MS-831: Jack, Joseph, and Morton Mandel Foundation Records, 1980 – 2008.
Series E: Mandel Foundation Israel, 1984 – 1999.

Box	Folder
D-1	1994

Gamoran, Adam. Correspondence, articles, and reports,
1994-1999.

Pages from this file are restricted and are not available online. Please
contact the American Jewish Archives for more information.

From: Adam Gamoran <gamoran@ssc.wisc.edu>
To: annette <annette@vms.huji.ac.il>
Date: 24 June 1999 21:22
Subject: update on indicators project

Dear Annette,

I'm writing to update you on the progress we've made with the Indicators project. The main focus since our meeting in February has been on Ellen and Bethamie's background papers on institutional quality and Jewish identity, respectively. I had a chance to review drafts of the papers in May. Subsequently the papers were revised and we distributed them (still as drafts) to members of our Professors Group for consultation.

Copies of the papers were also sent to you in late May. If you have any comments on the papers, we could incorporate them into another revision if we receive them by July 15. In any case, we look forward to discussing the papers and their implications for our work when we meet in August.

CONSULTATIONS ON BACKGROUND PAPERS

We held two meetings with our Professors Group about the papers. The first took place at the seminar in Los Angeles on June 4, and the second was held by conference call on June 15 with David Kaplan and Barbara Schneider, two members of the group who were unable to attend the seminar.

Discussion at the June 4 seminar was wide ranging, but we obtained several helpful suggestions. Overall, participants found both papers informative, useful, and interesting, and the conversations were quite spirited.

On the topic of identity, participants noted the lack of a developmental perspective in work on identity, an issue that may be particularly important for Jewish identity among diaspora Jews whose identities seem to shift and flux as they pass through different life stages. Interestingly,

Bethamie's forthcoming work on "Connections and Journeys" may help address this issue, albeit retrospectively. Another important comment, though outside the realm of Bethamie's paper, is that we need a clearer articulation of the relation between Jewish education and Jewish identity. A fourth point is that more information about empirical analyses of identity in other (non-Jewish) domains would be helpful. Finally, participants noted that more work needs to be done to prioritize among the many recommendations discussed at the end of the paper. Bethamie's proposals are compelling and many are new and creative, but because we cannot do everything we need more guidance on prioritizing.

In discussing the paper on institutions, participants stressed the importance of examining the outcomes as well as the input characteristics of Jewish institutions. A related concern is to link potential outcomes indicators as closely as possible to the activities and content of the institutions. These issues will need careful consideration in the future, but probably do not call for any revision in the paper per se. Participants found Ellen's recommendations for approaching the study of institutions quite reasonable, given the complexities involved. Her attention to both qualitative and quantitative strategies for addressing broader questions about the quality and effectiveness of institutions was particularly appreciated by the group.

The conference call with David Kaplan and Barbara Schneider focused on the methodological implications of the papers. Both readers found the papers "excellent, informative, sound, and of very high quality." Their enthusiasm for the papers' contents led them to offer many suggestions about how data on indicators could be collected, if the instruments for indicators were designed as recommended in the papers. They would like to see an indicators project that:

* is longitudinal for individuals as well as monitoring a system over time
* places individuals in their institutional and community contexts

While this is an exciting agenda, I cautioned that it is too ambitious for the present time. Consequently we discussed ways of beginning more modestly, perhaps by proceeding at two separate levels (individual across

02/09/99

the country, and institutions within selected communities) without an empirical linkage between the two. This would allow us to use the NJPS and its possible supplement for national data on individuals, and to obtain limited data on institutions within communities as suggested by Ellen. This more modest approach would have obvious limitations, in that it would not follow individuals over time, and would not link individuals to their particular institutions. However, it would satisfy the primary purpose of the indicators project (at least as I envision it), which is to provide data on current status and on changes over time for selected key elements of Jewish education.

The next step for these papers is to commission outside reviews from individuals with expertise in the Judaic worlds that are addressed by the papers (i.e., an expert on Jewish educational institutions, and an expert on Jewish identity). We expect to send each paper to one expert who will provide a written review. Ultimately, the papers will help guide our decisions as we plan the future of this project.

OTHER ACTIVITIES

David Kaplan has agreed to carry on with the analysis of secular national data sets which we began last year with our former research assistant. These analyses will serve as the basis for Indicator Reports, before we have our own data available. David is a professor of education at the University of Delaware who specializes in the analysis of large-scale survey data. Among his many publications is a recent article on the statistical validation of educational indicators. I am in close contact with him and we are delighted to have him on board. I expect to have a progress report on this work when we meet in August.

I have had preliminary discussions with Ellen, Bethamie, and Mark Gurvis about a supplement to the NJPS which could serve as an oppportunity to gather data specifically for the indicaotrs project. We are closely involved in the design of the instrument for the main NJPS survey, and it looks like many of our items will be incorporated. If that occurs, then we may propose to use the supplement to explore the institutional and community contexts in which the individuals are embedded (as perceived by the individuals).

02/09/99

Finally, Barbara Schneider has raised the possibility of using instruments from her national study of adolescent development, along with items designed for the indicators project, in a sample of Chicago day schools. I'm not sure if anything will come of this idea but we are discussing it.

I'd welcome any response you may have to these activities, and look forward to further discussions in August.

Best,

Adam

Jewish Educat—
institutions'
report

From: Danny Marom <marom@vms.huji.ac.il>
To: annette@vms.huji.ac.il <annette@vms.huji.ac.il>
Date: 07 September 1999 10:53
Subject: questions on "A Comprehensive Study of Jewish Schools in Chicago"

1. What do the researchers expect to learn from the study that is not already known?
2. Will there be any data on "control" groups which can produce useful comparative findings for policymakers? (eg. data on Jewish versus public education, Jewish versus private non-Jewish education, Jewish formal vs. informal education, Jewish formal education of one kind versus Jewish formal education of another).
3. Will the study relate to early childhood education programs? Could one argue that where a Jewish learner comes out of these programs determines more about their later Jewish identity than what happens at the day and afternoon school stage of development. Similarly, will the study relate to the ways in which day and afternoon schools do or do not prepare the Jewish learner for the challenges to Jewish identity which emerge at the university age and study context (Note: Lipsett claimed that this was the most critical stage in the development of the identity of American Jews).
4. What is the researchers' assumption about the relationship of Jewish literacy to identity? Would this assumption rule out the possibility that a learner with a low Jewish literacy level might have a strong Jewish identity or that a learner with a high Jewish literacy level might have a low Jewish identity?
5. How do the researchers intend to isolate the exclusive role of Jewish schooling on the development of Jewish identity? How will they know that high or low Jewish identity will not be the outcome of non-school factors?
6. What is the justification for the geographical definition of the Chicago area for the study? How does this justication relate to the aims and audience of the research?
7. Will the research look at how aims get decided, articulated and transmitted (eg. in the induction of new staff)?
8. What are the authors' assumptions concerning the role of Hebrew in Jewish identity and what is the justification for its centrality in the determination of Jewish literacy? Is this in line with the current commitments of American Jewry?
9. Have the researchers considered applying the categories suggested by Jim Coleman and others inspired by him in understanding the relative effectiveness and advantage of private over public education in America (eg. "intergenerational closure")?

Please let me know if you need any more. I'll be in the office as of 2:00 p.m.

Great meeting you yesterday. DM

(1) Update from Annette

(2)

(3) Chicago

07/09/99

From: Adam Gamoran <gamoran@ssc.wisc.edu>
To: Annette <annette@vms.huji.ac.il>
Date: 14:40 1999 אוגוסט 05 'יום חמישי
Subject: 1996 memo on lessons from lead communities

Annette,

It was great seeing you in Jerusalem, thanks for finding time for our meeting. We had a wonderful trip and were very sorry to leave.

I'll get you the brief memo on Barbara Schneider's idea for a survey in Chicago Jewish schools as soon as possible. In the meantime, attached is the memo we wrote in 1996 for Karen Barth, about lessons learned from working with and studying lead communities. Please pass it on to Shmuel if you think that's appropriate.

I received the message below after I returned, and can respond that I'd be happy to join you for a meeting in Madison on or after August 26. I have a heavy travel schedule in September so we should get the date set as soon as possible.

Adam

At 03:54 PM 7/18/99 +0300, you wrote:
>Dear Adam,
>
>Both the 26th and 28th are open at 8am. Call or write as convenient when you know which you prefer.
>
>Seymour and I think that a visit at Wisconsin might be useful at the end of the Summer. We may ask some people to join us in order to discuss R&D ideas. Might you be available for a meeting on or after August 26th?
>
>Thanks,
>
>Annette
>
>
>
><!DOCTYPE HTML PUBLIC "-//W3C//DTD W3 HTML//EN">
><HTML>
><HEAD>
>
><META content="text/html; charset=iso-8859-8-i" http-equiv=Content-Type>
><META content="MSHTML 4.72.2201.0" name=GENERATOR>
></HEAD>

From: Adam Gamoran <gamoran@ssc.wisc.edu>
To: Annette <annette@vms.huji.ac.il>
Date: 05 August 1999 14:40
Subject: 1996 memo on lessons from lead communities

Annette,

It was great seeing you in Jerusalem, thanks for finding time for our meeting. We had a wonderful trip and were very sorry to leave.

I'll get you the brief memo on Barbara Schneider's idea for a survey in Chicago Jewish schools as soon as possible. In the meantime, attached is the memo we wrote in 1996 for Karen Barth, about lessons learned from working with and studying lead communities. Please pass it on to Shmuel if you think that's appropriate.

I received the message below after I returned, and can respond that I'd be happy to join you for a meeting in Madison on or after August 26. I have a heavy travel schedule in September so we should get the date set as soon as possible.

Adam

At 03:54 PM 7/18/99 +0300, you wrote:
>Dear Adam,
>
>Both the 26th and 28th are open at 8am. Call or write as convenient when you know which you prefer.
>
>Seymour and I think that a visit at Wisconsin might be useful at the end of the Summer. We may ask some people to join us in order to discuss R&D ideas. Might you be available for a meeting on or after August 26th?
>
>Thanks,
>
>Annette
>
>
>
><!DOCTYPE HTML PUBLIC "-//W3C//DTD W3 HTML//EN">
><HTML>
><HEAD>
>
><META content="text/html; charset=iso-8859-8-i" http-equiv=Content-Type>
><META content="MSHTML 4.72.2201.0" name=GENERATOR>
></HEAD>
><BODY bgColor=#ffffff>

09/08/99

><DIV>
><P align=left dir=ltr>Dear Adam,</P>
><P align=left dir=ltr>Both the 26th and 28th are open at
>8am. Call or write as convenient when you know which you prefer.</P>
><P align=left dir=ltr>Seymour and I think that a visit at Wisconsin might be
>useful at the end of the Summer. We may ask some people to join us in order to
>discuss R&D ideas. Might you be available for a meeting on or after August
>26th?</P>
><P align=left dir=ltr>Thanks,</P>
><P align=left dir=ltr>Annette</P>
><P align=left dir=ltr> </P></DIV></BODY></HTML>
>

09/08/99

A Comprehensive Study of Jewish Schools in Chicago

Prepared by Barbara Schneider and Adam Gamoran
August 1999

Purpose:

We are proposing to conduct an intensive study of Jewish day and after-school programs in Chicago. This project will specifically examine: 1) what are the experiences of children in Jewish schools, focusing on the curricular organization including the content being taught and how it is taught; 2) what are the qualifications, activities, and experiences of teachers and administrators in Jewish schools; 3) how are Jewish day schools organized and financed, and how viable will these schools be in the next century; and 4) how does a sense of Jewish identity develop in children and what experiences foster such development in religious schools.

This project will be conducted by an interdisciplinary team chaired by Barbara Schneider and Adam Gamoran, and including Ellen Goldring, Bethamie Horowitz, David Kaplan, and Linda Waite. Throughout the team there is agreement on the importance of Jewish learning being organized around text and experience. This study will help to determine what text seems to influence young adults' sense of their Jewish life and what experiences help to reinforce these understandings.

Sample:

Learning about the work on adolescents being conducted at the University of Chicago, Barbara Schneider and Linda Waite were contacted to see if they would consider surveying all of the day and after-school Jewish schools and students in Chicago. Realizing that this is a major undertaking and recognizing the related work

recently completed by Adam Gamoran and Ellen Goldring, Barbara and Linda contacted them with the home that they would be interested in working together on a major study of Jewish education in Chicago. This project offers a unique opportunity to survey the entire Jewish school population in Chicago, making it perhaps the largest survey of the Jewish school population in the U.S.

Methods:

Relying on the rich expertise of the interdisciplinary team, we are proposing that the following data collection efforts be undertaken:

Survey of Schools, Administrators, and Teachers

There would be a survey of the school administrators and teachers. This instrument would be used to collect base-line information on how the schools are organized, where they draw their student populations, how long students stay in school, what teaching materials are provided to the staff, what are the evaluation criteria for administrators and staff, how administrators and teachers carry out their roles, what is the relationship between the school and the synagogue and the community at large, what ties does the school have to other schools, the wider Jewish community, programs in Israel, and so on. Teachers would be asked similar questions designed by Goldring and Gamoran with several new items on Jewish literacy and identity.

Student Survey and Interviews

Students in grades six through high school would be surveyed and asked questions about their experiences in these schools including the types of learning activities they engage in, what their interest is in maintaining a Jewish identity, what are their expectations for family life and how important is it for them to continue their Jewish

identity into adulthood. In addition to the surveys, approximately 100 students (fifty day and fifty after-school students) will be interviewed. These intensive interviews will be constructed around issues of Jewish learning and identity. The interviews with day school and after-school students will provide a more in-depth picture of Jewish family and school experiences and of Jewish learning and identity.

Possible Additional Components

In addition to the surveys, two additional components are under consideration. One is to develop and implement an instrument of Jewish literacy for both teachers and students. This would be a relatively concise instrument that would assess both Hebrew and fundamental ethical and historical questions that are uniquely Jewish.

A second additional component may be an ethnographic study of six schools (three day schools and three after-school programs). The purpose of this field study would be to obtain more fine-grained information on the experiences, constraints, and opportunities that Jewish schools, teachers, and students are encountering.

Scheduled Work Plan

During the coming year (1999-2000) the team will meet approximately three times, to construct instruments and methods for carrying out the work. In January a small pilot will be conducted to test the various instruments. Individuals in the Professors Group and others will be contacted for advice, especially regarding the possible Jewish literacy test and content questions posed to the teachers and students.

3

From: Annette Hochsteim <annette@vmss.hujii.ac.il>
To: damit@vms.huji.ac.ill <damit@vms.huji.ac.il>
Date: 25/07/99 08:04
Subject: Fw: RE:telecon

for my meeting with Adam please

thanks,
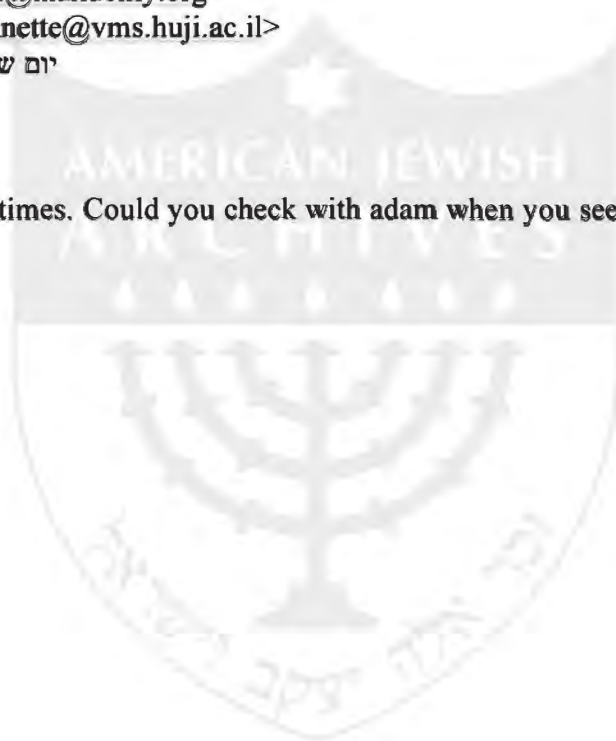
a
-----Original Message-----
From: Gail Dorph <gzdorph@mandelny.org>
To: "Annette Hochstein' <annette@vms.huji.ac.il>
Date: 17:48 1999 יולי 23 יום שישי
Subject: RE:telecon

>I'll check with ellen about times. Could you check with adam when you see
him. gail
>
>

From: Gail Dorph <gzdorph@mandelny.org>
To: 'Annette Hochstein' <annette@vms.huji.ac.il>
Date: 20 July 1999 14:56
Subject: RE: indicators phone call

my first thoughts.

I think we need to "make a plan" as to how we will be moving ahead on indicators.

what will happen to the two papers? will we try to move them ahead to creating indicators.

how will we develop a disseminarion plan for david kaplan's reports? and before that what's our process for reviewing the work?

-----Original Message-----
From: Annette Hochstein [SMTP:annette@vms.huji.ac.il]
Sent: Tuesday, July 20, 1999 7:05 AM
To: Gail Dorph
Subject: Re: indicators phone call

<< File: ATT00001.txt; charset = ISO-8859-7 >>

25/07/99

From: Elana Sztokman <ilanas@mandelschool.org.il>
To: Annette Hochstein <annette@vms.huji.ac.il>
Date: 19 July 1999 12:40
Subject: The Indicator articles

Annette -
I finished Ellen Goldring's article and I'm in the middle of Bethamie's. I
want to comment on the first while my ideas are fresh, and I'll write about
the second tomorrow.

First of all, I really enjoyed reading this paper. I think that she did a
thorough job in presenting indicator systems, and her critiques were right
on the mark.
There are two competing problems in determining the "quality" (whatever that
means) of Jewish education: one is the current lack of readily available
data, and the other is the difficulty in collecting data. What this says to
me is that there is probably no consensus out there about what Jewish
education is or what its goals are. Throughout the descriptions of indicator
systems, especially in Jewish education, I kept thinking that every single
item on the list of items to examine represents a value. The difficulty is
not in collecting data but in determining the value. So for instance, saying
we have to look at teacher degrees presupposes that the degree makes a good
teacher, and that we know what a good teacher is, and that the teacher is
what determines what the education is like. So I think that more important
than collecting the actual data in Jewish education is building this "list"
of values. And that requires answering a lot of difficult questions.
This is why I say there is not going to be a consensus in Jewish education.
When an institution, for example, says they want to instill a "love of
Torah", are they aware that this may come at the price of the goal of
knowledge? (So, for example, I can get kids to love learning by making
everything fun and exciting and eliminating tests -- but they may emerge
with less knowledge than the kids who were studying for tests all year.)
Schools look different because thay have different goals. That's okay, as
long as they're aware of this. I think that indicators reflect values. The
indicator itself is less important than simply identifying the HOW of the
value. So for instance, if a school says they want kids to emerge with total
dedication to the Jewish community, the most important question for it to
ask is, well, how do we do that? Answering that question is similar to
putting together the list of indicators ("techical and moral reductionism")
but without the pressure of actual measurement, energies can be spent on
building a school according to the vision -- programming, structuring,
training, materials, etc. It just seems that before getting to indicators,
there are other issues that need to be clarified.
So, what I'm saying is that I think that the first recommendation of the
paper -- to collect high-consensus widespread indicator data -- seems to me
very difficult. And I'm not sure that it's worthwhile. The second
recomendation I didn't fully understand. The third recomendation seems to me
right on the mark. I am a big supporter of qualitative measures. The

25/07/99

generalizability in quantitative studies is overrated. Case studies, profiles -- these can provide what she calls "rich" descriptions, a wealth of understanding that can actually help design Jewish schools.

Beyond all this, there is of course a basic fundamental problem with looking for indicators in all education but especially Jewish education. And that is, how do you actually know if a school is doing a "good" job? What is the job of a Jewish school? To keep kids in the fold? Or to make them exemplary members of society? These are two very different objectives. They affect issues of narrowness vs openness which are very, very complicated. Yet, what Jewish school is going to give up on either objective? So if you take the outcome approach, for example, a true indicator would be looking at graduates 10, 20 years down the line. Are they strong in their JEwishness? Are they good people? And if they are, how much can be attributed to the school? Obviously, these are difficult questions, and in the whole discussion of indicators, nobody suggested longitudinal studies. ANd yet it lies at the heart of educational vision and goals, and expecially of Jewish education. And there is no way to measure it. SO this puts the whole question of indicators into perspective, I think.

Anyway, those are my thoughts for now. I'll write about the second article tomorrow.
All the best, Elana

Elana Maryles Sztokman
The Mandel Foundation
Jerusalem

25/07/99

29/6/99

Dear Annette,

I'm writing to update you on the progress we've made with the Indicators project. The main focus since our meeting in February has been on Ellen and Bethamie's background papers on institutional quality and Jewish identity, respectively. I had a chance to review drafts of the papers in May. Subsequently the papers were revised and we distributed them (still as drafts) to members of our Professors Group for consultation.

Copies of the papers were also sent to you in late May. If you have any comments on the papers, we could incorporate them into another revision if we receive them by July 15. In any case, we look forward to discussing the papers and their implications for our work when we meet in August.

CONSULTATIONS ON BACKGROUND PAPERS

We held two meetings with our Professors Group about the papers. The first took place at the seminar in Los Angeles on June 4, and the second was held by conference call on June 15 with David Kaplan and Barbara Schneider, two members of the group who were unable to attend the seminar.

Discussion at the June 4 seminar was wide ranging, but we obtained several helpful suggestions. Overall, participants found both papers informative, useful, and interesting, and the conversations were quite spirited. On the topic of identity, participants noted the lack of a developmental perspective in work on identity, an issue that may be particularly important for Jewish identity among diaspora Jews whose identities seem to shift and flux as they pass through different life stages. Interestingly, Bethamie's forthcoming work on "Connections and Journeys" may help address this issue, albeit retrospectively. Another important comment, though outside the realm of Bethamie's paper, is that we need a clearer articulation of the relation between Jewish education and Jewish identity. A fourth point is that more information about empirical analyses of identity in other (non-Jewish) domains would be helpful. Finally, participants noted that more work needs to be done to prioritize among the many recommendations discussed at the end of the paper. Bethamie's proposals are compelling and many are new and creative, but because we cannot do everything we need more guidance on prioritizing.

In discussing the paper on institutions, participants stressed the importance of examining the outcomes as well as the input characteristics of Jewish institutions. A related concern is to link potential outcomes indicators as closely as possible to the activities and content of the institutions. These issues will need careful consideration in the future, but probably do not call for any revision in the paper per se. Participants found Ellen's recommendations for approaching the study of institutions quite reasonable, given the complexities involved. Her attention to both qualitative and quantitative strategies for addressing broader questions about the quality and effectiveness of institutions was particularly appreciated by the group.

