyses on comparable research material using only a limited proportion of all material to reduce the analysis time significantly.

The variation of the nature of the messages was high. The mean frequency of giving opinions was about 82 percent of all messages. The argumentation analysis showed that 87 percent of the messages consisted of argumentation. The other message categories showed a low frequency. The categories varied from .08 percent on checking consensus to 6.39 percent on giving information. A comparison of the differences between all categories except giving opinions shows that giving and asking information were two important characteristics of the deliberation process. Drawing conclusions was the next important category. Together, the three covered about 95 percent of all verbal interaction time during the analyzed sections of the curriculum conference. Asking opinions, starting discussion, and giving introductions took little time, and apparently the deliberation process does not need much of these kinds of interventions. This finding also indicates the emphasis on sharing arguments, information, and opinions to eventually approve conclusions—the essential components of deliberative curriculum decision-making processes.

Correlations between the percentages of the message categories by Sections A, B, and C varied from .9820 to .9938. Again, these findings imply that we could carry out further observation analyses on comparable research material using only a limited proportion of all material to reduce the analysis time significantly. Of all categories except giving opinions, however, drawing a conclusion showed the highest variation between the sections. Only Section B contained 8.78 percent for this category.

A comparison of the results of both the argumentation and observation analyses reveals an important discrepancy. How could many messages be classified as both argumentation and opinion? A message is either an argument or an opinion. The definitions used in the instruments, however, explain this phenomenon. Analysts had to mark sentences in the typed protocols that contained opinions and arguments. These sentences were labeled *argumentative*. Observers, however, had to classify messages as *giving opinions* whenever the senders reacted to a starting proposition, gave information, asked opinions, checked consensus, and drew a conclusion, as long as the message was not in the form of a question. A close look at the definitions

serving as instructions for the analysts and observers reveals some contamination of both the concepts of argumentation and opinion. Further studies need to separate these variables more carefully, and the research instruments will need some revision. Despite these weaknesses, a substantial proportion of the deliberation consisted of giving opinions and producing the arguments to back up these opinions.<sup>28</sup>

---

MARTIN MULDER is Lecturer, Department of Education, Division of Curriculum Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, the Netherlands.

Ben-Peretz, Miriam. *The Teacher-Curriculum Encounter: Freeing Teachers from the Tyranny of Texts*. Albany: State University of New York Press, 1990. 152 pp. \$39.50/\$12.95.

Teachers' development and use of curriculum and curriculum materials is the focus of this book. It discusses an array of new concepts and strategies (curriculum potential, interpretation, analysis, deliberation, and active teacher roles in development) and their implications for teacher education, staff development, and coping with teachers' concerns about curriculum.

Granheim, Marit, Maurice Kogan, and Ulf Lundgren. *Evaluation as Policy-making: Introducing Evaluation into a National Decentralized Educational System*. London: Jessica Kingsley, 1990. 256 pp. \$66.00.

Evaluation as an instrument for improving educational policy and practice is the subject of this analysis by evaluation scholars from Norway, the United States, England, and Sweden. The authors explore issues concerning evaluation and decentralization in the Norwegian context, with references to critical perspectives and schemes operating in Sweden and the United States.

---

<sup>28</sup>I want to express my thanks to J. te Brake, who generated the data for this study.

Copyright © 1991 by the Association for Supervision and Curriculum Development. All rights reserved.