The conference call with David Kaplan and Barbara Schneider focused on the methodological implications of the papers. Both readers found the papers "excellent, informative, sound, and of very high quality." Their enthusiasm for the papers' contents led them to offer many suggestions

about how data on indicators could be collected, if the instruments for indicators were designed as recommended in the papers. They would like to see an indicators project that:

* is longitudinal for individuals as well as monitoring a system over time

* places individuals in their institutional and community contexts

While this is an exciting agenda, I cautioned that it is too ambitious for the present time. Consequently we discussed ways of beginning more modestly, perhaps by proceeding at two separate levels (individual across the country, and institutions within selected communities) without an empirical linkage between the two. This would allow us to use the NJPS and its possible supplement for national data on individuals, and to obtain limited data on institutions within communities as suggested by Ellen. This more modest approach would have obvious limitations, in that it would not follow individuals over time, and would not link individuals to their particular institutions. However, it would satisfy the primary purpose of the indicators project (at least as I envision it), which is to provide data on current status and on changes over time for selected key elements of Jewish education.

The next step for these papers is to commission outside reviews from individuals with expertise in the Judaic worlds that are addressed by the papers (i.e., an expert on Jewish educational institutions, and an expert on Jewish identity). We expect to send each paper to one expert who will provide a written review. Ultimately, the papers will help guide our decisions as we plan the future of this project.

OTHER ACTIVITIES

David Kaplan has agreed to carry on with the analysis of secular national data sets which we began last year with our former research assistant. These analyses will serve as the basis for Indicator Reports, before we have our own data available. David is a professor of education at the University of Delaware who specializes in the analysis of large-scale survey data. Among his many publications is a recent article on the statistical validation of educational indicators. I am in close contact with him and we are delighted to have him on board. I expect to have a progress report on this work when we meet in August.

I have had preliminary discussions with Ellen, Bethamie, and Mark Gurvis about a supplement to the NJPS which could serve as an opportunity to gather data specifically for the indicators project. We are closely involved in the design of the instrument for the main NJPS survey, and it looks like many of our items will be incorporated. If that occurs, then we may propose to use the supplement to explore the institutional and community contexts in which the individuals are embedded (as perceived by the individuals).

Finally, Barbara Schneider has raised the possibility of using instruments from her national study of adolescent development, along with items designed for the indicators project, in a sample of Chicago day schools. I'm not sure if anything will come of this idea but we are discussing it.

I'd welcome any response you may have to these activities, and look forward to further discussions in August.

DRAFT -- COMMENTS WELCOME

ABILITY GROUPING AND THE
DISTRIBUTION AND EFFECTS OF CLASSROOM INSTRUCTION[1]

Adam Gamoran, University of Wisconsin-Madison
Martin Nystrand, University of Wisconsin-Madison
Mark Berends, RAND Corporation
Paul C. LePore, University of Wisconsin-Madison

November, 1993

ABILITY GROUPING AND THE
DISTRIBUTION AND EFFECTS OF CLASSROOM INSTRUCTION

ABSTRACT

Ability grouping appears to be a logical means of organizing a student body with diverse academic skills. Many observers contend, however, that when students are grouped according to their purported capacities for learning, high-achieving students receive better instruction and, consequently, increase their achievement advantage over students in other groups. This paper examines the kinds of instruction students receive in honors, regular, and remedial eighth- and ninth-grade English classes. It also assesses the links between instruction and achievement. The authors find that rates of student participation and discussion are higher in honors classes, contributing to the learning gaps between groups. Another finding is that rates of open-ended questions are similar across classes, but that honors students benefit more from such discourse because it more often occurs in the context of sustained study of literature.

ABILITY GROUPING AND THE
DISTRIBUTION AND EFFECTS OF CLASSROOM INSTRUCTION

Ability grouping is the practice of dividing students for instruction according to their purported capacities for learning. To many educators, ability grouping seems a sensible response to academic diversity among students, in that it allows teachers to tailor their instructional approaches to students' abilities (Wilson and Schmits 1978; National Education Association 1990). Critics of ability grouping, however, contend that the practice has harmful unintended consequences. They point out that when students are divided on the basis of academic criteria, they also tend to be segregated by social and economic characteristics (e.g., Rosenbaum 1976; Oakes 1991). Moreover, critics assert that although educators may attempt to provide appropriate instruction for each group of students, in practice those in low-ability classes tend to receive inferior instruction compared to their high-ability peers (e.g., Oakes 1985; Page 1991). This unequal allocation of instruction may result in a widening of the achievement gap between high-level and low-level classes over time (Oakes, Gamoran, and Page 1992).

Does ability grouping lead to inequitable opportunities for students assigned to different groups? What aspects of instruction are comparable across groups, and what aspects differ? To what extent does the differential allocation or impact of instruction account for inequality of achievement among students assigned to different ability groups? We address these questions with data from 92 eighth- and ninth-grade English classes in 18 midwest-American secondary schools, focusing on the links between ability grouping, classroom instruction, and student learning. To set the stage for our analyses, we first present the logic of ability grouping from an organizational perspective. This discussion both clarifies the rationale for ability grouping and illuminates its potential shortcomings. In addition, it leads us to introduce a new approach for measuring classroom instruction.

Ability Grouping as an Organizational Response to Diversity

Sociologists characterize ability grouping as an organizational response to diversity among students (e.g., Sorensen 1970; Gamoran 1986). By dividing students into subgroups that are more homogeneous than the population as a whole, schools operate like many other types of complex organizations. As Thompson (1967, p.70) explained: "Under norms of rationality, organizations facing heterogeneous task environments seek to identify homogeneous segments and establish structural units to deal with each." By adherence to "rationality," Thompson meant that the organization can accomplish its goals more efficiently when it allocates separate tasks to specialized subunits. Instead of conducting all types of work throughout the organization, more limited and specialized tasks can occur in smaller subunits. In schools, the subunits typically include grade levels, and ability groups within grades. Faced with a diverse input population, educators divide students into relatively homogeneous categories so they can deal with different groups of students in different ways. This is the logic behind ability grouping, and it stands behind differentiation in many types of complex organizations.

Problems of Academic Differentiation

Following this logic creates two sorts of difficulties for educational organizations. One set of problems has to do with raw materials, and the other relates to technology.

Raw materials. In the organizational conception, students are thought of as the "raw materials" of school systems. However, sorting students is not a neutral act. If students are divided not only on the basis of prior academic performance, but also according to family background, race, and ethnicity -- even if the latter divisions are unintentional and result from conditions external to the school -- then differentiation for the sake of efficiency conflicts with the desire for social integration within schools. Moreover, because ability grouping creates a status hierarchy, the assignment of students to groups constitutes a status allocation in which some are elevated above others. When

ability-group assignment is correlated with socio-demographic factors, it reinforces status distinctions that originate outside the school. Thus, what is a "rational" procedure (in Thompson's terms) for most types of organizations is problematic in schools.

Research on the formation of ability groups and curricular tracks gives weight to these concerns, indicating that ability-group assignment is correlated with socioeconomic status, race, and ethnicity (see Oakes, Gamoran, and Page 1992 for a review). Even though these associations mainly dissipate once test scores are taken into account, they indicate that the allocation of status within schools tends to coincide with social status outside schools, a situation that critics of ability grouping find objectionable.

Technology. The key to differentiation is technology, that is, the materials and activities through which the organization works to accomplish its goals (Thompson 1967; Perrow 1986).[2] Technology is central to differentiation because a prime reason for creating separate subunits is to facilitate varied technological processes under varied conditions. From an organizational perspective, the technology of school systems is classroom instruction, for that is the mechanism through which learning occurs (Parsons 1960; Gamoran and Dreeben 1986). Thus, one would expect to find varied instructional activities and/or materials in different ability groups, with greater emphasis on aspects of instruction that are most beneficial in each context.

Despite the apparent logic of this arrangement, uncertainties about the operation and effects of instruction raise doubts about how well the rationale for differentiation applies in school systems. First, cause/effect relations between teaching and learning are not well understood or documented (e.g., Weick 1976; Brophy and Good 1986). There is little consensus about what constitute the best teaching methods, so it is difficult for educators to know precisely how to vary their teaching for different groups. Second, instruction is a complex technology, and it is insufficient to conceive of it as a one-way act in which teachers merely apply treatment to objects. Instead, teachers interact with

students, who are not inert raw material, but sentient, intentional subjects (Nystramd 1992). Teachers have the dominant role, but precisely what action they should take is far from clear, in part because their efforts are sensitive to the characteristics and reactions of students (Jackson 1968). Particularly as students reach adolescence, their actions and intentions as well as those of teachers influence the quality of their classroom experiences (Metz 1978; Everhart 1983). Hence, although ability grouping seems at first to be a straightforward application of the common organizational strategy of differentiation, in practice it may not work out as planned, and it may well have unintended consequences. In particular, the extent to which ability grouping fosters effective instruction in all types of classes is open to question.

Past research on ability grouping speaks to these concerns in two ways. First, several observers have described instructional differences across tracks. Dar and Resh (1986), for example, argued that sorting students by ability creates a resource-rich environment for high-group students and deprivation for students in low groups, because the intellectual capacities of classmates constitute important classroom resources. Moreover, in many schools the teachers with the best reputations are assigned to honors classes, and less-experienced and/or less-successful teachers are relegated to remedial classes (Finley 1984; Talbert 1990). Also, secondary school teachers appear to spend less time preparing and are less enthusiastic with low-ability classes (Rosenbaum 1976; Vanfossen, Jones, and Spade 1987). Instruction in low-ability classes tends to be more fragmented, dwelling on isolated bits of information, and it progresses at a slower pace (Oakes 1985; Page 1991). By contrast, students in high-ability classes are more likely to engage in critical-thinking tasks and problem-solving (Oakes 1985). Whereas some of these instructional differences might be consistent with the view that ability grouping helps match instruction to students' needs (e.g., the slower pace of instruction in low groups), for the most part they support the claims of critics who argue that instruction is not just different, but of higher quality in high-level classes and inferior in classes for remedial students.

Second, most writers on grouping and tracking have concluded that sorting students by academic performance typically contributes to increasing inequality of achievement (e.g., Findley and Bryan 1971; Rosenbaum 1980; Gamoran and Berends 1987; Murphy and Hallinger 1987). Studies of broad curriculum tracking show clear evidence of widening achievement gaps among students assigned to different tracks, even after accounting for initial differences (e.g., Gamoran and Mare 1989). Research on ability grouping for particular subjects is more equivocal, with some studies indicating a rise in achievement inequality, but others not (Slavin 1987, 1990). At the secondary level, national surveys of ability grouping in Britain (Kerckhoff 1986) and the United States (Hoffer 1992) indicate that the net gap between students in high and low groups increases over time. These results imply that a given student —whether of high or low initial ability —would perform better if he or she were placed in a higher-level class —consistent with the critics' view, and contrary to the organizational aims of ability grouping.

Juxtaposing the observers' findings about instructional differences across groups, and the survey researchers' conclusions about achievement differences, one may hypothesize that variation in the quality of instruction leads to increasing inequality of achievement (Gamoran and Berends, 1987. To date, only a few studies have addressed this question directly. At the elementary level, Barr and Dreeben (1983), Rowan and Miracle (1983), and Gamoran (1986) showed that greater progress through the curriculum accounts for more learning among students in high-ability reading groups, compared to students in other groups. Evidence at the secondary level is lacking, probably because of the complexity of measuring instruction in secondary-school classes. Whereas content coverage and time allocated to various activities may suffice to characterize elementary-school reading (Barr and Dreeben 1983), instruction at the secondary level potentially involves too many complex ideas and too much give-and-take between teachers and students to be adequately measured by a simple count of materials and activities introduced by the teacher. Measures of instruction that are more

sensitive to interaction between teachers and students are needed to test the hypothesis at the secondary level. Since that is a central purpose of this study, we have developed new measures of instruction.

<u>Measuring Instruction in Secondary School Ability Groups</u>

We conceive of instruction more broadly than simply the lecture, recitation, and coverage through which teachers inform their students. Rather than defining instruction as what teachers "do to students," we define it in terms of how teachers and students interact; and to measure the extent of their interaction, we focus on the quality of their instructional discourse.

The most obvious features of high-quality instructional discourse are high student participation and correspondingly low offtask behavior, which both result, of course, when students and teachers interact extensively. Though important, high student participation is not a sufficient measure of high-quality instructional discourse, because the level of student activity alone indicates little about the nature of student engagement. Students are procedurally engaged when they pay attention, do their assignments, and consistently conform to the requirements of school tasks. Students who are procedurally engaged, however, are not necessarily intellectually engaged in the issues and content of their studies. For this reason, we need to examine variation across classes in the substantive quality of teacher-student discourse.

In addition to student participation, one of the most important features of high-quality instructional discourse is its coherence (Nystrand and Gamoran 1991b). Some teachers carefully frame lessons and activities in terms of previous lessons and activities, and, as a result, classroom talk frequently refers to previous classroom talk. More than this, these teachers also have their students discuss what they have read, write about what they have read, read and discuss before writing, and so forth. During question-and-answer exchanges, effective teachers also follow up on student responses by incorporating previous student answers into subsequent questions in a process linguists call uptake

(Cazden 1988; Collins 1982, 1986). This continuous interweaving of writing, reading, and talk helps students relate topics of instruction, and reinforces and builds upon previous learning.

Some teachers also increase the coherence of their instruction by asking questions that build upon students' concerns and interests, and, as a result, skillfully help students relate these concerns to the content of instruction and learning. For example, instead of asking only test questions (questions which are characteristic of recitation and which teachers ask when they are looking for particular answers), skillful teachers also ask authentic questions, or questions for which the teacher avoids prespecifying answers (e.g., How did you like the chapter you read last night? Did the story end the way you expected? Who do you think was the most important character, and why?). Authentic questions are effective because they promote student ownership and help students coherently relate the new information of instruction to what they already know and/or have experienced. Authentic questions are also important because they signal to students the teacher's interest in what students think, as well as the importance the teacher attaches to thinking and not just remembering (Nystrand and Gamoran 1988). Another aspect of high-quality discourse is discussion, the free exchange of opinions and information among teachers and students, without continual prompting by questions from the teacher.

Ability grouping and instructional discourse. Prior research suggests that the quality of discourse is higher in high-track classes and lower in low tracks. Procedurally, more offtask behavior occurs in low-track classes, teachers spend more time on discipline and less time on instruction, and students spend less time on homework (Oakes 1985). Substantively, instruction in low-track classes is more often fragmented, emphasizing isolated bits of information instead of sustained inquiry (Page 1991). A pilot study indicated that students in low-ability 8th- and 9th-grade English classes answered true-false, multiple choice, and fill-in-the-blanks questions 4 to 5 times as frequently as did their high-group counterparts (Nystrand and Gamoran 1988; Gamoran 1989). In responding to the

papers of students in low-ability classes, teachers commented around twice as much about spelling, punctuation, and grammar, and about half as much about content, compared to teachers' responses to high-track students' papers. Although teachers met students in both low- and high-track classes about as infrequently in writing conferences (about once a month on average), they discussed spelling 2.6 times as much with students in low-track classes in these conferences, and content 1.9 times as frequently with high-track students.

Not only is there reason to believe that high-quality discourse occurs more often in high-ability classes, but such instruction may be most important just where it occurs least. Scholars who write about at-risk students emphasize the need to promote ownership and meaningfulness in schoolwork to counteract the alienation that is common for such students (Wehlage et al. 1989; Wehlage and Smith 1992). To the extent that authentic questions and discussion serve these ends, their positive impact may be greater in low-ability classes in comparison to high-ability classes, where students may be more motivated by external rewards such as grades (Newmann 1992).

In addition, student misbehavior occurs and is treated differently in high and low tracks. As Metz (1978) observed, when high-track students disengage from schoolwork, they do so in a way that still allows them to carry out the task at hand. Passing notes, reading unrelated books, and making humorous remarks occur in the context of making it through the school day, while still getting one's schoolwork done. Thus, disruptive behavior in honors classes is less likely to impede students from carrying out their work, in comparison to regular and especially low-track classes where offtask behavior is part of students' rejection of classwork. Moreover, teachers react differently to misbehavior in high-ability classes. According to Metz (1978), students who are loud or speak out of turn may be seen as overeager but worth engaging in honors classes, whereas similar behavior generates reprimands in low-track classes. For these reasons, procedural disengagement may impede learning more in low-ability than in high-ability classes.

This pattern contradicts the organizational rationale for ability grouping. According to the organizational logic, ability grouping facilitates the matching of instruction to students' needs. Thus, if authentic discourse and discussion are especially beneficial for at-risk students, one would predict greater reliance on this approach in low-track classes. Yet available evidence suggests the opposite may be true. Similarly, if misbehavior is more harmful to low-ability students, a reason for differentiation would be to minimize disruptions in the learning contexts of low achievers; yet that does not seem to be the case, either. If the organizational aims of grouping are not fulfilled, as research to date seems to indicate, then differences in both the nature and the effects of instruction may explain, at least in part, why achievement gaps between ability groups widen over time.

The main purpose of the present study is to assess variation in the quality and the impact of instructional discourse across ability groups. Does the quality of instruction favor high- over low-ability groups? Does disruptive behavior occur more and cause more damage to achievement in low groups than in high groups? Addressing questions such as these will allow us to discern the extent to which instruction mediates the effects of ability grouping on achievement. In this way, we assess the hypothesis that instructional differences are the source of achievement differences across ability groups.

Data and Methods

The sample for this paper comes from a two-year study of 25 secondary schools. The schools were located in 9 communities in the American midwest, including rural, urban, and suburban areas, and public and Catholic schools. Overall about four English classes per school participated in the study, but this varied by the size of the school: 58 eighth-grade classes were distributed among 16 middle and junior high schools studied in 1987-88, and 54 ninth-grade classes were studied the following year in nine high schools for which the middle schools served as feeders. In smaller schools, all classes participated, and in larger schools, classes were selected to represent the different

ability-group levels defined by the school (e.g., honors or accelerated, regular, and basic or remedial). About 90 percent of students in the selected classes participated in the study.3

The analysis is restricted to 92 high, regular, and low classes in 10 junior high/middle schools and 8 high schools. Heterogeneous classes were excluded from the present analysis for three reasons: (1) In the ninth-grade study, heterogeneous classes were used only in a small, rural school, and a school-within-a-school in an urban school, so homogeneous/heterogeneous differences were confounded with school differences; (2) Standardized test scores, which serve as "ability" measures to help control for pre-existing differences among students from different tracks, did not exist for most of the eighth-grade heterogeneous classes; (3) The main issue for this paper is not the difference between homogeneous and heterogeneous classes, but the differences in the distribution and effects of instruction among the homogeneous classes. The 18 schools remaining in the analysis included two urban high schools and their three feeder junior highs in an ethnically diverse, mainly working-class area; one urban high school in a less diverse, more middle- to upper-middle-class locale; one suburban school and two feeder middle schools in an upper-middle-class community; one high school and one junior high from each of two small town/rural districts; and two Catholic high schools with three feeder K-8 schools, which served urban and suburban, predominantly middle- and upper-middle-class white students.

We visited each class four times, focusing mainly on the time spent in different activities and on the questions asked by teachers and students (see below). Students took tests and filled out questionnaires in the fall and spring, and teachers also filled out questionnaires in the spring. Of 1,968 students who began the year in the 92 classes, 1,750 (89 percent) participated in the study in the fall and spring. Listwise deletion of student-level missing data reduced the analysis sample to 1,564 students (89 percent of study participants, 79 percent of the total).

Background and Achievement Data

We measured learning with a year-end test of literature achievement. Because we assessed instruction as the quality of instructional discourse, we designed a test that required students to engage in discourse about the material they had covered during the year. The chances for detecting the effects of schooling are greater if one tests students on what they were actually taught, rather than on a standardized body of information (Walker and Schaffarzick 1974).

The test posed a series of questions about the novels, short stories, and plays that were assigned during the year. For each class, we selected five readings that had been covered, choosing items that were representative of the overall curricula. The questions ranged from simple recall ("Describe the ending of the story") to ones requiring in-depth understanding ("Relate the conflict of the story to its ending and theme"). The questions were the same for each class, but the stories differed, depending on what had been covered during the year. Each test was scored by two trained readers on dimensions such as extent of recall, depth of understanding, understanding of characters' motivations, and so on. When the scores differed by more than one point on any given dimension, the test was given a second reading. Scores from the two readers were averaged, and inter-rater reliability was calculated as correlations of .90 in the eighth-grade sample and .82 in the ninth-grade group. Means and standard deviations of this and all other variables are listed in Table 1.

Prior reading and writing skills. We administered two tests at the beginning of the year to account for prior differences among students in reading and writing skills. One was a multiple-choice test of reading comprehension, based on National Assessment of Educational Progress (NAEP) items. The eighth and ninth graders read different stories, but the results were calibrated on similar scales. This test also included a brief writing sample. The second test consisted of a fifteen-minute essay, for which eighth graders were asked to write about a person or event that was important to them, and the ninth graders about a special place or possession. This test was scored by two readers on level of

abstraction (Britton et al. 1975) and coherence of argumentation (Applebee, Langer, and Mullis 1985), and the marks were summed across dimensions and averaged across readers. The inter-rater correlation was .70.

"Ability". From school records, we obtained data on student performance on standardized tests administered by the districts. We recorded national percentile scores, which we transformed to normal curve equivalents. Unfortunately, the districts employed several different instruments, and while most were administered in the spring of the previous year, some were given in the previous fall, a full year before our arrival. This would not matter much if all the scores were truly normed to the national population, but the extent to which that is the case is unknown. To account for measurement error introduced by the standardized tests, we used the scores not as distinct variables, but as indicators of a common underlying trait, which we termed "ability." For each student, ability was indicated by a math score and a reading comprehension score. The measurement model for this latent variable yielded reliability estimates of about .54 for the math score and .44 for the reading score. These values are lower than is typical for such tests, presumably due to differences across districts.[4]

We used the ability measure despite its problems because of the danger that the effects of ability grouping could be inflated by unmeasured differences among students assigned to the different groups. Slavin (1990) has argued that all observed effects of grouping in correlational studies are likely due to such selection bias. While selection bias can never be completely ruled out in the absence of random assignment, the present study offers a more rigorous set of controls than has been used in nearly all comparable studies. In research on high school tracking with another rich data set, Gamoran and Mare (1989) found that using a similar set of controls eliminated the correlation between unobserved selection factors and outcomes.

Other background variables. Further controls for student background differences were indicated by dummy variables for sex (1 = female, 0 = male) and minority status (black or Hispanic = 1, others = 0). Last, student socioeconomic status was indicated by an unweighted linear composite of father's education, mother's education, the higher in status of father's or mother's occupation, and the availability of a list of home resources. These background data were drawn from student questionnaires.

Indicators of Ability-Group Positions

Recent writers have criticized survey studies of grouping and tracking for using ambiguous indicators of track positions (Gamoran 1989; Lucas 1990; Lucas and Gamoran 1991). U.S. studies of national data typically rely on student self-reports of whether their programs are best described as academic, general, or vocational. Although this indicator is useful when tracking is viewed as a social-psychological construct (see Berends 1992), its value as a structural indicator is limited. This is not so much because students may be incorrect; other data sources also carry the danger of unreliability (Gamoran and Berends 1987). Rather, the ambiguity of the survey indicator stems from an underlying assumption: that virtually all schools are in fact divided into academic, general, and vocational programs. Yet recent observers report that such programmatic tracking has waned, at least in formal terms (Oakes 1985; Moore and Davenport 1988). Instead, students in both junior and senior high schools tend to be stratified by performance on a subject-by-subject basis.

Whereas a student's track position (e.g., academic or general) is often ambiguous, there is little disagreement about the ranking of courses within a particular subject. In the case of English, the great majority of secondary schools distinguish among levels such as honors or accelerated, regular or average, and basic or remedial (Oakes 1985; Moore and Davenport 1988). For this study, English classes are categorized as honors (including classes labeled high, advanced, and accelerated), regular, and remedial (including classes termed low and basic). These categories were unambiguously

described by school staff. Student membership in particular classes was taken from class rosters and was verified by classroom teachers. The sample was not large enough to distinguish among schools having two, three, or four ability levels, but the grouping systems were similar across schools in that students were assigned to particular English classes based on judgments about their performance in English, rather than on how well they do in all subjects overall (see Slavin 1987, on the importance of this similarity). In four cases, teachers divided their time between two groups in a single room.

Measures of Instruction

For this study we have relied on seven key indicators of instructional discourse. More indicators were available in the data, but we narrowed our focus on the basis of preliminary exploratory factor analyses, inspection of reliabilities in confirmatory factor analyses, and the theoretical centrality of particular indicators (Gamoran, Berends, and Nystrand 1990; Nystrand and Gamoran 1991a). In an early analysis of the eighth-grade data, we had used a composite indicator of discourse quality (Gamoran and Nystrand 1990), but we learned subsequently that there was no single underlying factor that incorporated the diverse measures (Gamoran, Berends, and Nystrand 1990). Consequently, we now use the seven variables as indicators of distinct aspects of discourse quality.

We obtained three measures of student participation. Two came from the spring student questionnaire: students' reported frequency of completing their writing and reading assignments. (See the appendix for the wording of questionnaire items.) The other came from classroom observations: the percentage of students visibly offtask during question-answer sessions.

For discourse coherence, we used a composite of six teacher-questionnaire items that asked about the interconnections among different classroom activities: the extent to which teachers asked students to (a) write about what they read; (b) discuss writing topics before writing; (c) discuss readings; (d) relate readings to other readings; (e) relate discussions to previous discussions; and (f) discuss what other students have written about (see appendix).

Uptake was computed as the percentage of questions that followed up on what someone had said previously, averaged over the four observations. In the following exchange, for example, the teacher's second question uses uptake:

Teacher:	Why did Atticus need Aunt Alexandra at this time?
Student:	To keep Scout safe.
Teacher:	Why would Scout be safe with Aunt Alexandra?

In this dialogue, which occurred during a discussion of <u>To Kill a Mockingbird</u> in a ninth-grade class, the teacher had specific answers in mind. Even though she was asking the student to draw conclusions rather than simply recite the story, her questions tested students' knowledge and use of information instead of encouraging them to construct new ideas, and we refer to them as "test questions." In contrast, authentic questions treat students' ideas as legitimate knowledge in their own right. A bit later in the lesson, this teacher asked students to speculate about alternative paths the story might have taken: "What are some ideas for Atticus not having Aunt Alexandra come?" Here, she was asking an authentic question, showing interest in students' ideas rather than testing for a prespecified answer. We computed the percentage of teacher questions that were authentic, averaged across the four observations, as another indicator of discourse quality.

Finally, we counted the number of minutes per day devoted to discussion. Discussion is defined more narrowly than simply teacher-student discourse; it refers to the free exchange of information among teachers and students, without the usual question-response-evaluation structure of ordinary recitation (Mehan 1979). Often during discussion, students speak to one another without interruption by the teacher (Nystrand and Gamoran 1991b). We focused on discussion because from the standpoint of instructional discourse, it is qualitatively different than other classroom activities which are heavily dominated by teachers.

<u>Statistical Models</u>

Our initial questions are descriptive. They concern the compositions of the different groups and the differences among groups in the quality of instructional discourse. Subsequently, we turn to the analytic questions of net achievement differences between groups, the effects of instruction on achievement, and the extent to which variation in the distribution and effects of instruction produce ability-group differences in achievement.

To address the analytic issues, we used maximum likelihood methods. We chose this approach because it permitted us to specify the latent "ability" construct described above. It also provided tests for the comparative fits of various alternative model specifications. We divided the data into the three groups: honors, regular, and remedial classes. First, we set aside the instructional data and estimated models of ability-group differences in achievement net of the exogenous variables (sex, minority status, SES, fall reading and writing performance, and ability). After selecting a baseline model and estimating achievement differences between groups, we added the instructional data to the model. We compared the fits of models in which the instructional variables were constrained to be the same across ability groups with models that permitted different effects in different groups. After selecting the best-fitting model, we re-examined the achievement gaps between groups under various instructional circumstances.

<div align="center">

Results

</div>

Does ability grouping divide students on more than solely academic dimensions? Previous writers have maintained that minority students and economically disadvantaged students are overrepresented in low-status groups and tracks (e.g., Oakes 1990), and our data conform to that pattern. As Table 1 shows, whereas the sample as a whole consisted of nearly 20 percent minority students, honors classes had just half that proportion while remedial eighth- and ninth-grade English classes averaged more than twice the total sample mean. The contrast is even greater if we focus on

the district in our sample with the highest proportion of minority students. In that district, located in a working-class urban area, 52 percent of the students were black or Hispanic, but the proportion minority was 26 percent in the honors classes, 52 percent in regular classes, and 65 percent in remedial classes. Similar patterns appear for the social class composition of the ability groups: In the total sample, honors classes averaged .37 standardized units above the mean in SES while low-ability classes stood at .42 standardized units below the mean.

These findings are far from new, and they cannot be used as evidence that assignment procedures were discriminatory. As previous studies have shown, the direct impact of socio-demographic conditions on track assignment is small, compared to the overwhelming importance of academic performance (e.g., Gamoran and Mare 1989; Gamoran 1992). The point here is to show that sorting students is not neutral with respect to social and economic inequality: the allocation of status within schools tends to coincide with status differences in the wider society, and if ability grouping magnifies achievement inequality, then it tends to reinforce initial differences (see also Vanfossen, Jones, and Spade 1987; Gamoran and Mare 1989; Lucas and Gamoran 1991).

Distribution of Instruction among Ability Groups

Are there inequities among ability groups in the conditions of instruction? Table 2 displays class-level means of instructional variables for the different types of classes. As expected, students in honors classes exhibit the most consistent participation, and students in remedial classes are least engaged in their schoolwork. These findings replicate those of other studies, both in their consistent patterns and in that the differences, while statistically significant, are not large substantively (Oakes 1985; Gamoran and Berends 1987).

In contrast, most aspects of instructional discourse did not differ significantly between class types. If anything, regular classes contained higher proportions of authentic questions and questions with uptake, as well as a higher degree of coherence, but these differences are not statistically

significant. Only discussion favored honors classes over other classes, but it was not a common occurrence even there, averaging only about 75 seconds per day. Authenticity and uptake were also infrequent, with less than a quarter of questions having one or both of these qualities in regular classes, and smaller proportions elsewhere. Coherence was reported to be more common, with teachers citing an average of about twelve activities per week, or more than twice each day, in which reading, writing, and classroom talk were interwoven.

The results are consistent with descriptions of classroom discourse as dominated by teachers and emphasizing reproduction rather than production of knowledge (Mehan 1979; Goodlad 1984). We did not find evidence of especially fragmented and recitation-oriented instruction in low-ability classes. Although we observed significantly more discussion in high-ability classes, it remains to be seen whether this difference is related to achievement gaps in light of its infrequency. Similarly, even though participation was greater in honors classes and lower in remedial classes, it is not yet clear whether these differences help account for achievement gaps between ability groups. An additional possibility is that differences between groups in the effects of instruction, rather than differences in instructional means, produce differential achievement. To assess the importance of the findings so far we must examine the contributions of instructional conditions to student learning.

The Effects of Ability Grouping on Achievement

Before bringing together grouping, instruction, and learning, we need first to determine whether students in different types of classes obtained varied achievement, net of pre-existing conditions. To address this question, we estimated a model in which the effects of all background variables were constrained to be equal across honors, regular, and remedial classes. This model fit the data reasonably well, with a chi-square of 55.89 and 32 degrees of freedom.[5] Table 3 shows that each of the background conditions contributes significantly to literature achievement. Girls, non-minorities, and high-SES students scored higher than boys, minority students, and the economically

disadvantaged, respectively. Students with higher initial test scores and higher estimated ability also performed better on our test at the end of the year.

Since all the effects were constrained to be equal, the only differences among models for the three groups are in the intercepts. Consequently, the intercepts reveal differences in achievement between groups, net of background conditions. These show gaps of .843 points between the honors and regular classes, and another 1.147 points between the regular and low-ability classes. These differences are not large – they constitute about 12 percent and 17 percent, respectively, of the total-sample standard deviation – but because they occurred within a single academic year, they need to be taken seriously.[6]

To test for statistical significance of ability-group differences in achievement, we cannot compare the intercepts with their standard errors; that tests whether the intercepts differ from zero, and we need to test whether they differ from each other. This question is addressed by comparing this model to another in which the intercepts are constrained to be equal across groups. We estimated the equal-intercept model and found that its fit was significantly poorer, yielding a chi-square of 66.62 with 34 degrees of freedom. The chi-square difference between these two nested models is 10.73, with 2 degrees of freedom, a difference that is significant at p < .01. Hence, we conclude that the type of class students attended made a small but significant difference for their achievement.

Ability Grouping, Instruction, and Achievement

To what extent were the achievement differences produced by variation in the distribution and effects of instruction? We first included the instructional variables using the same specification as we used for the background variables – that is, no differences between class types in the effects of instruction on achievement.[7] Fit statistics for this model are presented in the first row of Table 4. This model fit the data fairly well, but we had reason to question the assumption of equal instructional effects across groups. Our conceptual formulation, and some preliminary analyses, suggested that

offtask behavior, authenticity, and discussion might have varied effects, and we estimated this model next. As shown in the second row of Table 4, this model fit significantly better. Subsequent modifications, however, failed to improve the fit. Hence, the data suggest that completion of reading and writing, coherence, and uptake exert similar effects in honors, regular, and remedial classes, but the effects of offtask, authenticity, and discussion vary by class type.

Effects of instruction on achievement. Table 5 displays the results of the best-fitting model. Each of the variables with similar effects across groups contributes positively to achievement: students who report completing more of their reading and writing scored higher, as did those in classes with more uptake and more coherence among instructional activities. The effects of the other instructional variables are more complex: Offtask behavior led to lower achievement in regular and remedial classes, but not in honors classes; authentic questions resulted in higher achievement in honors classes but lower achievement in remedial classes; and discussion benefitted honors students but reduced achievement for those in regular classes.

The participation variables appear to exert substantial effects. For example, a 10 percent increase in offtask behavior would reduce achievement by about one and a quarter to nearly two points in the remedial and regular classes, a loss of almost 20 percent to 30 percent of a standard deviation. Similarly, students who did half their reading and writing assignments would score more than two points lower than those who completed all their work, other things being equal. Among the discourse variables, the impact of coherence stands out: A class with twice the average level of coherence (a difference of almost two standard deviations) would raise achievement by nearly two points on average, other things being equal. This is especially important because activities that promote coherence occur with regular frequency, according to teachers, although they are more common in some classes than others. Effects of other discourse variables are more modest, but still large enough to be substantively as well as statistically meaningful. In light of these instructional

effects, we can ask whether differences in levels and/or effects of instruction account for achievement differences across ability groups.

Accounting for achievement gaps: Variables exerting the same effects in all groups. Not only do completion of reading and writing assignments contribute to achievement overall, but they also contribute to achievement gaps between ability groups. This is because students in honors classes tended to complete higher proportions of their work than students in regular classes, who finished more of their work than those in remedial classes (see Table 2). If these conditions were more equal across classes, our results imply that achievement would be more evenly distributed.

As noted above, coherence and uptake have significant positive effects on literature achievement. However, they do not contribute to achievement gaps, because they do not vary significantly or consistently among the three class types.

Accounting for achievement gaps: Variables exerting different effects in different groups. Authentic questions, discussion, and offtask behavior do not have the same effects on literature achievement at each ability-group level. Whether or not they contribute to gaps in achievement depends on how frequently they occur in honors, regular, and remedial classes. At the very least, they demonstrate that the same type of discourse can result in unequal achievement in the different types of classes. The greater the incidence of each, the wider the achievement gaps. For example, if there were no authenticity, discussion, or offtask behavior, the intercepts would capture all of the differences in achievement, suggesting little difference between regular and low groups (-7.081 - -7.144 = .063) and higher achievement in regular than high classes (-8.502 - -7.081 = -1.421). At low levels, say 15 percent authenticity, half a minute per day in discussion, and no offtask, achievement would be roughly similar across all class types. More realistically, when instructional conditions are at the averages for all classes, achievement is similar in regular and low

classes but about two points higher in high-ability classes. Hence, achievement gaps result from a combination of differences in the levels and the effects of instructional conditions.

Interpreting the varied effects. How should we interpret the differences across groups in the effects of offtask behavior, authenticity, and discussion? Results for offtask were anticipated by previous research: We predicted that offtask behavior might be more harmful in lower-status classes, because such activity reflects resistance to schooling in such classes, whereas in honors classes misbehavior does not necessarily indicate rejection of schoolwork. In addition, high-group students who are not themselves offtask may be less distracted by offtask behavior than students in regular and remedial classes.

Although the data are consistent with earlier work, they conflict with the organizational goals of ability grouping, which call for alignment of classroom conditions to students' needs. Instead of less offtask where it is more damaging, we observed less offtask behavior where it was inconsequential. We cannot interpret this finding conclusively, however, because we do not know whether the low-ability students would have been similarly disruptive in a mixed-ability class. Perhaps their levels of misbehavior would be even higher in a heterogeneous context, and since our models do not speak to that possibility, we cannot fully dismiss this aspect of the rationale for ability grouping.

In light of the positive (though non-significant) coefficient for offtask in high groups, another interpretation must be considered: Offtask behavior may occur in honors classes after students have mastered the material. Perhaps students who have figured out the answers and completed their work are afterwards likely to relax and misbehave. In that case, high achievement may lead to offtask behavior, rather than the reverse. This interpretation challenges the causal ordering specified in our model.

Authenticity occurred with roughly similar frequency across class types, yet it was beneficial in high-ability classes and detrimental in low groups. This finding is not consistent with the organizational logic of differentiation, which encourages greater emphasis on activities where they are more effective, but it is also inconsistent with our speculation based on prior research that authentic discourse offers greater benefits in low-ability classes than elsewhere. We found just the opposite. Results for discussion were more closely aligned with the organizational perspective, in that high-ability classes had more discussion time and also derived greater benefits from discussion, but this finding contradicted our expectation that discussion would benefit low-ability students most of all.

How did these results come about? To better understand the pattern, we returned to the data to examine the content of authentic questions and the contexts of discussions in different classes. We discovered that teachers in honors English classes were much more likely to ask authentic questions about literature, whereas authentic questions in remedial English classes pertained to a wide variety of topics. One teacher in a remedial class, for example, asked authentic questions about test-taking: "How do most of you feel about test-taking?" Another example was brainstorming: "What things would you associate with lying in the sun?" By contrast, authentic questions in high-ability English classes generally focused on ideas and issues found in literary texts. Overall, we counted 73.4 percent of authentic questions had to do with literature texts in honors English classes, but only 31.3 percent of authentic questions in remedial English classes concerned the texts students were reading.

The pattern for discussion was similar but the interpretation is less clear-cut. Almost all the instances of discussion in honors English classes concerned literature, whereas only half the discussions in remedial English classes were about texts students were reading. However, two-thirds of the discussions in regular English classes were on literary texts. Thus, we are not able to account for the substantial negative effect of discussion in the regular classes. Discussion was so infrequent in

regular classes, however, averaging only 12 seconds per day (see Table 2), that the negative coefficient may not be meaningful.

Like the results for offtask, the pattern for authenticity could also be interpreted as reflecting a mis-specified causal sequence. This interpretation would suggest that authentic questions about topics other than literature are the teacher's response to, rather than a contribution to, low-track students' poor performance in literature. At present, we are unable to test among these competing causal chains. More generally, we cannot test whether high-quality instruction produces higher achievement, or higher achievement leads to better instruction. Our model does not presume a causal order between participation and discourse variables —we assume these conditions are interrelated — but on the basis of our controls for prior ability and achievement, we have assumed that instructional conditions affect year-end achievement rather than the reverse.

Conclusions

The results of this study indicate that differences in the nature and effects of classroom instruction constitute an important part of the explanation for widening achievement gaps among students assigned to classes at varied levels. The findings are thus consistent with a long-standing hypothesis, although the particular pattern of results complicates the story. As has been widely assumed (but not documented empirically), the greater conformity to instructional demands among students in honors classes accounts for part of their achievement advantage, compared to students in remedial classes. Not only do high-ability students turn in their assignments more often, but they are less often offtask; and yet offtask behavior is detrimental only where it is more frequent, in middle- and low-ability classes.

Differences in the quality of instructional discourse were smaller than we expected on the basis of prior research. Only discussion time was clearly more evident in honors classes than elsewhere, and its absolute level was small in all types of classes. This difference turned out to be

potent for achievement inequality, however, because discussion only benefitted students in the high-level classes. Authenticity was also consequential for achievement gaps, but not in the way originally expected: It occurred with similar frequency across classes, but was beneficial to high-ability students and detrimental to those in low-ability classes. This occurred, we believe, because of a key difference in the context of authentic discourse: whereas authentic questions in honors classes generally concerned literature, this was not the case in remedial classes.

On the one hand, one might argue that the types of authentic questions and discussions that occurred in low-ability classes were just what was called for. Perhaps by holding brainstorming sessions, talking about test-taking, and so on, these classes were meeting students' needs. Although these encounters failed to improve (and perhaps impeded) literature achievement, they may well have contributed in other areas. On the other hand, this conclusion admits defeat in the effort to engage low-achieving students in serious academic work. Students in remedial classes were not denied access to authentic questions, but they had far fewer opportunities to address such questions in the context of literature, one of the major foci of secondary school English. Hence, it was not the interactive style but the content of the interaction that favored honors over regular and remedial classes.

To the extent that ability grouping continues to be used, our results suggest that achievement inequality could be reduced by raising the caliber of both instructional content and instructional discourse in regular and remedial classes. In addition, the findings indicate that high-quality instructional discourse, characterized by student participation, coherence, discussion, authenticity, and uptake, can improve student learning when it occurs in the context of substantive academic content. The data failed to support our speculation that authentic discourse and discussion are especially beneficial to academically at-risk students – perhaps not because such instruction is ineffective, but because little of this type of instruction occurred in low-ability classes studying literature -- and this remains, in our view, an open question.

Although schools follow a common organizational procedure when they divide students on academic criteria, the procedure engenders atypical consequences, because sorting students is not a neutral act, and because instruction is an uncertain and complex technology. Ability grouping divides students on social as well as cognitive characteristics, so by magnifying achievement inequality it contributes to overall achievement inequality among social groups. Moreover, the teachers in our sample did not establish effective instructional conditions in all types of classes; some aspects of instruction were equally effective and equally distributed (coherence and uptake), but other forms were unequally distributed (student completion of work) or exerted unequal effects (authentic questions) or both (offtask behavior and discussion). This pattern indicates that the practice of ability grouping must be reconsidered, and if not replaced with other organizing principles -- an option to which these analyses do not speak --then the quality of experiences in regular and remedial classes must be improved -- as clearly indicated by the results of this study.

Notes

[2] We refer to "technology" in the organizational sense, and not merely in reference to electronic aids to instruction, as the term is used in the educational vernacular.

[3] About half the cases in the analysis are students who were included in the study twice, once as eighth graders and a second time in ninth grade. These students are represented twice in the data set. Although this may artificially increase the correlations among the predictors to some degree, the increase does not appear serious. The eighth- and ninth-grade data were obtained in separate years, and measures of classroom instruction, the key independent variables, were completely independent from one year to the next. We could find no meaningful differences between students who participated once and those who participated twice (Berends, 1992), and we gain much statistical power by pooling the data across grades. We also tested for differences between grades in the effects of the background variables on achievement, and found no significant differences.

[4] The schools were less successful than we were at obtaining data from all students; about 15 percent of students for whom we had complete data on background and prior achievement lacked standardized test results. Scores for these students were imputed on a district-by-district basis from our background and prior achievement data.

[5] The fit of the model could be improved slightly by allowing all background variables to have different effects across groups (chi-square difference = 24.62, d. f. difference = 12, difference is significant at $p = .016$). We chose to estimate the more constrained model because (a) we had no

strong prior grounds for predicting between-group differences in effects of background variables; (b) the relaxed model would greatly complicate the estimation of track effects; and (c) after instructional variables are added, the improved fit from allowing varied background effects is not statistically significant (chi-square difference = 19.01, d.f. difference = 12, p = .088).

[6] All maximum-likelihood estimates were obtained using LISREL (Joreskog and Sorbom, 1987). We compared these results to an ordinary least squares (OLS) regression in which the standardized test scores were included as single-item variables. This analysis yielded track effects that were considerably larger, at close to one and a half points for each gap. Hence, our model is a more conservative test for track effects, and probably does a better job of accounting for pre-existing differences among students assigned to different types of classes, compared to an OLS analysis.

[7] To simplify the model, we did not specify causal paths from the background variables to the instructional indicators, but left these relations as zero-order correlations. This specification does not affect the estimation of direct effects of background and instruction on achievement.

References

Applebee, A., J. Langer, and I. Mullis. 1985. Writing: Trends across the Decade. 1974-84. Princeton: Educational Testing Service.

Barr, Rebecca, and Robert Dreeben. 1983. How Schools Work. Chicago: University of Chicago Press.

Berends, Mark. 1992. Effects on School Orientations and Achievement in the Transition from Middle to High School. Unpublished Ph.D. thesis. Department of Sociology, University of Wisconsin-Madison.

Britton, J., T. Burgess, N. Martin, A. McLeod, and H. Rosen. 1975. The Development of Writing Abilities, 11-18. London: Macmillan.

Brophy, Jere, and Thomas L. Good. 1986. "Teacher Behavior and Student Achievement." In Handbook of Research on Teaching, 3rd edition, edited By M. C. Witrock. New York: Macmillan.

Cazden, Courtney. 1988. Classroom Discourse: The Language of Teaching and Learning. Portsmouth, NH: Heinemann.

Collins, J. 1982. "Discourse Style, Classroom Interaction, and Differential Treatment." Journal of Reading Behavior 14:429-437.

Collins, J. 1986. "Differential Instruction in Reading." In The Social Construction of Literacy, edited by J. Cook-Gumperz. Cambridge: Cambridge University Press.

Dar, Yehezkel, and Nura Resh. 1986. "Classroom Intellectual Composition and Academic Achievement." American Educational Research Journal 23:357-374.

Everhart, Robert. 1983. Reading, Writing, and Resistance: Adolescence and Labor in a Junior High School. Boston: Routledge and Kegan Paul.

Findley W. G., and M. M. Bryant. 1971. Ability Grouping: 1970. Athens, GA: Center for Educational Improvement.

Finley, Merilee K. 1984. "Teachers and Tracking in a Comprehensive High School." Sociology of Education 57:233-243.

Gamoran, Adam. 1986. "Instructional and Institutional Effects of Ability Grouping." Sociology of Education 59:85-198.

_____. 1987. "The Stratification of High School Learning Opportunities." Sociology of Education 60:135-155.

_____. 1989. "Measuring Curriculum Differentiation." American Journal of Education 97:129-143.

_____. 1992. "Access to Excellence: Assignment to Honors English Classes in the Transition from Middle to High School. Educational Evaluation and Policy Analysis 14:185-204.

Gamoran, Adam, and Mark Berends. 1987. "The Effects of Stratification in Secondary Schools: Synthesis of Survey and Ethnographic Research. Review of Educational Research 57:415-435.

Gamoran, Adam, Mark Berends, and Martin Nystrand. 1990. "Classroom Instruction and the Effects of Ability Grouping: A Structural Model." Paper presented at the annual meeting of the American Educational Research Association, Boston.

Gamoran, Adam, and Robert Dreeben. 1986. "Coupling and Control in Educational Organizations." Administrative Science Quarterly 31:612-632.

Gamoran, Adam, and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Reinforcement, Compensation, or Neutrality?" American Journal of Sociology 94:1146-1183.

Gamoran, Adam, and Martin Nystrand. 1990. "Tracking, Instruction, and Achievement." Paper presented at the World Congress of Sociology, Madrid.

Goodlad, John I. 1984. A Place Called School. New York: McGraw-Hill.

Hoffer, Thomas. 1992. "Middle School Ability Grouping and Student Achievement in Science and Mathematics." Educational Evaluation and Policy Analysis 14: 205-227.

Jackson, Philip W. 1968. Life in Classrooms. Chicago: University of Chicago Press.

Joreskog, Karl G., and Dag Sorbom. 1987. LISREL 7 User's Reference Guide. Morehead, IN: Scientific Software.

Kerckhoff, Alan C. 1986. "Effects of Ability Grouping in British Secondary Schools." American Sociological Review 51:842-858.

Lucas, Samuel R. 1990. "Course-Based Indicators of Curricular Track Locations." Unpublished M.S. thesis, University of Wisconsin-Madison.

Lucas, Samuel R., and Adam Gamoran. 1991. "Race and Track Assignment: A Reconsideration with Course-Based Indicators of Track Locations." Paper presented at the annual meeting of the American Sociological Association, Cincinnati.

Mehan, Hugh. 1979. Learning Lessons. Cambridge, MA: Harvard University Press.

Metz, Mary H. 1978. Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools. Berkeley, CA: University of California Press.

Murphy, Joseph, and Philip Hallinger. 1989. "Equity as Access to Learning: Curricular and Instructional Treatment Differences." Journal of Curriculum Studies 19:341-360.

Moore, Donald R., and Suzanne Davenport. 1988. The New Improved Sorting Machine. Madison, WI: National Center on Effective Secondary Schools.

National Education Association. 1990. Academic Tracking. Report of the NEA Executive Committee Subcommittee on Academic Tracking. Washington, DC: National Education Association.

Newmann, Fred M., editor. 1992. Student Engagement and Achievement in American Secondary Schools. New York: Teacher College Press.

Nystrand, Martin. 1992. "Dialogic Instruction and Conceptual Change." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Nystrand, Martin, and Adam Gamoran. 1988. "A Study of Instruction as Discourse." Paper presented at the annual meetings of the American Educational Research Association.

_____. 1991a. "Student Engagement, Instructional Discourse, and Literature Achievement." Research in the Teaching of English 25:261-290.

_____. 1991b. "Student Engagement: When Recitation Becomes Conversation." Pp. 257-276 in Contemporary Research on Teaching. edited by H. Waxman and H. Walberg. Berkeley, CA: McCutchan.

Oakes, Jeannie. 1985. Keeping Track: How Schools Structure Inequality. New Haven, CT: Yale University Press.

_____. 1990. Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science. Santa Monica: RAND.

Oakes, Jeannie, Adam Gamoran, and Reba N. Page. 1992. "Curriculum Differentiation: Opportunities, Outcomes, and Meanings." Pp. 570 -608 in Handbook of Research on Curriculum. edited by P. W. Jackson. Washington, DC: American Educational Research Association.

Page, Reba N. 1991. Lower Track Classrooms: A Curricular and Cultural Perspective. New York: Teachers College Press.

Parsons, Talcott. 1960. Structure and Process in Modern Society. Glencoe, IL: Free Press.

Perrow, Charles. 1986. Complex Organizations: A Critical Essay. Third edition. New York: McGraw-Hill.

Rosenbaum, James E. 1976. Making Inequality: The Hidden Curriculum of High School Tracking. New York: Wiley.

Rowan, Brian, and Andrew W. Miracle, Jr.. 1983. "Systems of Ability Grouping and the Stratification of Achievement in Elementary Schools." Sociology of Education 56:133-144.

Slavin, Robert E. 1987. "Ability Grouping and Achievement in Elementary Schools: A Best-Evidence Synthesis." Review of Educational Research 57:293-336.

_____. 1990. "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." Review of Educational Research 60:471-499.

Talbert, Joan. 1990. "Teacher Tracking: Exacerbating Inequalities in the High School." Paper presented at the annual meeting of the American Educational Research Association, Boston.

Thompson, James D. 1967. Organizations in Action. New York: McGraw-Hill.

Vanfossen, Beth E., James D. Jones, and Joan Z. Spade. 1987. "Curriculum Tracking and Status Maintenance." Sociology of Education 60:104-122.

Walker, David F., and Jon Schaffarzick. 1974. "Comparing Curricula." Review of Educational Research 44:83-111.

Weick, Karl E. 1986. "Educational Organizations as Loosely Coupled Systems." Educational Administration Quarterly 21:1-19.

Wehlage, Gary G., Robert A. Rutter, Gregory A. Smith, Nancy Lesko, and Ricardo R. Fernandez. 1989. Reducing the Risk: Schools as Communities of Support. London: Falmer.

Wehlage, Gary G., and Gregory A. Smith. 1992. "Building New Programs for Students At-Risk.
  Pp. 92-118 in Student Engagement and Achievement in American Secondary Schools, edited
  by F. M. Newmann. New York: Teachers College Press.

Wilson, Barry J. and Donald W. Schmits. 1978. "What's New in Ability Grouping?" Phi Delta
  Kappan 59:535-36.

.

APPENDIX
Wording of Questionnaire Items

<u>Student questionnaire</u>

Completion of reading: "About how often do you <u>complete your reading assignments</u> for this class?"

Completion of writing: "About how often do you <u>complete your writing assignments</u> for this class?"

> Response categories: Never, almost never, less than half the time, about half the time, more than half the time, most of the time, every time. Responses in these categories were scored 0, 10, 33, 50, 67, 90, 100, respectively.

<u>Teacher questionnaire</u>

Discourse coherence:

> "About how often do students in your class write about (or in response to) things they have read?"
> "About how often do you discuss writing topics with your students before asking them to write?"
> "About how often do you and your class discuss the readings you assign?"
> "When you ask students about their reading assignments in class, how frequently do you attempt to do each of the following?
> > "Ask them to relate what they have read to their other readings"
> "About how often does your class relate its discussion to previous discussions you have had?"
> "About how often do you and your class discuss things students have written about?"

> Response categories: Never, less than once a month, once a month, two to three times a month, once a week, more than once a week, every day. Responses in these categories were scored in a monthly scale of 0, .5, 1, 2.5, 4, 10, 20. Then they were summed across items, and divided by 4 to convert to a times-per-week scale.

Table II. Means and Standard Deviations of Variables.[a]

Variable	Total	Honors	Regular	Remedial	Source of Data
Dependent Variable					
Literature Achievement	15.822 (6.776)	19.774 (6.110)	15.625 (5.933)	9.838 (5.268)	Researcher-administered
Background Variables					
Sex (female = 1)	0.504 (0.500)	0.512 (0.500)	0.503 (0.500)	0.491 (0.501)	Student questionnaire
Minority (black or Hispanic = 1)	0.187 (0.390)	0.096 (0.295)	0.154 (0.361)	0.426 (0.495)	Student questionnaire
SES	0.001 (0.815)	0.372 (0.738)	-0.071 (0.768)	-0.415 (0.808)	Student questionnaire
Fall Reading Score	27.418 (7.630)	29.913 (7.185)	28.594 (6.430)	20.097 (6.906)	Researcher-administered
Fall Writing Score	5.995 (1.390)	6.767 (1.366)	5.714 (1.247)	5.488 (1.274)	Researcher-administered
Standardized Math Score[b]	64.237 (19.240)	80.412 (14.462)	60.828 (16.074)	46.847 (13.258)	School records
Standardized Reading Score[b]	62.110 (17.771)	77.662 (13.344)	59.167 (13.559)	44.476 (13.032)	School records
Instructional Variables					
Percent of Reading Completed	83.700 (23.612)	86.971 (19.433)	83.101 (24.704)	79.938 (26.104)	Student questionnaire
Percent of Writing Completed	87.387 (20.364)	92.112 (15.904)	86.475 (20.753)	82.076 (23.948)	Student questionnaire
Percent Offtask	4.202 (5.666)	2.220 (2.673)	3.794 (3.854)	8.583 (9.718)	Classroom observation
Percent of Authentic Questions	20.554 (16.900)	16.991 (12.125)	24.234 (18.849)	16.404 (15.839)	Classroom observation
Percent of Questions with Uptake	19.967 (11.566)	20.553 (8.440)	21.656 (12.399)	14.396 (11.999)	Classroom observation
Minutes of Discussion Time	0.563 (1.409)	1.163 (2.041)	0.167 (0.323)	0.653 (1.582)	Classroom observation
Discourse Coherence[c]	11.970 (6.514)	11.016 (6.570)	13.109 (6.464)	10.436 (5.989)	Teacher questionnaire
Number of Students	1564	480	793	291	

[a] Standard deviations are in parentheses.
[b] Normal curve equivalent of national percentile.
[c] In scale of times per week (see appendix for questionnaire items).

Table 2. Class-level means of instructional conditions.

INSTRUCTIONAL VARIABLE	CLASS TYPE		
	Honors	Regular	Remedial
Participation			
Percent of reading completed*	87.791	81.986	80.417
Percent of writing completed*	91.306	84.657	82.546
Offtask in class*	2.043	4.079	6.840
Discourse			
Percent authentic teacher questions	16.635	22.943	16.975
Percent of questions with uptake	19.409	21.315	17.059
Minutes of discussion per day*	1.224	.200	.643
Coherence of instruction	10.865	13.351	10.367
Number of classes	24	44	24

*F-test for differences between class types is significant at p ≤ .05.

Table 3. Maximum likelihood estimates of background effects on literature achievement in eighth- and ninth-grade ability-grouped English classes. N=1564 students.

INDEPENDENT VARIABLES	Effect	Standard Error
Background		
Sex (female = 1)	1.051**	.263
Minority (black or Hispanic=1)	-1.292**	.347
SES	.661**	.179
Fall reading score	.292***	.020
Fall writing score	.629***	.108
Ability	.101***	.018
Intercepts		
Honors classes	-1.707	1.337
Regular classes	-2.550*	1.081
Remedial classes	-3.697**	.930

Chi-square equals 55.89 with 32 degrees of freedom.

* coefficient is twice its standard error
** coefficient is three times its standard error
*** coefficient is four times its standard error

Table 4. Alternative models of the effects of background and instruction on achievement.

	MODEL	Chi-square	Degrees of freedom	COMPARISON TO PREVIOUS MODEL		
				Chi-square difference	Difference in degrees of freedom	P
(1)	Same effects of instruction in each ability group	123.69	67			
(2)	Varied effects of offtask, authenticity, and discussion	86.33	61	37.36	6	< .01
(3)	Model (2) plus varied effects of uptake	86.19	59	0.14	2	> .50
(4)	Model (3) plus varied effects of writing completed	85.74	57	0.45	2	> .50
(5)	Model (4) plus varied effects of reading completed	85.65	55	1.09	2	> .50
(6)	Model (5) plus varied effects of coherence	85.17	53	0.48	2	> .50

Table 5. Maximum likelihood estimates of background and instructional effects on literature achievement in eighth- and ninth-grade ability-grouped English classes. N=11564 students.

INDEPENDENT VARIABLES	Effect	Standard Error
Background		
Sex (female = 1)	1.188***	.252
Minority (black or Hispanic=1)	-.652	.339
SES	.155	.174
Fall reading score	.202***	.024
Fall writing score	.512***	.103
Ability	.121***	.018
Instruct on		
Completion of reading	.022**	.006
Completion of writing	.025**	.007
Offtask in class		
Honors classes	.149	.092
Regular classes	-.193***	.044
Remedial classes	-.124***	.028
Authentic teacher questions		
Honors classes	.056*	.022
Regular classes	.003	.010
Remedial classes	-.050*	.017
Uptake	.063***	.013
Discussion		
Honors classes	.277*	.129
Regular classes	-1.510*	.591
Remedial classes	.045	.169
Discourse coherence	.158***	.022
Intercepts		
Honors classes	-8.502***	1.385
Regular classes	-7.081***	1.207
Remedial classes	-7.144***	1.061

Chi-square equals 86.33 with 61 degrees of freedom.

* coefficient is twice its standard error
** coefficient is three times its standard error
*** coefficient is four times its standard error

Middle School Ability Grouping and Student Achievement in Science and Mathematics

Thomas B. Hoffer
Northern Illinois University

This paper analyzes the effects of middle school ability grouping on cognitive achievement in mathematics and science. In contrast to most previous research on tracking, this analysis compares outcomes in grouped and nongrouped schools. The hypotheses tested here are, first, that ability grouping raises the aggregate level of student achievement and, second, that ability grouping achieves this end by increasing the learning of all students. Comparing average student achievement growth from the seventh to the ninth grades in grouped and nongrouped schools shows that overall gains from ability grouping in either subject are negligible, controlling for differences in student social background and initial levels of achievement. Comparing the achievement growth of nongrouped students and high- and low-group students shows that high-group placement generally has a weak positive effect while low-group placement has a stronger negative effect. Ability grouping thus appears to benefit advanced students, to harm slower students, and to have a negligible overall effect as the benefits and liabilities cancel each other out.

Why do students differ in their learning of basic skills and knowledge? What might schools and school systems do to raise achievement levels and reduce learning inequalities? Initial efforts to answer these questions emphasized the importance of student social backgrounds and school social compositions (Coleman, et al., 1966; Jencks, et al., 1972). Subsequent efforts to identify the mechanisms of these general factors emphasized the importance of socialization differences among students, particularly the encouragement for academic success that students receive from parents, teachers, and peers (Jencks, Crouse, & Mueser, 1983; Sewell, Haller, & Portes, 1969), and school-based opportunity differences (Kerckhoff, 1976). The most important school organization variable has proved to be student ability group or curriculum-program assignment. A number of careful studies have shown that student placement exerts a powerful independent effect on achievement and attainments and explains part of the total effects of

social class on these outcomes (Alexander & McDill, 1976; Gamoran, 1987; Gamoran & Mare, 1989; Hauser, Sewell, & Alwin, 1976; Heyns, 1974; Vanfossen, Jones, & Spade, 1987).

An important practical conclusion of the research on ability grouping and high school tracking is that these practices constitute significant barriers to the goals of greater equality of opportunity and outcomes and to higher average academic achievement (Oakes, 1985, 1990). This conclusion and the research from which it is derived have recently been called sharply into question, however. Slavin (1990) concluded from his best evidence review of some 25 studies that ability grouping has no significant overall effects on secondary school achievement. Further, Slavin found that the effects of grouping are negligible for students at all levels of prior achievement. If Slavin is correct, then the considerable body of sociological research on the effects of tracking on cognitive achievement is simply wrong and is properly discarded.

Natriello, G., Pallas, A. N., & Alexander, K. L. (1989). On the right track? Curriculum and academic achievement. *Sociology of Education, 62,* 109-118.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement, *Research in the teaching of English, 25, 261-290.*

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven, CT: Yale University Press.

Oakes, J., Gamoran, A., & Page, R. N (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P. W. Jackson (Ed), *Handbook of research on curriculum* (pp. 570-608). New York: Macmillan.

Parsons, T. (1959). The school class as a social system: Some of its functions in American society. *Harvard Education Review, 29,* 297-313.

Rehberg, R. A.,& Rosenthal, E. R (1978) *Class and merit in the American high school.* New York: Longman.

Rohlen, T. P (1983). *Japan's high schools.* Berkeley, CA: University of California Press.

Rosenbaum, J. E. (1976). *Making inequality.* New York: Wiley

Sorensen, A. B. (1987). The organizational differentiation of students in school as an opportunity structure. In M T. Hallinan (Ed), *The social organization of schools: New conceptualizations of the learning process* (pp. 103-129). New York: Plenum

Useem, E. (1991). Student selection into course sequences in mathematics: The impact of parental involvement and school policies. *Journal of Research on Adolescence, 1,* 231 250.

Yogev, A. (1981). Determinants of early educational career in Israel: Further evidence for the sponsorship thesis. *Sociology of Education, 54,* 181-195.

Author

ADAM GAMORAN, Professor of Sociology and Educational Policy Studies, Department of Sociology, University of Wisconsin Madison, 1180 Observatory Drive, Madison, WI 53706. *Specializations:* stratification and resource allocation in school systems.

Synthesis of Research

Is Ability Grouping Equitable?

Adam Gamoran

Grouping and tracking do not increase overall achievement in schools, but they do promote inequity, research suggests. To reduce inequality, we should decrease the use of both practices, and, where ability grouping is retained, improve its use.

Ability grouping is one of the most common responses to the problem of providing for student differences. but is it an *equitable* response? Few questions about education have evoked more controversy.

Grouping has different effects in different circumstances. As currently practiced, it typically leads to *inequitable* outcomes. To place the debate in its proper perspective, we must remember that decisions about grouping are preliminary and that what matters most comes next: decisions about what to do with students *after* they've been assigned to classes. Given poor instruction, neither heterogeneous nor homogeneous grouping can be effective; with excellent instruction, either may succeed.

Drawing on the best research we have on grouping, I want to describe conditions that make one system or the other more likely to result in high achievement that is equitably distributed. Then I'll look at the challenges educators face depending on which approach to grouping they take. But, first, let's clarify two terms.

Tracking vs. Grouping

"Curriculum tracking" and "ability grouping" are sometimes used interchangeably. I use "tracking" to mean broad, programmatic divisions that separate students for all academic subjects. For example, high school tracks divide students into academic, general, and vocational programs.

Elementary schools "track" students when they divide them into separate classes for the entire day.

I use "ability grouping" to refer to divisions among students for particular subjects, such as special class assignments for math or within-class groups for reading. "Ability," strictly speaking, however, is not usually the criterion for grouping. Rather, students are typically divided according to measured or perceived performance in school. Because school performance is related to social inequality outside the school, such divisions contribute to the separation of students from different racial, ethnic, and social backgrounds (Oakes et al. 1992).

Achievement Effects of Grouping and Tracking

To consider the effects of ability grouping, we need to keep two questions in mind. First, how does grouping affect the overall *level* of achievement in the school? This is a question about "productivity." Would the school produce higher achievement if ability grouping were eliminated?

Second, how does grouping affect the *distribution* of achievement in the school? This is a question about "inequality." Would achievement be more equally distributed in the absence of ability grouping? In the past, advocates of grouping have tended to focus on the first question, and critics have emphasized the second. To engage in a balanced discussion, we must examine both.

Grouping and productivity. Little evidence supports the claim that tracking or grouping by ability produces higher overall achievement than heterogeneous grouping. At the

elementary level, most grouping systems fail to raise achievement. Some forms of subject-specific grouping—particularly within-class grouping for math and cross-grade grouping for reading—tend to have positive effects on overall achievement (Slavin 1987). The issue has received less attention at the secondary level, probably because almost all American secondary schools have some degree of tracking (Oakes 1985).

In a well-designed British study, Fogelman (1983) and Kerckhoff (1986) followed more than 9,000 students in grouped and ungrouped secondary schools for a five-year period, finding little difference in average scores on standardized tests of math and reading achievement.[1] The absence of overall differences between types of schools, however, masked important differences that occurred *within* the grouped schools.

Grouping and inequality. In the British study, there were no average differences between grouped and ungrouped schools because within the grouped schools, high-group students performed better than similar students in ungrouped schools, but low-group students did worse. Students in remedial classes performed especially poorly compared to ungrouped

students with similar family backgrounds and initial achievement. With low-group losses offsetting high-group gains, the effects on productivity were about zero, but the impact on inequality was substantial.

In the United States, high school tracking results in similar increases in inequality. In a national survey that followed more than 20,000 students from grades 10–12, academic track students gained significantly more on tests of math, science, reading, vocabulary, writing, and civics, compared to similar students in general and vocational tracks (Gamoran 1987).[2] In fact, achievement gaps between students in different tracks widened more than the overall disparity between students who dropped out of school after 10th grade and those who stayed in school. This means that which program a student pursued in high school mattered more for achievement than whether or not he or she was in school. Unfortunately, studies like this one do not show whether increasing inequality occurred in the context of rising or falling achievement for the school as a whole, because tracked and untracked schools were not compared.

> Students at all abilities—not just high achievers—benefit from collaborative projects.

Elementary school studies also show increasing inequality over time (Weinstein 1976, Hallinan and Sorensen 1983, Gamoran 1986). Even when overall achievement rises, inequality may grow because high-group students often gain more than students in low-ability groups (Oakes et al, 1992).

Slavin's "best evidence syntheses." Perhaps the most comprehensive and careful reviews of research on ability grouping are Robert Slavin's reports of grouping and achievement in elementary (1987) and secondary (1990) schools. Other than the elementary school exceptions noted above, Slavin argued that ability grouping has no effects on either productivity or inequality: grouped and ungrouped schools produce about the same level of achievement, and neither high, nor low, nor average groups obtain any special benefit or suffer a particular loss due to grouping. Slavin reached these conclusions after examining a diverse array of studies conducted over a 60-year period. Some of the studies showed positive effects; others yielded negative results, for productivity and inequality, as a result of ability grouping. Because the results averaged out to about zero, Slavin concluded that ability grouping has no effects and that the effects that appeared in many studies resulted from random or systematic errors of measurement (Slavin 1990).

I think another interpretation is more likely: the diversity of results does not mean the true effects are zero but, rather, that ability grouping has different effects depending on where and how it is implemented. The studies Slavin reviewed provided almost no information on what

Which program a student pursued in high school mattered more for achievement than whether or not he or she was in school!

occurred inside the classrooms after students were assigned. In some studies, teachers may have provided exactly the same instruction to the grouped and ungrouped classes, and there would be little reason to expect achievement benefits *or* detriments to ability grouping. In other studies, teaching quality may have favored one group or the other, leading to outcomes that differed by group. Slavin's ultimate conclusion echoes a finding that is more than half a century old: ability grouping has no effects on achievement unless teachers use it to provide different instruction to different groups."

I conclude that grouping and tracking rarely add to overall achievement in a school, but they often contribute to inequality. This finding is most consistent for high school tracking, but it is not uncommon in other forms and at other levels. Typically, it means that high-track students are gaining and low-track students are falling farther behind. But the effects of ability grouping are not the same in every context, and we need to discover how they come about in order to improve productivity and reduce inequality.

Sources of Achievement Inequality

Why does tracking often benefit high achievers but not their counterparts in other groups? Most research on grouping and achievement has failed

to consider how students were treated after they were assigned to their classes. Fortunately, a number of case studies and a few surveys provide information on what goes on in different groups and tracks. These reports suggest that the quality of instruction and the climate for learning favors high-level groups and honors classes over low groups and remedial classes.

Unequal instruction. At the elementary level, several researchers have documented fast-paced reading instruction in high-level groups and slow-moving progress in low groups. This occurs for both within-class and between-class grouping (Barr and Dreeben 1983, Gamoran 1986, Rowan and Miracle 1983). From these studies, one cannot tell whether slower instruction in low groups meets the needs of these students or unnecessarily holds them back. When middle- and low-group students of similar prior achievement are compared, middle-group students gain more, suggesting that slow-paced instruction contributes to the low-group deficit. This interpretation is bolstered by a recent survey of elementary school mathematics classes, in which middle- and low-group students were significantly more likely than high-group students to say their class was too easy (Coley et. al. 1992). Other researchers indicate that low reading groups offer a less conducive learning environment, with more interruptions than middle and high groups (Allington 1980, Eder 1981)"

Differences in context and climate have also been described at the secondary level. First, college-track students take more academic courses than students in other tracks, contributing to their achievement

Ending Ability Grouping Is a Moral Imperative

Cloyd Hastings

C utting in-depth educational research, such well-intended writers as Robert Slavin and Jeannne Oakes have attacked ability grouping. Their reliance upon quantitative methodologies does not sufficiently distinguish them from the supporters of homogeneous grouping. Both share a common belief in the power to persuade and influence others through statistical data. This common dependence upon numerical data is the cause for a continuing battle. It blinds the world to a different paradigm.

The answer to the debate on ability grouping is not to be found in new research. There exists a body of philosophic absolutes that should include this statement: The ability grouping of students for educational opportunities in a democratic society is ethically unacceptable.

We need not justify this with research, for it is a statement of principle, not of science. It should become a moral imperative along with the beliefs that slavery is immoral and that all people are

created equal under the law.

Our individualism is a defining element of our membership in society; it should not exclude us. We must accept and celebrate diversity because we are all different. We must believe in the fundamental worth and dignity of each person.

The individual is fundamental to democracy and most religions. The individual should be fundamental to all educational decisions. Because much of our thinking about mass education practices is derived from factory model thinking, commitment to the individual will be more difficult to implement in public education. We now have, however, 100 years of knowledge and technology that was unavailable to the developers of mass education, and we have new models.

For example, a bicycle company in Japan is filling orders for individualized bikes. In a nation that has established itself as a champion of mass production techniques, the Japanese have discovered a way to customize production on a mass

level. This is the challenge facing American education. How do we customize educational opportunities and experiences on a mass level?

The bicycle company starts with what is common, and defining, about the product and then incorporates what the customer believes is necessary to fulfill the concept of a bike. In education, we must start with what all learners need and then customize based upon the individual.

We need to stop standardizing expectations based upon aggregated data and begin to customize based upon disaggregated knowledge of the individual. Standardized testing used for sorting, categorizing, and labeling must be ended. Accountability in terms of student progress can, and must, be maintained on an individual continuum and not on a group continuum. Difficult? Yes! Challenging? Yes! Impossible? No! ◼

Cloyd Hastings is Principal, McCoy Elementary School, 2425 McCoy Rd., Carrollton, TX 75006.

advantage (Gamoran 1987). Second, observers report that high-track teachers are more enthusiastic and spend more time preparing (Rosenbaum 1976, Oakes 1991). Teachers may compete for the opportunity to teach honors and accelerated classes, and those with more experience or better reputations tend to win the privilege (Finley 1984, Oakes 1991). Although problem solving and critical thinking are not especially common, they are more likely to occur in high

tracks than low tracks (Oakes 1985, Gamoran and Nystrand 1990). In contrast, low-track instruction tends to be fragmented, emphasizing worksheets and recitation (Page 1992). Teachers in low-track classes spend more time on behavior management and less time on instruction (Oakes 1985).

Unequal behavior and attitudes among students. These differences cannot be ascribed solely to teachers, however, because *students' responses*

to instruction also differ across tracks and ability groups. Low-track students are off-task more often, spend less time on homework, and turn in fewer assignments (Oakes 1985, Gamoran and Nystrand 1990). Current data do not indicate whether low-track students respond less well because instruction is less engaging or whether instruction is less engaging because students are not responsive. Both processes are probably at work. Case study writers have long

contended that tracking polarizes the student body into "pro-school" and "anti-school" groups (for example, Lacey 1970, Abraham 1989). The latest survey research supports this claim: Berends (1991) found that college- and noncollege-track students differ more over time in the extent of disciplinary problems, in engagement with schoolwork, and in expectations for future schooling.

What Can Be Done?

Although the research is not definitive, it does suggest two actions: reduce the use of tracking and grouping and improve the way ability grouping is used where it is retained.

Reduce the use of tracking and grouping. Generally, the more rigid the tracking system, the more research studies have found no benefits to overall school achievement and serious detriments to equity. Students who report being assigned to different tracks in high school become more unequal in their achievement over time, and the increase in inequality is greatest in schools where students rarely change tracks (Gamoran 1992). In elementary schools, between-class grouping for the entire school day is least likely to show any benefits (Slavin 1987). As Slavin (1987) explains, rigid tracking systems are likely to fail because when a single division by ability is made for all subjects, classes remain heterogeneous on most skills, so there is no improvement in the fit between students' needs and the provision of instruction. In addition, rigid tracking systems may be more likely to induce polarized attitudes toward schooling (Gamoran 1992). In moving to reduce

Observers report that high-track teachers are more enthusiastic and spend more time preparing lessons. In contrast, teachers in low-track classes spend more time on behavior management and less on instruction.

the use of grouping, then, the first step should be to eliminate the most rigid forms of tracking, such as broad, inflexible program assignment in high schools and between-class tracking for the whole day in elementary schools.

Efforts to reduce tracking must grapple with the fact that in at least some cases, high-track students perform better than similar students in heterogeneous classes. The elimination of grouping must be accompanied by staff development opportunities for teachers to learn strategies for enhancing the learning of all students in classes that are more diverse than those to which they are accustomed. At the same time, those who strive to maintain ability grouping out of concern for high-track students must come to grips with the growth in inequality that occurs in many cases.

Improve the use of ability grouping. To the extent that grouping is not

completely eliminated, it must be implemented more effectively than is typical. First, it is essential to avoid locking in teachers and students to their track assignments. Permanent assignments result in a vicious cycle, in which the expectations of teachers and students enter a downward spiral (Page 1992). Schools must make at least two sorts of investments to bring greater flexibility to their grouping systems: (1) they must reassess students' capabilities and take new information into account when making assignment decisions, and (2) they must enable students to make up curricular material they may have missed—for example, in tutorials during the school year or the summer—so that those who are ready to advance are not held back by lack of curriculum coverage. The latter requires investment not just by schools, but by students as well, who must undertake extra work to catch up. Implementing more flexible grouping systems also means rotating teachers so that all students have opportunities to learn from the most effective teachers and to prevent the loss of morale that sometimes occurs for teachers who are assigned to low tracks year after year.

Second, those who use ability grouping must improve instruction in low groups. This could, at the same time, reduce the inequality that often results from grouping and raise the overall level of achievement in the school. This recommendation is extremely difficult to follow—indeed, were it not so difficult, ability grouping would be a lot less controversial. It is difficult because (1) by virtue of their assignment, teachers

and students in low tracks have low expectations for academic work; and (2) low-track students often resist challenging academic work. One observer found that low-track students preferred worksheets to discussion, because the seatwork kept private what students did and did not know (Metz 1978).

Is it even possible? Can high-quality instruction ever take place in low-status groups? We have many more examples of unsuccessful low-track classes than successful ones, but there are some circumstances under which low-group students receive effective instruction. At the elementary level, grouping systems that divide students on the basis of skills closely related to the curriculum and those that adjust curriculum and instruction to address students' needs are more likely to be effective. This conclusion is based on studies of within-class grouping for math and cross-grade, subject-specific grouping for reading (Slavin 1987), but the conclusion is probably generally valid.

At the secondary level, a few case studies suggest that low-track classes may serve their remedial purpose—that is, they allow students to catch up, or at least prevent them from falling further behind—under the following conditions:

- Teachers hold high expectations, manifested by their emphasis on academic work.
- Teachers exert extra effort, compared to their efforts in other classes.
- Teachers and students have opportunities for extensive oral interaction.
- There is no procedure in place that assigns weak or less experienced teachers to the lower track (Page and Valli 1990, Gamoran 1991).

Current data do not indicate whether low-track students respond less well because instruction is less engaging or whether instruction is less engaging because students are not responsive. Both processes are probably at work.

These case studies rely on private schools mostly with middle-class students, and we have as yet no evidence that they generalize well to other situations.

One 9th grade English teacher I observed, whose low-group students kept pace with their peers in other classes, told her students: "I know it's not easy, you guys—I know it's not easy—but we're not going to read *Weekly Reader* in this class. All right? You deserve to have this information, so stick with it." With such a persistent teacher, and equally persistent students, low-track classes may be effective, but the phenomenon is too rare for one to have confidence that it will become the general case anytime soon. All the more reason to curtail tracking and grouping where possible. ■

The British study is remarkable in its comprehensiveness: it began with nearly every child born in England, Scotland, and Wales during the first week of March 1958 and followed them from birth to age 23. The ability-grouping analyses covered the period from age 11 to 16. The study is also

especially valuable because it includes a large number of comparable schools that used and did not use tracking, or "streaming" as it is called in Britain. In the United States, it is impossible to find a representative sample of secondary schools in which students are not grouped in math and English.

These differential gains occurred for students who were statistically equated in prior achievement and background characteristics. In general, students in the different tracks are far from equal in these areas, so the gross differences between tracks were much larger.

'Slavin has stated: "For ability grouping to be effective at the elementary level, it must create true homogeneity on the specific skill being taught, and instruction must be closely tailored to students' levels of performance" (1987, p. 323). For the secondary level, he remarked: "The lesson to be drawn from research on ability grouping may be that unless teaching methods are systematically changed, school organization has little impact on student achievement" (1990, p. 491). Compare these to what Ethel L. Cornell concluded in 1936: "The results of ability grouping seem to depend less upon the fact of grouping itself than upon ... the differentiations in [curriculum] content, method, and speed, and the technique of the teacher" (p. 304).

References

Abraham, J. (1989). "Testing Hargreaves' and Lacey's Differentiation-Polarization Theory in a Setted Comprehensive." *British Journal of Sociology* 40: 46-81.

Allington, R. L. (1980). "Teacher Interruption Behaviors During Primary-Grade Oral Reading." *Journal of Educational Psychology* 72: 371-374.

Barr, R., and R. Dreeben. (1983). *How Schools Work.* Chicago: University of Chicago Press.

Berends, M. (1991). *High School Tracking and Students' School Orientations.* Madison, Wis.: National Center on Effective Secondary Schools.

Coley, R. J., R. Ekstrom, J. Gantt, A. M. Villegas, R. Mitchell, and S. M. Watts.

Highlights of Research on
Ability Grouping and Achievement

■ Ability grouping rarely benefits overall achievement, but it can contribute to inequality of achievement, as students in high groups gain and low-group students fall farther behind. The more rigid the tracking system, the more likely these patterns are to emerge.

■ When students are grouped according to skills that are closely related to the curriculum, and when curriculum and instruction are tailored to students' capacities, ability grouping may raise achievement. Research at the elementary level supports this claim more so than at the secondary level, where there are few examples of effective instruction in low-ability classes.

■ The use of ability grouping should be curtailed, starting with its most rigid forms: permanent program assignments in high schools and between-class grouping for the whole school day in elementary schools.

■ Where grouping is not eliminated, its implementation must be improved: neither teachers nor students should be locked into their assignments, and the quality of instruction in low groups must be raised.

(1992). *On the Right Track*. Princeton, N.J.: Educational Testing Service.

Cornell, E. L. (1936). "Effects of Ability Grouping Determinable from Published Studies." In *The Grouping of Pupils*. Yearbook of the National Society for the Study of Education, vol. 35, part I, edited by G. M. Whipple. Bloomington, Ill.: Public School Riblishing Co.

Eder, D. (1981). "Ability Grouping as a Self-Fulfilling Prophecy: A Microanalysis of Teacher-Student Interaction." *Sociology of Education* 54: 151-161.

Finley, M. K. (1984). "Teachers and Tracking in a Comprehensive High School." *Sociology of Education* 57: 233-243.

Fogelman, K. (1983). "Ability Grouping in the Secondary School." In *Growing Up in Great Britain: Papers from the National Child Development Study*, edited by K. Fogelman. London: Macmillan.

Gamoran, A. (1986). "Instructional and Institutional Effects of Ability Grouping." *Sociology of Education* 59: 185-198.

Gamoran, A. (1987). "The Stratification of High School Learning Opportunities." *Sociology of Education* 60: 135-155.

Gamoran, A. (1991). *Alternative Uses of Ability Grouping: Can We Bring High-Quality Instruction to Low-Ability Classes?* Madison, Wis.: Center on Organization and Restructuring of Schools.

Gamoran, A. (1992). *The Variable Effects of High School Tracking*. Madison, Wis.: Center on Organization and Restructuring of Schools.

Gamoran, A., and M. Nystrand. (1990). "Tracking, Instruction, and Achievement." Paper presented at the World Congress of Sociology, Madrid.

Hallinan, M. T., and A. B. Sorensen. (1983). "The Formation and Stability of Instructional Groups." *American Sociological Review* 48: 838-851.

Kerckhoff, A. C. (1986). "Effects of Ability Grouping in British Secondary Schools." *American Sociological Review* 51: 842-858.

Lacey, C. (1970). *Hightown Grammar*. Manchester: Manchester University Press.

Metz, M. H. (1978). *Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools*. Berkeley: University of California Press.

Oakes, J. (1985). *Keeping Track: How Schools Structure Inequality*. New Haven, Conn.: Yale University Press.

Oakes, J. (1991). *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science*. Santa Monica, Calif.: RAND.

Oakes, J., A. Gamoran, and R. N. Page. (1992). "Curriculum Differentiation: Opportunities, Outcomes, and Mean-

ings." In *Handbook of Research on Curriculum*, edited by P. W. Jackson. Washington, D.C.: American Educational Research Association.

Page, R. N. (1992). *Lower Track Classrooms: A Curricular and Cultural Perspective*. New York: Teachers College Press.

Page, R. N., and L. Valli, editors. (1990). *Curriculum Differentiation: Interpretive Studies in U.S. Secondary Schools*. Albany, N.Y.: SUNY Press.

Rosenbaum, J. E. (1976). *Making Inequality: The Hidden Curriculum of High School Tracking*. New York: Wiley.

Rowan, B., and A. W. Miracle, Jr. (1983). "Systems of Ability Grouping and the Stratification of Achievement in Elementary Schools." *Sociology of Education* 56: 133-144.

Slavin, R. E. (1987). "Ability Grouping and Achievement in Elementary Schools: A Best-Evidence Synthesis." *Review of Educational Research* 57: 293-336.

Slavin, R. E. (1990). "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." *Review of Educational Research* 60: 471-499.

Weinstein, R. S. (1976). "Reading Group Membership in First Grade: Teacher Behavior and Student Experiences Over Time." *Journal of Educational Psychology* 68: 103-116.

Adam Gamoran is Professor of Sociology and Educational Policy Studies, University of Wisconsin-Madison, Center on Organization and Restructuring of Schools, 1025 W. Johnson St., Madison, WI 53706.

kinko's
The new way to office.

2741 University Ave. Madison WI. • Telephone: (608) 238-2741 • Fax; (608) 238-5392

נכנס | בקבל
ת. תוש היב .4 | המתקבל:
הכניס .ב | .2 VA..
.5 | 4.
.4 לימוד |

fax cover sheet

Date __May 9 1997__   Number of pages __R5__ (including cover page)

to: Name __Haidar Abyris__   from: Name __Adam Gamoran__

Company __Mundeh Institute__   Company __U of Wisconsin__

Telephone _____   Telephone __(608) 263-4253__

Fax __011 - 972 - 2 - 566 - 2887__   Fax __(608) 265-5389__

Comments Haddar,

   This fax contains two documents:

① Evaluation of TEI (18 pages)

② the Leading Bud. caters Project (6 pages)

Please confirm that the fax came through cleanly. Thanks,
                                    Adam

More than 800 locations worldwide. For the location nearest you, call 1-800-2-KINKOS.

EVALUATION OF THE CIJE TEACHER EDUCATOR INSTITUTE (TEI)

May 1997

Adam Gamoran, Ellen Goldring, and Bill Robinson
Council for Initiatives in Jewish Education

CONTENTS

1. Overview. (Reprinted from *Pathways: A Guide to Evaluating Programs in Jewish Settings,* by Adrienne Bank.)

2. Instrument: Professional Development Survey

3. Instrument: Interview of TEI Participants

Evaluating Professional Development Programs:
The Teacher-Educator Institute of the Council for Initiatives in Jewish Education (CIJE)

[The authors of this section, who designed the evaluation described in the example, are Adam Gamoran, Ellen Goldring, and Bill Robinson.]

Program Description

As a catalyst for systemic change in Jewish education, CIJE's mission includes the transformation of "the supplementary school into an institution where exciting learning takes place, where students are stimulated by what they encounter, and where a love of Jewish learning and the commitment to Jewish living is the hallmark of the institution." To accomplish this task, supplementary schools must become places where "exciting, innovative teaching by knowledgeable and committed educators" takes place. While research undertaken by CIJE has shown that Jewish educators are committed to a career in Jewish education, the research also has highlighted the substantial deficiencies of Jewish educators in formal Judaic training. One way to address these issues is to transform the types of professional development programs being offered to Jewish educators in their schools and communities from one-shot workshops focused on giving educators new techniques to more extensive and content-rich professional development opportunities that increase the capacity of Jewish educators to learn in and from their practice.

To develop this option, the Teacher-Educator Institute (TEI) brings together teams of educational leaders from different communities and denominations to inquire through reflective practice into the nature of good Jewish teaching and good professional development. Teams are a central element of TEI's change strategy – facilitating the development of local cohorts of educators who have shared an intense learning experience, developed a shared vocabulary and mode of educational discourse, and wrestled with conception of good teaching and learning and professional development. TEI models the type of professional development opportunities that the educational leaders may offer to the teachers in their particular schools and communities through such activities as:

- Investigations of videotaped lessons
- Curricular investigations
- Investigations into actual teachers' practices

Participants attend six seminars lasting about four days each over the course of two years. Between seminars participants are asked to complete exercises, such as observation of teachers and design experiments. At the end of the two years, participants are expected to have developed:

- Improved understandings of teaching and learning, Jewish content, teachers as learners, and professional development
- A personal repertoire of strategies for designing, implementing, and assessing professional development opportunities
- The ability to articulate a vision of good Jewish teaching, images of worthwhile professional development, and the relationship between the two

Program Evaluation

The focus of the evaluation is to examine:

- Changes in the conceptions and practices of the TEI participants
- Changes in the culture of the schools in which TEI participants will be working
- Changes in the formal professional development programs offered to supplementary school educators in their community and schools

We determined that if TEI was to produce change it was most likely to be observable in these three areas.

Formal Professional Development Programs

To evaluate community wide changes in professional development programs an operationalized set of ideal characteristics of professional development programs was developed:

- **Content**: the program is designed to contribute to the Judaic content knowledge of the participating educators
- **Audience**: the program is designed for a specific group of educators
- **Sessions**: the program is a series of sessions designed to address a coherent theme
- **Groups**: the program requires educators to attend as school teams
- **Practice**: the program is designed to help educators reflect on and apply learning to their practice
- **Plan**: the program is part of a comprehensive plan, sustained over time, for the ongoing professional development of the educators
- **Incentives**: incentives are provided to encourage the participation of educators in the program
- **Evaluation**: the program contains a worthwhile evaluation process

A survey of the professional development programs offered by the central agency and supplementary schools was administered in five of the communities that sent teams to TEI. The data yielded a base-line map against which change can be measured when the survey is re-administered in a few years. In addition, the findings from this survey have been shared with the TEI faculty, who in turn have used the data in community presentations designed to mobilize the lay and professionals in the participating communities to change their current professional

development offerings to be more in accordance with the articulated characteristics of good professional development.

Conceptions and Practices of the TEI Participants

To evaluate changes in the conceptions and practices of the participants, interview protocols were designed in consultation with the TEI faculty and an outside expert in professional development. The first interview protocol, administered to participants from the same five communities prior to their participation in TEI, focused on:

- Their past work as teacher-educators
- Their current relations with professional colleagues
- Their prior learning experiences
- Their images of good professional development

The findings provided both a base-line picture of the participants and insights into the nature of the environment in which they are expected to create change. A second (interim) interview protocol, administered during their participation in the TEI seminars, focused on the perceived significance of TEI to the participants and the influence it has had on their work with Jewish educators. In addition, it also probes for changes in their conceptions of good professional development and perception of their own educational needs. The findings from these interviews will help the TEI faculty understand the impact of TEI and reveal any unexpected aspects of participation. Timely reporting of base-line and interim findings will allow the faculty to adjust elements of the program to have a better chance of reaching the intended goals. In addition, the interim interviews allow the evaluation team to begin to understand the processes and conditions through which change will or will not occur in the participating communities and schools. Later interview protocols will be developed based on the findings from these initial interviews and the survey, as well as any changes in the program.

Culture of the Schools

To evaluate changes in the culture of schools and to continue monitoring changes in the practices of the TEI participants, case studies of the participants will be conducted following their completion of the TEI seminars. Given limited resources, we decided to conduct the case studies in only two of the participating communities. This combination of limited case studies and more widespread interviews and surveys was the result of a conscious decision to find an optimum balance between getting fairly easily obtained data from a large number of participants and procuring potentially richer data on actual changes in participants' practices and their effects that are more difficult to obtain from a small number of participants. The case studies will involve observation of actual teacher-educator opportunities designed and implemented by the TEI participants and interviews with participating teachers about the significance and influence of these opportunities on their learning and teaching.

General Comments

The TEI evaluation is currently a work-in-progress. It is important to understand that not only should the program designers be responsive to interim findings of an evaluation, but the evaluation designers should be responsive to their own findings in (re)designing future elements of the evaluation. Both are iterative approaches. Nevertheless, one must not lose track of the initial purposes of any evaluation; otherwise, at the end you may be left wondering as to whether or not the program actually "worked." In the evaluation of TEI we have been and continue to be responsive to the results of our interim findings and to changes in the program itself; yet, we remain committed to evaluating the program against the goals of TEI that were articulated at the beginning. In designing and implementing evaluations of professional development programs, success often hinges upon maintaining the proper balance between formative and summative purposes, between the breadth and depth of data gathering activities, and between focusing on the initial goals of the program and ongoing changes in the program.

Council for Initiatives in Jewish Education

Teacher-Educator Institute

PROFESSIONAL DEVELOPMENT PROGRAM SURVEY

IMPORTANT

Please complete a Professional Development Program Survey for each in-service program that is offered by your institution. Answer all of the questions as completely as possible. If you have any difficulty in answering a particular question, explain why next to the question.

In addition, please complete a single questionnaire for your institution's staff meetings ONLY if the meetings contain an ongoing professional development component as part of an institutional staff development plan.

Please include a COURSE OUTLINE and EVALUATION INSTRUMENTS for the program.

Name of Program_____

Structure of Program:
(Check one)

☐ Single Session

☐ Course

☑ Study Group

Sponsoring Institution_____

Name and Title of Person Completing Survey _____

_____

I. PROGRAM AUDIENCE

The following questions ask you about the educators who attend the program.

1. The program participants work in the following ROLES;
   (Check all that apply)

   ☐ a. Teacher

   ☐ b. Teacher Aide

   ☐ c. Educational Director or Principal

   ☐ d. Assistant Educational Director or Principal

   ☐ e. Department Head (e.g., Hebrew department chair, director of primary program)

   ☐ f. Tutor

   ☐ g. Central Agency Staff

   ☐ h. Other (specify) _____

2. The program participants work in the following SETTINGS:
   (Check all that apply)

   ☐ a. Day School

   ☐ b. Supplementary School

   ☐ c. Pre-school

   ☐ d. Adult Education

   ☐ e. Central Agency

   ☐ f. Other (specify) _____

3. The program participants work in schools with the following AFFILIATIONS:
   (Check all that apply)

   ☐ a. Reform

   ☐ b. Conservative

   ☐ c. Traditional

   ☐ d. Orthodox

   ☐ e. Reconstructionist

   ☐ f. Community

   ☐ g. Jewish Community Center

   ☐ h. Other (specify) _____

4. The program participants work with the following POPULATIONS:
   (Check all that apply)

   ☐ a. Pre-school

   ☐ b. Kindergarten

   ☐ c. Elementary School (1 - 5)

   ☐ d. Middle School (6 -8)

   ☐ e. High School (9 - 12)

   ☐ f. Adults

   ☐ g. Other (specify) _____

5. The program participants have the following level(s) of EXPERIENCE:
   (Check all that apply)

   ☐ a. Novice in Jewish Education (5 years or less)

   ☐ b. Experienced in Jewish Education

   ☐ c. Other (specify)_____

   ☐ d. I don't know.

6. The program participants have the following type(s) of FORMAL TRAINING:
   (Check all that apply)

   ☐ a. No Formal Training

   ☐ b. Degree or Certification in Education

   ☐ c. Degree or Certification in Jewish Content

   ☐ d. Degree or Certification in Educational Administration/Leadership

   ☐ e. Other (specify) _____

   ☐ f. I don't know.

II. PROGRAM DESIGN

The following questions ask you about the subject matter, goals, and organization of the program. (Please include a course outline.)

7. Please describe the subject matter that is covered in the program (e.g., Hebrew language, life cycle, lesson planning, drama techniques)?

_____

_____

_____

_____

_____

_____

8. If the program were successful, how would participants be different as a result of their participation?

_____

_____

_____

_____

_____

9. What is (are) the primary format(s) of the program?
   (Check all that apply)

   ☐ a. Lecture

   ☐ b. Lecture-discussion

   ☐ c. Text-study Session

   ☐ d. Mentoring

   ☐ e. Coaching (working on specific skills)

   ☐ f. Videotape Analysis

   ☐ g. Curriculum Development or Analysis

   ☐ h. Classroom Experimentation and Reporting Back

   ☐ i. Investigating Problems of Practice

   ☐ j. Demonstration Teaching

   ☐ k. Other _____

10. Who are the faculty of the program?
    (Check all that apply)

    ☐ a. Teachers

    ☐ b. Principals or Educational Directors

    ☐ c. Central Agency Staff

    ☐ d. Rabbis

    ☐ e. Lay Leaders

    ☐ f. Outside Experts (specify)_____

    ☐ g. Other (specify)_____

11. Who designed the program?
    (Check all that apply)

- [ ] a. Teachers
- [ ] b. Principals or Educational Directors
- [ ] c. Central Agency Staff
- [ ] d. Rabbis
- [ ] e. Lay Leaders
- [ ] f. Outside Experts (specify) _____
- [ ] g. Other (specify) _____

12. Were the specific people who served as faculty also involved in designing the program?

Yes [1]   No [2]

III. PROGRAM PARTICIPANTS

The following questions ask you for additional information about the program participants.

13. Typically, how many people attend the program as participants? _____

14. Do participants attend as individuals, members of a school team, or along with their entire faculty?
    (Check all that apply)

- [ ] a. Individuals
- [ ] b. School Team without Principal
- [ ] c. School Team with Principal
- [ ] d. Entire Faculty
- [ ] e. Other (specify) _____

If you checked more than one response, please explain.

_____

_____

15. What incentives and/or supports are available to participants?
   (Check all that apply)

☐   a. None

☐   b. Stipend (How much?),_____

☐   c. Salary increase (How much?)_____

☐   d. Release Time   d. Release Time

☐   e. Academic Credits

☐   f. License or Certification

☐   g. CEU (Continuing Education Units)/
       SDU (Self Development Units)

☐   h. Trip to Israel

☐   i. Required by Contract

☐   j. Other (specify)_____

16. Are incentives provided directly to the school(s) for their educators' participation in the program?

Yes ☐   No ☐ 2

If Yes, please describe the incentive(s) and the criteria for awarding it.

_____

_____

IV. PROGRAM MEETINGS

The following questions ask you about the duration and intensity of the program, as well as the relationship between program meetings and other programs.

17. In total, how many meetings occur during the course of the program?_____

18. How often do the meetings occur? _____

19. On average, how many hours is each meeting of the program? _____

20. Over what period of time does the entire program run?_____

21. Is the content of one meeting related to the content of subsequent meetings?

Yes ☐ 1   NO ☒ 2

If Yes, please explain.

_____

_____

_____

_____

_____

_____

_____

22. Is there a relationship between the content of this program and any other program being offered in the community?

Yes ☐ 1   No ☐ 2

If Yes, please explain.

_____

_____

_____

_____

_____

_____

_____

V. PROGRAM EVALUATION

The following questions ask you about the evaluation of the program.
(Please include evaluation instruments.)

23. Is the program being evaluated?

Yes [1]   No [2]   (If No, you have completed this questionnaire.)

24. What is the focus of the evaluation?
(Check all that apply)

☐ a. Participants' Satisfaction

☐ b. Participants' Knowledge

☐ c. Participants' Attitudes

☐ d. Participants' Skills

☐ e. Students' Classroom Behaviors

☐ f. Students' Knowledge

☐ g. Students' Attitudes

☐ h. Other (specify)_____

How is this information collected (e.g., participant self-report, observations of teachers in their own classrooms)?

_____

_____

25. Who designed the evaluation?
(Check all that apply)

☐ a. Faculty

☐ b. Participants

☐ c. Outside Experts (specify) _____

☐ d. Other (specify)_____

Thank you very much for your cooperation!

CIJE Teacher-Educator Institute (TEI)
MEF Evaluation
INTERVIEW PROTOCOL:
EDUCATIONAL DIRECTORS
(COHORT #1)
(Revised May 1996)

The following are the questions and probes for the interviews with the supplementary school and pre-school educational directors in the SECOND COHORT.

A. BACKGROUND

I would like to begin our interview with some brief questions about your background.

1. How did you come to be in the position that you currently hold?
    [Probe: How long have you held your current position?]

2. For how many hours per week are you contracted to work in your current position?
        [Probes: do you currently hold any other positions in Jewish education? In general education? In other areas? For how many hours per week are you contracted to work in these other positions?]

3. What other positions have you held in Jewish education?
    [Probes: For how many years have you held a leadership position in Jewish education? In total, how long have you worked in Jewish education?]

4. What types of educational experiences have you had that have prepared you for your current position?
        [Probe for formal degrees and certification/licensure in Jewish studies, education, and administration/leadership. What kinds of formal and informal Jewish educational experiences did you receive as a young person?]

5. What have you been doing over the last two years to continue developing yourself as a Jewish educational leader?
        [Probe for formal and informal professional development experiences in Jewish studies, education and administration/leadership.]

B. PROFESSIONAL DEVELOPMENT PRACTICES

The next set of questions concerns the work you do in your position as [fill in].

1. Describe to me something that you did with teachers in the last five years about which you felt particularly good?
   [Probes; What did you do? What did the other teachers do? Why did you decide to do this? Why do you feel particularly good about this? How did the teachers respond? How would you do it differently now?]

2. Describe to me the types of interactions you have had with teachers during the past year.
   [Probe for both formal and informal interaction. Specifically, probe for interactions which support their teachers' work. In regard to this, probe for why they decided to do this, how they planned it, how the teachers responded and how would they do it differently. In addition, probe for differences among teachers and if their support addresses these differences.- Probe for learning of Judaic content.]

3. Describe to me the types of interactions you have had with other educational leaders during the past year.
   [Probe for what happens at Council meetings. Probe for informal interactions and other formal events, such as mentoring groups. Probe for both other educational directors and central agency staff. Probe; Who do you go to for support in your work?]

C. PARTICIPATION IN TEI AND ITS RELATION TO PRACTICE

Now let's talk about your participation in CIJE's Teacher=Educator Institute (TEI) and its relations to your current work.

1. Are there specific things about your experience in TEI that you have found to be particularly significant?
   [Probes: Why is this significant? Hove you encountered this before? Where? Has your experience in TEI changed the way you think about this? How?]

2. Are those playing a role in your current work? How?
   [Probe: How does this illustrate what you have experienced in TEI?]

ד)

3 Please describe in detail some professional development work you are doing, which has been influenced by your participation in TEI?

[Probe: Describe the typical encounter. What did you do? What did the participants do?]

The following questions focus on this particular work.

4. Why did you decide to do this?

[Probes: Did you want to learn something from doing this? What? Do you expect the others to be different as a result of their participation? how? What do you think this activity will lead to these changes?]

5. How did you plan this?

[Probe as to how they decided to do specific aspects: How did you decide to do this task or present this concept? With whom did you talk? When? Where? In what ways? What is your relationship to this person? Why this (these) person(s) and not others?]

6. How did the participants respond to this?

[Probes: Can you describe to me a particular incidence that makes you think this? Has their attitude toward the program changed over time? How do you know this?]

7. Would you do it differently now? In what ways?

[Probe for if they have learned anything from doing this? What? Probe for what they may have learned about the participants, the participants' learning, the subject matter, the context, and their own abilities.]

8. Do you consider this to be a good professional development experience? Why?

Now I'd like to return to you experiences in TEI.

9. How does TEI compare to other professional development experiences you have attended?

[Probes: In what ways is TEI more valuable for you personally and professionally? In what ways is it less valuable?]

10. What do you think you need to learn about professional development?

)

C. FUTURE PROFESSIONAL DEVELOPMENT PRACTICES

The final set of questions deal with your plans for the coming year and into the future.

1, What issues of professional development do you want to work on with your teachers during the coming year?
[Probe as to why s/he has decided to work on these particular issues]

2. Please describe how you plan to address these issues.
[Probes: What will you be doing? What will the participants be doing? Will other educators being helping you in your work? In what ways?]

3 In five years from now, what would you like to see happening in regard to the professional development off the educators in your community?
[Probe: Who are you referring to? What do you see as the obstacles to achieving this vision?]

a 5,/c ˙ y ·

facsimile
TRANSMITTAL

to: Annette Hochstem (Mandel Institute)

fax#:   011-972-2-566-2837

re: Paper on Gender Differences (per Adam's e-mail)

date: March 20, 1997

pages: 29, including cover sheet.

Article of his

Adam Gamoran et al

on gender differences

From the desk of...

Bill Robinson
Staff Researcher
CIJE
1525 Wood Creek Trail
Roswell, Georgia 30076

(770) 552-0930
Fax: (770) 998-0860

GENDER DIFFERENCES AMONG TEACHERS IN JEWISH SCHOOLS:
A STUDY OF THREE COMMUNITIES

ABSTRACT

Researchers have documented gender differences among teachers, including differences in salary levels, in secular, private, and Catholic high schools. While recent community studies have made data on teachers in Jewish schools available, gender differences have not been investigated. This paper reports findings from a survey of Jewish school teachers in three communities: Atlanta, Baltimore, and Milwaukee. The findings suggest important differences between women and men teachers. Women tend to enter Jewish education primarily in order to work with children, while men are more concerned with learning and teaching about Judaism. Similarly, while a substantially greater percentage of women have been trained in education as compared to Jewish studies, the reverse holds for men. While women and men have similar lengths of experience and both overwhelmingly plan to stay in Jewish education, a substantially greater percentage of men as compared to women see Jewish education as a career. Finally, women receive lower salaries and fewer are offered health benefits than men. Using a linear regression analysis, gender was shown to be a significant predictor of salary differences, even when controlling for various setting and personal characteristics.

what "% men" "significant"
16%

GENDER DIFFERENCES AMONG TEACHERS IN JEWISH SCHOOLS: A STUDY OF THREE COMMUNITIES

Bill Robinson
Adam Gamoran
Ellen Goldring

Council for Initiatives in Jewish Education

GENDER DIFFERENCES AMONG TEACHERS IN JEWISH SCHOOLS:

A STUDY OF THREE COMMUNITIES

BACKGROUND

In 1990, researchers at the University of Michigan found that women high school teachers nation-wide earned an average of $2,300 to $3,300 less than men who teach in high schools (Lee and Smith, 1990). The study used data from a sample of 8,894 teachers in 377 high schools compiled during the 1983-84 school year as part of the U.S. Education Department's ongoing *High School and Beyond* study. Even when controlling for educational backgrounds, experience, and differing wage levels across cities, the authors of this study found that women teachers in public, private, and Catholic high schools still earn less than men. These findings conform to a general pattern of gender-based salary differences in the workplace, which has been documented for decades.

While much research has been conducted on issues of gender equity within the classroom and its effects on students, gender differences among teachers in salary level and other important career-related issues has not received as widespread attention. A few studies (Lee, Smith, and Cioci, 1993; Huberman 1993; Kalaian and Freeman, 1994), have pointed toward specific gender differences in teachers' reasons for choosing an educational career, their orientation to pre-service training, their commitment to a career in education, and their perceptions of leadership."

Recent community-wide studies of teachers in Jewish schools in Boston, Los Angeles, and Miami, in addition to CIJE'S study of teachers in Atlanta, Baltimore, and Milwaukee have provided valuable information about the backgrounds, careers, and work conditions of Jewish educators (Aron and Phillips, 1988; Gamoran et al., 1997; Naava, Margolis, and Weisner, 1992; Sheskin 1988). However, none of these studies have focused on gender differences. Considering the amount of gender inequality among teachers that has appeared in other

2

contexts, it is important to find out whether the same condition holds in Jewish schools. The
existence and degree of gender differences may have important implications for the recruitment,
training, and retention of teachers in Jewish schools.

This report explores gender differences in three related areas: career paths, Judaic and
educational backgrounds, and current work conditions. It seeks to answer the following
questions:

- *Do teachers differ by gender in their reasons for entering Jewish education?*

- *Do teachers differ by gender in the length of their experience and their commitment to the
  profession of Jewish Education?*

- *Do teachers differ by gender in their early childhood Jewish education?*

- *Do teachers differ by gender in their formal training as Jewish educators?*

- *Do teachers differ by gender in regard to the conditions of their work (i.e. hours, salary,
  benefits)?*

METHOD OF STUDY

In 1992-93, the Council for Initiatives in Jewish Education (CIJE) in collaboration with the
Jewish communities of Atlanta, Baltimore, and Milwaukee conducted a study of all Judaica
teachers in the day schools, congregational schools, and pre-schools in those communities. A
survey was administered to the entire population of Judaica teachers (1192), and a response
rate of 82% (983 teachers) was obtained. Formal in-depth interviews were conducted with 125
Jewish educators, including teachers and educational directors of day schools, congregational
schools, and pre-schools, as well as central agency staff and Jewish educators in higher
education. The findings on teachers are highlighted in CIJE's Policy Brief (Gamoran et al., 1994)
and reported more completely by Gamoran et al. (1997).

The data for this paper are taken primarily from this survey. Data from the in-depth

3

interviews highlight the quantitative findings. In analyzing and reporting the results, we have avoided sampling inferences (e.g., t-tests) because we are analyzing population figures, not samples.∧ Data from all three communities are combined for all analyses. Despite some differences, the teachers in each community are largely similar. The broad comparability of results from the three communities – delineated in the study mentioned above – suggests that the gender differences and similarities presented here are likely to resemble that of many other Jewish communities.

FINDINGS

DEMOGRAPHICS

Eighty-four percent of the teachers (804 teachers)[1] in the three communities of Atlanta, Baltimore, and Milwaukee are women. Overall, teachers are divided fairly evenly among day schools (31% or 302 teachers), congregational schools (40% or 392 teachers) and pre-schools (29% or 283 teachers)   However, almost all pre-school teachers (99%) are women, while 29% of day school teachers and 18% of congregational school teachers are men. Among Orthodox day schools, the percentage of men rises to 45%, while in non-Orthodox day schools ?pen only account for 6% of the teachers. In total, almost half (48%) of the male teachers work in Orthodox day schools, while 43% work in congregational schools.

Almost all (87%) of the teachers are Jewish, the 3% who are not Jewish are all women. Sixty-two percent of male teachers are Orthodox, while women are spread fairly evenly among the denominations: 33% Reform; 27% Conservative; 26% Orthodox; 8% Traditional; and 5% Other. Men and women are similarly represented in all age categories: The mean age of both groups is 36. Eighty percent of women and 84% of men are married; thirteen percent of women and 14% of men are single   Three percent of men are separated, divorced, or widowed,

---

[1] There were 22 cases with missing data on gender.

whereas 7% of women have this marital status. Lastly, while the large majority of men (89%) and women (86%) teachers were born in the United States, 8% of women were born in Israel compared to 3% of men. Ninety-four percent of the Israeli-born teachers in the three communities are women.

JEWISH EDUCATION AS A CAREER

Entering Jewish Education

Most teachers enter Jewish education for its intrinsic, as opposed to any extrinsic, rewards (Gamoran et al., 1997). *Do teachers differ by gender in their reasons for entering Jewish education?*

As Table 1 indicates, men tend to value those intrinsic rewards associates with the teaching and learning of Judaism more than women do, though most women did value those highly. Eighty-five percent of men as compared to 63% of women reported that "teaching about Judaism" was a very important reason for entering Jewish education. Similarly, a greater percentage of men indicated that "love of Judaism" and "learning about Judaism" were very important to them [2]

In contrast, greater percentages of women favored rewards associated with teaching children as important factors in choosing to enter Jewish education. Eighty-two percent of women, as compared to 63% of men, reported "working with children" as very important. Similarly, though percentages were low for both groups, more than twice the percentage of women than men saw "recognition as a teacher" as a very important reward.

In regard to the extrinsic rewards of teaching, more women (46%) than men (14%)

---

[2] Gender differences, overall, hold across setting (day school and congregational school) and denomination of the school (Orthodox and non-Orthodox). The only exception is that similar percentages of men and women in congregational schools and non-Orthodox schools reported "learning more about Judaism" as a very important reason for entering into Jewish education.

5

tended to consider the "part-time nature of the profession" as a very important inducement to entering the field. Also, more women (46%) than men (20%) entered Jewish education because it could provide a "supplement to their income." Seemingly, when men enter Jewish education it is more likely that they desire a full-time position in which the salary from Jewish education would be their main source of income. Findings on this issue = 56% of men as compared to 12% of women consider their salary from Jewish education to be their main source of income = mutually confirm this proposition.

When asked about the factors that influenced their decisions to work in the school at which they are currently employed their answers corroborate the previous findings. As Table 2 indicates, the highest percentages of both men and women reported that scheduling was an important consideration though 89% of women compared to 78% of men indicted this as a consideration. While the religious character of the school was ranked second by men (in regard to the percentage of respondents who indicated it to be a factor), it was the fourth most important consideration for women (as a group).

Both sets of findings illustrate differences between men and women in the factors used in considering whether or not to enter the field of Jewish education and in selecting a particular school at which to work   For men, religious (Jewish) considerations seem to dominate. For women, teaching children in a flexible work environment seems most important. Interviews conducted with Jewish educators highlight these differences.

A woman teacher told about beginning to teach Sunday school in order to pay for her tuition in a graduate social work program.

> In thinking about what I really loved to do during those two years that I was in graduate school, I realized it was the teaching. I liked my Sunday morning better than anything else, better than social work school.

Moreover, many women related how the part-time nature of the profession facilitated their entrance into the field

6

I had my third child, and I was feeling like I needed to get out and do something, but I couldn't do something on a full-time basis. [Working as a Jewish educator] seemed to coincide with what I needed at the time.

I worked first in the public schools. When my children were little, I could only accept the half-day kind of job, so that is how I originally started working [in Jewish education].

Experience and Commitment

As a group, Jewish teachers show considerable stability. Only 6% of all teachers were in the first year of Jewish education when they responded to the survey, while thirty-eight percent had taught for more than 10 years (Gamoran et al.1994) In addition, only 6% plan to leave Jewish education during the next several years (Gamoran et al., 1997). As another measure of commitment, when asked if they considered Jewish education to be a "career," 69% of full-time teachers and 54% of part-time teachers said "Yes." *Do teachers differ by gender in the length of their experience and their commitment to the profession of Jewish education?*

As Table 3 illustrates, both men and women show considerable stability. Slightly more than two-thirds of both men and women have worked in Jewish education for six years or more. The slightly higher percentage of men compared to women who have worked in Jewish education for 21 or more years may be accounted for by the growth in non-Orthodox day schools and pre-schools over the last two decades.

In regard to their length of employment in their current position (see Table 3), there are no substantial gender differences. For both men and women, approximately 60% have worked in their current position for only 5 years or less.

The future plans of men and women similarly show little differences (see Table 4). Only 6% of men and 7% of women plan to leave Jewish education. Also, 67% of men and 64% of women plan to remain in their same position.

These findings indicate that both men and women, regardless of their diverse reasons for entering Jewish education, tend to stay for a considerable period of time. Yet, do they see their participation in Jewish education in the same way? As Table 5 shows, while almost three-

3 ; quarters of the men consider Jewish education as their career, women report a somewhat lower percentage (60%). This may be due to the larger percentage of women as compared to men who work part-time in Jewish education (see Table 8). However, if we only examine the findings for full-time educators (those working 25 hours or more), the gender difference is even greater (see Table 5). While almost all men who teach full-time consider Jewish education to be a career, only 62% of women who teach full-time feel the same way.

Summary

Men and women indicated substantial differences in their reasons for entering Jewish education. Men tended to view their decision as one that would provide them with the opportunity to learn continually and teach about Judaism. Similarly, their religious character of the school was a strong factor in their determination of where to work. In contrast, women viewed their choice of entering into Jewish education as an opportunity to teach children. The flexible and part-time nature of Jewish schooling facilitated their entrance and was the primary consideration in deciding at which Jewish school to work.

However, once they entered the field of Jewish education and selected a school at which to work, their career paths become similar. Both men and women have stayed in Jewish education for a considerable length of time, and both are comparably new to their current positions though they overwhelmingly plan to stay in them. Nevertheless, their conceptualization of their work seems to be substantially different. Even when only examining the findings for full-time teachers, a substantially lesser proportion of women as compared to men view Jewish education as a career.

JUDAIC AND EDUCATIONAL BACKGROUNDS

Early Jewish Education

Compared to the general population (see Kosmin's *Highlights of the CJF 1990 National*

8

Jewish Population Survey, 1991), a greater percentage of Jewish teachers had a formal Jewish education as children. While 22% of men and 38% of women in the general population had no Jewish education as children, Gamoran et al. (1994) reported that only 10% of teachers were not formally educated as Jews during childhood. Differences by gender were not reported. *Do teachers differ by gender in their early childhood Jewish education? Do they differ by gander in their formal training as Jewish educators?*

As indicated in Table 6, fifty-four percent of men reported attending a day school, yeshiva, or school in Israel, and only 2% indicated not attending any Jewish school before the age of 13. In comparison, only 30% of women attended a day school or school in Israel, while 15% did not attend any school before 13. Similarly, while 61% of men attended a day school, yeshiva, or school in Israel after the age of 13, only 30% of women did. In addition, while only 15% of men did not have any formal Jewish education after the age of 13, 36% of women did not. These gender differences seem to follow the pattern in the general population.

Formal Training

Gamoran et al. (1997) argued that preparation for a career in Jewish education should consist of formal training in both education and Jewish studies. Formal training is defined by having a degree or certification in that area. Overall, 19% of Jewish teaches have training in both education and Jewish studies, while 34% are trained in neither. As table 7 indicates, men and women illustrate similar proportions. Twenty-one percent of men and 18% of women have formal training in both education and Jewish studies, while 37% of men and 33% of women lack formal training in both areas.

The largest percentage of teachers (48%) have formal training in either education of Jewish studies. Differences between men and women are substantial here. While only 26% of women can be considered to have formal training in Jewish studies, 56% of men have training in Jewish studies (see Table 7). [Forty-one percent of men with training in Jewish studies have

rabbinic ordination or smicha.] In contrast, while only 28% of men can be considered to have formal training in education, 59% of women have training in education. The figures present almost a mirror image of each other. In accordance with the emphasis on educational training found among women 56% of women as compared to 29% of men have previous experience working in general) education.[3]

Summary

Men come to the profession of Jewish education with a stronger Judaic background than women due to their early childhood education and additional formal training (often rabbinic). At the same time, women approach their work in Jewish education with a stronger foundation in educational pedagogy gained either through study or experience in general education. Perhaps not surprisingly, these findings are consistent with their stated reasons for entering the profession. As mentioned earlier, men tended to enter Jewish education to continue their life-long engagement with Judaism, while women tended to enter Jewish education in order to teach children.

CURRENT WORK CONDITIONS

Full-time Employment?

The field of Jewish education offers primarily part-time employment opportunities for teachers. Sixty-eight percent of teachers in Jewish schools are part-time (Gamoran et al., 1994). Consequently, salary levels tend to remain low and benefits, such as health and pension plans, are unavailable to most teachers (Gamoran et al., 1997). Yet, do teachers differ by gender in regard to the conditions of their work?

---

[3] Gender differences exist among congregational school teachers, but these do not follow the same pattern as described for the total population of teachers. Among teachers in congregational schools, a greater percentage of men (60%) as compared to women (41%) are trained in neither Jewish studies or education. In addition, while a greater percentage of women (50%) as compared to men (24%) are trained in education, almost the same percentage of men (28%) as women (22%) are trained in Jewish studies.

10

As Table 6 illustrates, a greater percentage of men (46%) as compared to women (29%) work full-time in Jewish education. Among those who work full-time in Jewish education, 73% of men and 72% of women do so at one school. The remaining teachers put together the equivalent of full-time employment (25 hours of more) through working at two or more schools.

Salaries

While salary levels, overall, are low in Jewish education, they are even lower for women than they are for men As indicated in Table 9, while 41% of men earn over $30,000, only 3% of women take home such high earnings. Instead 44% of women earn less than $5,000, and another 44% earn between $5,000 and $19,999. The distribution of men's salaries is bimodal with over three-quarters of men located either between $1,000 and $5,000 or over $30,000. By contrast, the distribution among women has a single mode between $1,000 and $5,000 followed by a quick drop and then a gradual tapering off in subsequent categories.

Some of this wage gap may be due to the larger percentage of men as compared to women who teach full-time. However, including only those who teach full-time, the differences are even greater (see Table 9). While three-quarters of men who teach full-time earn over $30,000, only 8% of women who teach full-time take home similar earnings. Instead, half of the women working full-time earn less than $15,000.[*]

Benefits

While employer contributions to a health plan, overall, are unavailable to most teachers, they are less available to women than men. As Table 10 illustrates, a greater percentage of men (36%) as compared to women (24%) reported that they were offered health benefits from their schools. When only full-time teachers are considered, the difference is even greater: 61% of

---

[*] For congregational school teachers, there are no substantial differences between men and women in salary levels. Slightly more than three-quarters of both groups received between $1,000 and $4,999. The lack of difference in salary levels exist among congregational school teachers despite the larger percentage of women as compared to men who are trained in education and a similar percentage of men and women who are trained in Jewish studies (see Footnote 2.)

11

men and 35% of women who work full-time reported the availability of health benefits.

There is not a substantial difference between men and women in regard to pension benefits, as only one-quarter of both groups has that option (see Table 10).

Summary

The findings illustrate that, even when controlling for hours of employment (full-time vs. part-time), substantial differences exist in salary level and health benefits offered to women as compared to men These differences exist despite the fact that men and women have similar stability and commitment to the field of Jewish education, and similar percentages of men and women are trained in both or neither education and/or Jewish studies. As the earlier findings indicated, they do differ in regard to the emphasis on a Judaic or general education background. However, it would be doubtful if this alone accounted for all or most of the reported differences.

EXPLAINING DIFFERENCES IN SALARY LEVELS

To explore factors that may account for the differences in salary levels between men and women teachers, a linear regression was used with reported salary levels as the dependent variable.[5] This variable is coded as a scale of 1 to 8 with each point corresponding to the salary categories listed in Table 9, which range from less than $1,000 to $30,000 or more. *The primary hypothesis is that while gender differences exist in salary levels among educators, this may be due to other factors, such as hours employed and professional training.* The gender of the respondent is initially the only variable entered into the equation, as shown in Table 11. This shows that gender by itself, is a significant predictor of salary level, though the findings also

---

[5]The authors feel confident in the validity of the respondents' answers to the question on salary, but they have certain reservations in regard to their answers to the question on benefits. Informal interviews with educational leaders in the three communities suggest that teachers may not be aware of all the benefits available to them. Thus, the authors have decided to investigate further only the issue of salary differences between men and women.

12

indicate that gender only accounts for 10% of the variation in salary levels.6

Next, three setting-related variables are entered into the equation in order to account for more of the variation in salary level and to determine if gender is still significantly related when other variables are considered. The findings indicate that hours of employment (full-time or part-time) and the setting (pre-school, day school, or congregational school) are related to differences in salary level. Not surprisingly, full-time educators and day school educators earn more than part-time educators and those who work in pre-schools or congregational schools or pre-schools. The drop in the coefficient for gender between the first column (1.72) and the second column (.93) indicates that almost half of the raw gender difference is attributable to setting and hours of employment. Still gender remains a significant predictor of salary level even after controlling for setting and hours. These variables together account for 65% of the variation in salary levels.

Next, six variables related to the background and career of the respondent are entered into the equation. The findings indicate that experience in Jewish education, formal training in education, and formal training in Jewish studies are all significantly related to differences in salary level. Only the respondents' willingness to leave Jewish education is not significantly or substantially related to salary level. Together, these variables account for 68% of the variation in salary level.7 Notably even when controlling for all of these personal characteristics, gender is still a significant predictor of salary level.

Lastly, considering the possibility that ideological differences between the denominations may influence salary levels of teachers, a variable indicating if the setting in which the respondent worked was Orthodox was entered into the equation. This also was significantly

---

6 Significance levels are reported here purely as a convention. As the data are based on a population, sampling inferences such as significance tests are not really appropriate.

7 The linear regression was run with additional independent variable that indicated whether or not the respondents considered their work in Jewish education as a "career." The results did not differ much from those described in Table 11; the significance and strength of the relationships remained relatively the same.

related to differences in salary levels. Even after accounting for all of these factors, gender was still found to be a significant predictor of salary level.

Is the gender difference meaningful? The coefficient of 1.72 in the first column (see Table 11) means that on average, males tend to be ahead of females by almost two categories on the salary scale (see Table 9). After controlling for other relevant conditions, that difference drops to .39 or slightly less than one salary category. This difference is still larger than the gap between experienced and inexperienced teachers (a maximum of .62). It is also larger than the gap between trained and untrained teachers; a teacher trained in both education and Jewish studies would be about .59 categories ahead of an untrained teacher (.26 + .33 = .59). Viewed in this way, the gender difference in salaries must be regarded as substantial.

DISCUSSION OF FINDINGS

The findings from the CIJE study in the three communities of Atlanta, Baltimore, and Milwaukee suggest that important gender differences exist among teachers of Judaica. These findings have important implications for the recruitment, training, and retention of teachers in Jewish schools.

JUDAISM OR WORKING WITH CHILDREN

Women tend to enter Jewish education primarily because they enjoy teaching children, and the structure of Jewish schooling allows them flexibility in how much and when they work. Their emphasis on being formally training in education, while lacking training in Jewish studies, is consistent with these reasons   In contrast, men tend to enter Jewish education primarily because of their continued interest in Judaism. Similarly, their emphasis on being formally trained in Judaic studies (as well as their more intensive early childhood Jewish education), while

lacking training in education, also seems consistent with their reasons for entering Jewish education. In a similar vein, Kalaian and Freeman (1994) found that women teacher candidates were more cognizant of the ambiguities and uncertainties of teaching -- a practice in which the appropriate representation of subject matter in the classroom depends on the aptitudes and learning histories of the particular students. It is important to note that men and women both valued working with children and teaching/learning about Judaism and that some women have training in Judaic studies while some men have training in education. However, the differences between men and women in both areas are substantial and potentially meaningful.

As mentioned earlier, Jewish educators should be formally prepared in both education and Jewish studies. Shulman and his colleagues (Shulman, 1986; Wilson, 1988; Wilson, Shulman, and Richert 1987) have suggested that successful teaching requires teachers to have knowledge of pedagogy (education), knowledge of content (Jewish studies), and pedagogical-content knowledge (knowing how to bridge the gap between the learner and the subject matter). If men tend to enter Jewish education only with knowledge of Jewish studies and women are acquiring only educational knowledge, this poses complex problems for developing in-service programs that attempt to address these deficiencies. Not only must teacher-educators consider how to develop the pedagogical-content knowledge of all teachers, their approaches must take into account the seemingly gender-linked nature of teachers' knowledge -- men's knowledge of content and women's knowledge of pedagogy. Perhaps, in-service programs need to go "against the grain." Programs designed to contribute to the pedagogical proficiency of those (mostly male) teachers who are deficient in this area should be designed for with their particular learning styles of men in mind Similarly, in-service programs designed to enhance content knowledge should be designed to fit with the ways women tend to learn (Belenky et al., 1986).

15

HAVING A CAREER IN JEWISH EDUCATION

Despite these initial differences, men and women have, perhaps surprisingly, similar lengths of experience in Jewish education. They also show a similar degree of tenure in their current school, and both groups overwhelmingly intend to stay in Jewish education. However, they differ in their perception of whether or not their work in Jewish education is a "career." While a slight majority of women (57%) see Jewish education as a career, almost three-quarters of men (76%) do so. The differences are even greater when considering only full-time teachers.

We can only venture an explanation at this time as to why this difference exists. Jewish education presents applicants with few opportunities for advancement. Within individual schools, teachers are grouped together with little stratification in positions, responsibilities, or salaries. Above them exist a handful of educational leadership positions, such as educational director of a congregational school, department head of a day school, or central agency staff. For the majority of teachers, upward mobility is not a possibility. Coupled with this is the finding from a survey of educational leaders (Goldring et al., 1996) - completed at the same time and in the same cities as the survey of Jewish teachers - that approximately one-third of the education leaders are men. This is compared with only 16% of teachers who are men. Seemingly, while vertical career advancement is limited for Jewish teachers as a whole, women may face additional difficulties. Perhaps, they do not consider Jewish education as a "career" because there is no opportunity for career advancement. Perhaps, in addition, many of those women who were interested in pursuing a "career" left or never entered the field of Jewish education. If so, the smaller percentage of women as compared to men who view Jewish education as a "career" is symptomatic of the difficulties involved in recruiting and retaining career-oriented women in a field with limited opportunities for advancement (especially for women), low salaries (especially for women), and a lack of prestige due to having been considered "Women's work." This topic will be examined in a future paper, which will explore differences between educational leaders

and teachers in the three communities.

SALARIES

The most dramatic gender differences among Jewish educators, though perhaps the least surprising, are found in their work conditions. The data show that while almost half (46%) of the men teachers work full-time, most women (71%) work part-time. Yet, this does not account for differences in salary and the availability of health benefits found among men and women teachers. Counting only full-time teachers, three-quarters of men earn over $30,000 In contrast, only 8% of women earn a similar salary. Half of the women working full-time earn less than $15,000. In addition, while 61% of full-time men teachers are offered employer contributions to a health plan, only 45% of full-time women teachers are so offered.

A linear regression analysis was conducted to determine the factors that may account for the salary discrepancy Many factors were shown to predict salary differences - hours of employment (full-time vs part-time); setting (day school, congregational school, pre-school); length of experience in Jewish education; training in Jewish studies; training in education; and the religious character of the school. Nevertheless, even when controlling for all of these factors, gender was still significantly related to differences in salary.

Does this pattern indicate gender discrimination in Jewish education? Although we have no direct evidence on discrimination, inequities among teachers who are otherwise comparable (e.g., in experience, in formal training) must raise discrimination as a possibility. This finding is similar to the findings of the study conducted by Lee and Smith (1990) on salary differences of high school teachers in public, private, and Catholic schools, as described earlier. Jewish education is not immune to the conditions permitting gender discrimination in the secular and non-Jewish religious worlds.

REFERENCES

Acker. S. (1992). Creating Careers: Women Teachers at Work. *Curriculum Inquiry,* 22(2), 141–163.

Aron, I. and Phillips. B (1988). *Findings of the Los Angeles BJE Teacher Census.* Paper presented at the Conference of the Research Network on Jewish Education. New York.

Belenky, M. F., Clmchy, B. M., Goldberger, N. R., Tarule, J. M. (1986). *Women's Ways of Knowing; The Development of Self, Voice, and Mind.* Basic Books.

Frank, N., Margolis. D . and Weisner, A. (1992). *The Jewish School Teacher: Preliminary Report of a Study.* Boston: Bureau of Jewish Education of Greater Boston.

Gamoran, A., Goldring, E., Goodman, R. L., Robinson, B., and Tammivaara, J. (1994). *Policy Brief: Background and Training of Teachers in Jewish Schools.* New York: Council for Initiatives in Jewish Education.

Gamoran, A., Goldring, E., Goodman, R. L., Robinson, B., and Tammivaara, J. (1997). *Teachers in Jewish Schools A Study of Three Communities.* New York: Council for Initiatives in Jewish Education.

Goldring, E., Gamoran, A., and Robinson, B. (1996). Educational Leaders in Jewish Schools. *Private School Monitor.* 18(1), 6–13.

Huberman, M. (1993) *The Lives of Teachers.* New York: Teachers College Press.

Kalaian, H. and Freeman, D. (1994). Gender Differences in Self-Confidence and Educational Beliefs among Secondary Teacher Candidates. *Teaching and Teacher Education,* 10(6), 647–658.

18

Kosmin, B. A., Goldstein, S., Waksberg, J., Lerer, N., Keysar, A., and Scheckner, J. (1991). *Highlights of the CJF National Jewish Population Survey*. New York: Council of Jewish Federations.

Lee, V. E. and Smith J. B. (1990). Gender Equality in Teachers' Salaries: A Multi-Level Approach. *Educational Evaluation and Policy Analysis*, 12(Spring), 57-81.

Lee, V. E. and Smith J. B. (1993). Teachers and Principals: Gender-Related Perceptions of Leadership and Power in Secondary Schools. *Educational Evaluation and Policy Analysis*, 15(Summer), 153-180

Sheskin, I. (1988) *The Miami Jewish Educator Study*. Miami: The Central Agency for Jewish Education.

Shulman, L. S (1986) Those Who Understand: Knowledge Growth In Teaching. *Educational Researcher*, 15(2), 4.14

Wilson, S. M. (1938) *Understanding Historical Understanding: Subject Matter Knowledge and the Teaching of American History*. Unpublished Dissertation, Stanford University, Palo Alto, CA.

Wilson, S. M., Shulman, L. S., and Richert, A. (1987). "150 Different Ways of Knowing": Representations of Knowledge in Teaching. In J. Calderhead (ed), *Exploring Teacher Thinking* (pp. 104-124). Sussex England: Holt, Rinehart, and Winston.

TABLE 1:   REASONS FOR ENTERING JEWISH EDUCATION

Reasons	% of educators who indicated reason as "Very Important"	
	Women	Men
Working with Children	82%	63%
Teaching about Judaism	63%	85%
Love of Judaism	61%	80%
Learning about Judaism	51%	61%
Part-time Nature of the Profession	46%	14%
Supplement to Income	45%	20%
Recognition as a Teacher	29%	12%

TABLE 2:   FACTORS IN CONSIDERING WHERE TO WORK

Factors	% of educators who indicate factor as affecting their decision	
	Women	Men
Hours & Days Available	89%	78%
Location	76%	70%
Reputation of School & Students	67%	62%
Religious Orientation	67%	76%
Salary	49%	58%
Friends Who Work There	47%	44%

TABLE 3:   LENGTH OF EMPLOYMENT

Years Employed in Jewish Education

	Women	Men
5 Years or Less	33%	31%
6 to 10 Years	30%	24%
11 to 20 Years	25%	22%
21 or More Years	12%	23%

Years Employed in Current School

	Women	Men
5 Years or Less	60%	58%
6 to 10 Years	24%	19%
11 to 20 Years	13%	15%
21 or More Years	4%	8%

TABLE 4:   FUTURE PLANS

Plans	Women	Men
Continue in Same Position	64%	69%
Change Schools or Position	6%	8%
Leave Jewish Education	6%	7%
Don't Know	19%	11%
Other	5%	5%

TABLE 5   JEWISH EDUCATION IS A CAREER?

% of teachers who considered Jewish education to be their career

Women	57%
(Only women who work full-time in Jewish education)	62%
Men	72%
(Only men who work full-time in Jewish education)	94%

TABLE 6: EARLY CHILDHOOD JEWISH EDUCATION

BEFORE 13

Type	% of teachers who attended this type of Jewish school	
	Women	Men
None	15%	2%
1 Day/Week Congregational School	28%	15%
2 or More Days/Week Congregational School	27%	29%
Day School, Yeshiva or School in Israel	30%	54%

AFTER 13

Type	% of teachers who attended this type of Jewish school	
	Women	Men
None	36%	15%
1 Day/Week Congregational School	22%	11%
2 or More Days/Week Congregational School	13%	13%
Day School, Yeshiva. or School in Israel	30%	61%

TABLE 7:   FORMAL TRAINING

Areas	% of teachers with a degree or certification in these areas	
	Women	**Men**
Both Jewish Studies and Education	18%	21%
Only Jewish Studies	8%	34%
Only Education	41%	7%
Neither Jewish Studies or Education	33%	37%

TABLE 8:   FULL-TIME?

% of teacher who work full-time (25 hours or more) in Jewish education

	Women	Men
Full-time	29%	46%

TABLE 9: SALARY

	Women	(Full-time Only)	Men	(Full-time Only)
Less than $1,000	3%	- -	3%	- -
$1,000 - $4,999	41%	(3%)	36%	(2%)
$5,000 - $9,999	17%	(18%)	4%	- -
$10,000 - $14,999	17%	(29%)	3%	- -
$15,000 - $19,999	10%	(17%)	3%	(3%)
$20,000 - $24,990	6%	(13%)	3%	(5%)
$25,000 - $29,99')	4%	(11%)	6%	(15%)
$30,000 or More	3%	(8%)	41%	(75%)

TABLE 10: BENEFITS

Type of Benefit	% of teachers who reported being offered the type of benefit			
	Women	(Full-time Only)	Men	(Full-time Only)
Employer Contribution to a Health Plan	24%	(45%)	36%	(61%)
Pension Plan	23%	(18%)	25%	(6%)

TABLE III   EXPLAINING DIFFERENCES IN SALARY

Differences among groups of individuals and settings in salaries from Jewish education.

Independent Variables

Gender (Male = 1)	1.72** (.17)	.93** (.11)	.94** (.11)	.88** (.11)
Congregational School		-2.20** (.11)	-1.88** (.11)	-1.69** (.12)
Pre-School		-1.64** (.11)	-1.33** (.11)	-1.21** (.11)
Work Full-time (25+ Hours)		1.88** (.10)	1.88** (.10)	1.90** (.10)
Experience 6-10 Years			.33** (.10)	.33** (.10)
Experience 11-20 Years			.59** (.11)	.60** (.11)
Experience 21+ Years			.63** (.13)	.62** (.13)
Plans to Leave Jewish Education			.01 (.15)	.02 (.15)
Trained in Education			.25** (.08)	.26** (.08)
Trained in Jewish Studies			.40** (.10)	.33** (.10)
Orthodox Setting				.37** (.12)
Constant	3.36** (.07)	4.20** (.09)	3.38** (.13)	3.21** (.14)
R^2	.10	.65	.68	.68

* $p < .05$ ** $p < .01$

Note: Metric regression coefficients with standard errors in parentheses. N = 914 teachers. Equation also includes controls for missing data on sex, works full-time, experience, trained in education, trained in Jewish studies, and plans to leave Jewish education.

TABLE 12: VARIABLES USED IN EXPLAINING DIFFERENCES IN SALARY

Independent Variables	Mean	Standard Deviation
Salary	3.64	2.00
Sex (Male = 1)	.16	.37
Day School	.31	.46
Congregational School	.40	.49
Pre-school	.30	.46
Works Full-time (25 Hours)	.30	.46
Experience Less than 5 Years	.26	.44
Experience 6 - 10 Years	.28	.45
Experience 11 - 20 Years	.24	.43
Experience 21 - Years	.14	.35
Plans to Leave Jewish Education	.07	.26
Trained in Education	.53	.50
Trained in Jewish Studies	.30	.46
Orthodox Setting	.22	.41
Non-Orthodox Setting	.78	.41
Missing Sex	.01	.09
Missing Full-time	.06	.23
Missing Experience	.03	.16
Missing Plans to Leave	.06	.23
Missing Trained in Education	.03	.17
Missing Trained in Jewish Studies	.04	.19

Note: N = 914 teache's

Council for Initiatives in Jewish Education

Fax Cover Sheet

Date sent: Time sent   No. of Pages (incl. cover): 14

To: Seymour Fox; Annette Hochstein   From: Robin Mencher

Organization: Mandel Institute

Phone Number: Phone Number: 212-532-2360

Fax Number: Fax Number: 212-532-2646

COMMENTS:

I will also give this article to SF tomorrow when he is in our office.

November 8, 1993

To: S2 (Sir, Name)

From: Adam

In thinking about "e situation" in messy situations, W *W attached article may be of interest.

How Shall We Study Comprehensive, Collaborative Services for Children and Families?

MICHAEL S. KNAPP

Researchers and evaluators confront difficult challenges in studying comprehensive, collaborative services for children and families. These challenges appear in the interaction of multiple professional perspectives, specification of independent and dependent variables, attribution of effects to causes, and sensitive nature of the programmatic treatment. Given limited knowledge about these complex interventions, they will best be understood through studies that are strongly conceptualized, descriptive, comparative, constructively skeptical, positioned from the bottom up, and (when appropriate) collaborative.

The recent burgeoning of interest and activity in the integration of education and human services should be a signal to researchers and evaluators that there is work to be done. New solutions to old problems—including newly rediscovered ones—have a way of suddenly capturing the attention of policymakers, advocates, and the public. Claims about the new solution proliferate, as do pilot versions, labels for the activity, and purported distinctions among these pilots. Along with these claims, labels, and pilot experiments comes advice to policymakers, practitioners, and funders. And all at once there is a need to do careful, probing research and hardheaded evaluation, to sort among the claims, characterize what pilot initiatives have indeed demonstrated, and discover what the sound and fury signifies.

We are at such a point once again with the integration of education and human services. Voices calling for comprehensive, collaborative services as a solution to the needs of the "high-risk" family and child have built to a crescendo across the past decade and especially the last half dozen years. A number of demonstration projects, both great and small, have been launched. And policymakers are scrambling to make comprehensive, collaborative services standard operating procedure. In so doing, all participants are reenacting a drama of several decades past, when federal initiatives set in motion a wave of interest in service integration (Agranoff, 1991). The reinvention of this programmatic thrust in the late 1980s and early 1990s has some attributes of the earlier episode, only now it is being played out on a grander scale, and with a greater sense of urgency.

This area of social endeavor poses special problems for researchers and evaluators. The purpose of this article is to explore these problems and suggest some ways they might be productively overcome. I accomplish this task by first characterizing this "new" evaluation problem and the literature that addresses it, along with several premises about "good" research and evaluation. Following that, I identify and "unpack" five issues facing researchers and evaluators, and finally I briefly review some ways of meeting these challenges.

A "New" Problem for Research or Evaluation and an Emerging Literature

The very act of naming the target of inquiry hints at the complexity of the research task. With some trepidation, I have chosen *comprehensive, collaborative services for children and families* instead of half a dozen other phrases, knowing that any choice will leave someone feeling left out or underappreciated. I will use the simpler phrase *comprehensive, collaborative services* throughout much of the article (with apologies to educators who wish not to view what they do as a "service"). But in so doing, I include most of what is said about *the integration of education and human services, school-linked services, services integration, interprofessional collaboration, coordinated services for children, and family support*—once again, acknowledging that meaningful distinctions can be drawn among these terms.

The difficulty for those who wish to study comprehensive, collaborative services, however labeled, stems from their complexity and flexibility, the nature of collaborative effort, and the convergence of different disciplines. Complexity derives from the sheer number of players, stakeholders, and levels of the system, as multiple services lodged in different agency or disciplinary contexts, each operating from its own premises about good practice and the "client" or "consumer," join forces in some fashion to influence the life prospects of high-risk families and children. The extent to which their efforts are collaborative defies easy conceptualization, no less description or assessment. The boundaries of research and evaluation design stretch further to handle the idiosyncratic tailoring of effort that is frequently part of collaborative practice and the interplay among agencies or other collaborating partners. Finally, the act of studying such endeavors engages researchers from traditions that do not normally communicate with one another.

MICHAEL S. KNAPP *is an Associate Professor of Educational Leadership and Policy Studies at the University of Washington, College of Education, Mail Stop DQ-12, University of Washington, Seattle, WA 98195. His areas of specialization are policy implementation, policy-to-practice connections, and the education of disenfranchised groups.*

... variety of arenas have been undertaken of comprehensive, collaborative services, so there are a number of examples of what might or might not be useful ways to understand it. These studies are embedded in a larger advocacy literature, which both makes the case for the integration of education and human services (e.g., Levy, Kagan, & Copple, 1992; National Association of State Boards of Education, 1991; National Commission to Prevent Infant Mortality, 1991) and offers advice on how to do so (e.g., Chynoweth & Dyer, 1991; Melaville, Blank, & Asayesh, 1993). The burgeoning of this advocacy literature lends urgency to the task of research and evaluation, because as is typical with compelling ideas about social intervention, enthusiasm outstrips evidence at a rapid rate.

Much of the research and evaluation literature related to comprehensive, collaborative services is fugitive; various attempts to capture what is there have been undertaken recently, among them several comprehensive reviews (e.g., Crowson & Boyd, 1993; Gomby, 1992), selective analyses of effective practices (e.g., Schorr, 1988), annotated bibliographies (e.g., Chaudry, Maurer, Oshinsky, & Mackie, 1992), and the activities of several technical assistance centers such as the National Center for Services Integration and the National Center for Children in Poverty.

Case descriptions and single project assessments predominate in this literature, and they are, understandably, a mixed bag (see Crowson & Boyd, 1993, for a characterization of the literature). Add to that several surveys (e.g., American Public Welfare Association, 1992; Chang, Gardner, Watahara, Brown, & Robles, 1991; Kagan, Rivera, & Lamb-Parker, 1990) and multiple-project comparative studies (e.g., Marzke, Chimerine, Morrill, & Marks, 1992), as well as formal attempts to study statewide initiatives of several kinds (e.g., State Reorganization Commission, 1989; Wagner et al., 1994) and other demonstration projects (e.g., New Beginnings Team, 1990; Nucci & Smylie, 1991; Wehlage, Smith, & Lipman, 1992). Some of these studies employ elaborate, multiyear designs, especially those done in conjunction with the more ambitious initiatives.

The methodological literature to date is thin. Although some would argue that the existing, voluminous literatures regarding the study of social interventions are adequate to the task of studying comprehensive, collaborative services, others have begun to recognize the special methodological issues that arise. Attempts to address the methodological questions have often approached the matter straightforwardly, for example by adapting conventional experimental methods to the evaluation task (see Gomby & Larson, 1992). Other treatments note special challenges to evaluation stemming from the complex, incomplete implementation that characterizes so many collaborative ventures (Kagan, 1991). Stemming from the family support literature, a more radical critique has put the spotlight on assumptions underlying conventional efforts and suggested a more participatory alternative framework (see Weiss & Greene, 1992). Work in this tradition argues for a new vision of evaluation and research, emphasizing a collaborative relationship between those studied and those who are carrying out research. Another alternative vision of evaluation places more emphasis on evaluation as an ongoing, developmental learning process serving both internal and exter-

... viewed by one extensive study of a services-integration evaluation and its impact on policy and practice (Stake, 1986). In addition, several interdisciplinary meetings have begun to assemble wisdom about the evaluation task (e.g., Family Impact Seminar, 1993a; also the AERA/OERI Invitational Conference mentioned in the notes).

Some Premises About Research and Evaluation in This Area

The methodological writers have begun to pinpoint several important dimensions of the problem. But before launching into a discussion of method, it is important to be explicit about several assumptions.

First, discussion of method includes considerations of paradigm, that is, the assumptions we make about how the world works and what constitutes evidence and knowledge about it. Although there are fundamental differences at issue, this article assumes that alternative paradigms can support complementary examinations of a phenomenon, especially one of this complexity. In particular, paradigms supporting qualitative and quantitative studies are necessary both to make sense of comprehensive, collaborative services and to respond to the many audiences who wish to understand these social interventions.

Second, I am assuming that good conceptualization of what is to be (or what has been) studied is essential to research and evaluation. Underconceptualization plagues much research; empirical "fishing" expeditions unguided by a sense of what concepts are relevant and what relationships are likely to yield little of value. Being clear about what one is studying is half the battle. Much of what is talked about as a problem of measurement boils down to the task of constructing and operationalizing theories of social needs and the means for addressing these needs.

Third, I am making few distinctions in this case between *research* and *evaluation*. For this topic of study, both are concerned with systematic learning about the design, conduct, and impacts of a form of social intervention aimed at a broad range of human needs. To be sure, studies commissioned as evaluations are overtly political—that is, more directly constrained by stakeholder interests and expectations—and pursue a more explicitly value-laden set of purposes than research by any other name. However, in the realm of popular social interventions, all research has political and evaluative overtones, and, regardless of intention, may be enlisted in the debates regarding the merits of one or another initiative. The terms *research* and *evaluation* will therefore be used somewhat interchangeably in this article, though doing so obscures some important subtleties.

A few further comments will define how the term *evaluation* is used in this article. Drawing on work by authors such as Scriven (1974), Cronbach and Associates (1980), and Patton (1978), among others, the term is broadly construed to include a wide range of systematic attempts to make sense of social interventions for broad stakeholder audiences and policy communities. In particular, I am not restricting the term to investigations aimed at figuring out whether initially stated program goals are achieved and to what degree; such designs typically pay too little attention to the evolution of intentions over time and to the unanticipated facets of implementation or effects that crop up along the way. Nor am I assuming that randomized, ex-

_____ _____, xvmauve anu summanve evaluation purposes, or process versus impact studies. In other words, I am assuming that all evaluations are in a fundamental sense formative (see Cronbach & Associates, 1980) and that proper attributions of impact to cause can only be made by understanding the process that produced the impacts (see Patton, 1978).

Issues Confronting Research and Evaluation

Five sets of issues confront researchers and evaluators wishing to make sense of comprehensive, collaborative services for children and families. These issues are present to some degree in studying many complex interventions, but they are demonstrably acute in this case.

1. *Engaging divergent participants' perspectives:* For whom and with whom are we undertaking research on comprehensive, collaborative services? How should the perspectives of different research and service disciplines, professionals and consumers, and diverse agencies be reflected in the design, conduct, and interpretation of studies?

2. *Characterizing (and measuring) the elusive independent variable:* What exactly is it that we are studying?

3. *Locating (and measuring) the bottom line:* What would indicate that delivering human services in a comprehensive, collaborative form had achieved some desirable ends? What ends would be included in a such an evaluation—health, education, welfare, the reform of human service systems, or all of the above?

4. *Attributing results to influences:* Given so many possible influences, what is to be taken as the result of what?

5. *Studying sensitive processes and outcomes:* How do we capture what is going on without intruding on the subtle (and often confidential) interaction between service providers and consumers of services?

These issues have been framed using conventional causal terms—independent and dependent variables—not to imply that a particular research paradigm is more appropriate, but rather to use a language that is most widely understood by members of the research and evaluation community.

Engaging Divergent Perspectives: Can We Speak Everyone's Language?

Comprehensive, collaborative services efforts—and attempts to study them—inevitably involve the perspectives of different stakeholders and participants. Almost by definition, more than one professional discipline and the traditions of research that are typically used to study these activities are implicated. In addition, the perspectives of clients or consumers are relevant to understanding what is going on and even to framing questions and research designs. Finally, because a given initiative usually involves more than service delivery, the perspectives of different agency leaders and policymaking or sponsoring groups are central to understanding comprehensive, collaborative services as a systems phenomenon.

How should all these perspectives be represented in the design, conduct, and interpretation of research? There is no easy answer, and the answer always reflects the political

research traditions. In approaching an intervention involving school-based health and mental health clinics, for example, investigators steeped in educational evaluation, social work research, or health research are likely to zero in on different facets of the intervention, measure different things, and construct accounts of the program's effectiveness on different bases. All three perspectives would be helpful in framing and carrying out a research strategy. To arrive at such a strategy means communicating across disciplinary boundaries about assumptions, focus, productive measures, acceptable evidence, and so on. At the least, results may need "simultaneous translation" (as at the United Nations) to make sure that different research communities understand each other (e.g., this article may need to be translated into terms that would scan to individuals primarily engaged in public health research or social work research).

Although the language problem just described can be and often is addressed in a given study (e.g., through multidisciplinary teams of researchers), a more difficult language gap yawns between those who carry out research and those who are studied. Some researchers seek to close this gap by engaging the consumers of collaborative services as collaborators in the act of studying these services (e.g., Weiss & Greene, 1992). Although there are obvious advantages to the researcher (and the consumer) in doing so—among them, increased access to participants, the prospect of better quality data, and more accurate rendering of the participants' perspectives and experiences—there are also possible trade-offs in time, complexity, analytical distance, and the sophistication of research designs.

Evaluative studies carried out in the public eye add a third set of perspectives that must be engaged and accommodated—those of powerful stakeholders who are involved in the initiative under study, have an interest in its outcomes, or sponsor the evaluation. If nothing else, this fact reduces the researcher's room for maneuvering, necessitating compromises that may "buy" an audience's support while weakening the study's evidence base or design logic.

Characterizing (and Measuring) the Elusive Independent Variable: Is There One?

Like other broad domains of social reform (e.g., school restructuring), the integration of education and human services takes many forms and has different meanings. This makes for an independent variable—the programmatic factors presumed to bring about results for individuals or systems—of some complexity. In many manifestations of comprehensive, collaborative services, the notion of the independent variable itself ceases to be a *fixed treatment*, as conventionally assumed by experimental research designs, and becomes instead a *menu of possibilities* accompanied by a series of supports that facilitate consumers' interaction with these possibilities.[1]

The meanings of comprehensive, collaborative services range from relatively low-intensity efforts to coordinate the work of different professionals to intensive, highly integrated arrangements; some writers reserve the term *coordination* for the least intensive end of this continuum and

- *Enhanced referral* of children or families for professional help of one kind or another (e.g., as in community-based programs described in Marzke et al., 1992).
- *Coordinated management of "cases,"* as when children or families require more than one specialized human service (see James, Smith, & Mann, 1991).
- *Colocation of services,* such as health or mental health professionals in a school building—a key feature of "full-service schools" (Dryfoos, 1994)—or various specialists in a community multiservice center (see Marzke et al., 1992).
- *Enhanced communication and information sharing* among providers of different human services through joint databases, liaison activity, and agreements (e.g., regarding confidentiality) that encourage information sharing, argued by some to be essential to family-centered, coordinated services (e.g., see Coulton, 1992).
- *Sharing of resources,* as in discussions of fiscal strategies supporting coordinated services, the commingling of funds originally intended for separate services, or the pooling of nonfiscal resources (see Cutler, 1994; Farrow & Joe, 1992; Garvin & Young, 1993; Kirst, 1994).
- *Reconceptualization of human services,* by altering the conceptions of existing professional roles (e.g., subsuming a kind of counseling function in the teaching role), developing new roles such as "integrated services specialist" (see Wilson, Karasoff, & Nolan, 1993) or even rethinking the relationship between professionals and consumers, as in the conceptions of family-centered, client-responsive service (Weiss & Greene, 1992) or "consumer-guided" schools (Hooper-Briar & Lawson, 1994).
- *Joint planning and execution of services,* for example, in various teaming arrangements, where different professionals (and others, such as parents) pool ideas, orchestrate a plan for helping children or families that draws on the expertise of more than one distipline, and in varying degrees carry out the plan through joint effort (see Robison, 1993; Hooper-Briar & Lawson, 1994).

Comprehensive, collaborative services may involve one or virtually any combination of these meanings. In addition, such initiatives often take place on multiple levels of the human service system and may be designed to change the way that system functions (Agranoff, 1991; General Accounting Office, 1992). Nearly all intend to integrate efforts at the service delivery level, but that often requires some integration one level up, among individuals and organizations providing the first layer of management support to direct service providers—school principals, clinic directors, field supervisors for outreach workers—and at the policymaking level as well, among school districts, regional or state social service agencies, and so on. Indeed, efforts to mount comprehensive, collaborative services may target changes in the actual services available to families and children (e.g., Philliber Research Associates, 1994), the service-providing system (e.g., White, 1993), or both (e.g., Wehlage et al., 1992).

The fact that so many kinds of arrangements share the same generic label cries out for ways to conceptualize the differences in terms of common dimensions, and there

boyd, 1993; Kagan, 1991; Golden, 1991; Morrill, Reisner, Chimerine, & Marks, 1991; Schorr & Both, 1991). At a minimum, the following dimensions of difference are involved. First as noted previously, comprehensive, collaborative services initiatives may address system reform primarily, the actual services provided to particular families and children, or both. Second, the arrangements differ in the extent to which distinct services are actually changed or redefined through collaborative effort, or simply relocated or made more accessible. Third, the degree to which resources, control, and power are shared among the collaborating partners varies. Fourth, the scale and scope of arrangements vary enormously, from local arrangements involving only two service sectors to massive statewide initiatives that bring together many sectors. Finally, arrangements differ in what might be termed the flexibility or mutability of treatment—that is, the degree to which the services provided to any given child or family are individually tailored, and even changeable over time.

The last dimension generates some of the biggest puzzles for researchers. If each consumer accesses the human service system in a different way, or in a way that changes over time, then there may be no programmatic independent variable to study. Or put another way, it is extremely difficult to characterize what comprises collaborative service over a given period of time. One researcher discussing collaborative arrangements for young children put it as follows:

> Since collaborations are designed to be flexible and meet changing needs, their implementation is never complete. No precise definition of implementation exists because it is a highly idiosyncratic and mutable condition. Indeed, the strength of collaborations is that they are tailored to meet changing local circumstances. For example, it is not uncommon to find collaborations that deem themselves well implemented one day and fledgling the next. Such changes are predictable and underscore the evaluation dilemma; while implementation flux is a practical necessity, it remains an empirical nightmare. (Kagan, 1991, p. 74)

Because the independent variable has many meanings, both across and within collaborative services arrangements, researchers and evaluators may often be talking past each other, and not talking about the same thing, even within the same study. Beyond the question of figuring out what is being studied and regardless of which conception of collaborative services we employ, the intervention is almost always going to comprise multiple, often separate components. Simply multiplying the number of independent variables (as in multivariate correlational designs) is no real answer; one rapidly runs out of analytic capacity to handle and interpret the many discrete variables that come to mind, and one misses the "glue" that may bind these elements together into a more integrated whole.

The researcher is left with difficult questions: How to describe the independent variable(s) under study? What are its conceptual boundaries? What isn't part of the independent variable(s)? What are the most meaningful units (and levels) of analysis? What indicators most efficiently capture the presence and mutability of the independent variable(s)?

...rts varied as the independent variable(s) may be in studies of comprehensive, collaborative services, so may the dependent variable(s) be. In part a reflection of the differences in perspective and paradigms held by the different services that are integrated, the initiatives under study can aim at remarkably different outcomes, among them the academic achievement and attainment of children, their social adjustment or health status, family welfare, and so on. The temptation to which ambitious collaborative services efforts often succumb is to say, in effect, "all of the above."

Whatever the stated goals of a collaborative services arrangement, the researcher's attempts to pinpoint outcomes face three issues: (a) the large number of possible outcomes, (b) the interdependence among them (including developmental interdependence over time), and (c) the range of abstraction from discrete, modest outcomes (e.g., children immunized by age 2) to those that are more global and complex (e.g., children adequately educated for further education and work roles). Consider the following child and youth outcomes, offered by one group of researchers as a core list around which outcome accountability might be developed (adapted from Schorr, 1994). The outcomes are conceptualized as higher rates of:

- Healthy births (as indicated by decreases in low birthweight babies and births to school-age mothers; high utilization of prenatal care).
- 2-year-olds immunized.
- Children ready for school (as indicated by completion of immunizations, detection and remediation of preventable health problems, no signs of abuse or neglect, or school readiness measures as identified by preschool or kindergarten).
- Children succeeding in elementary, middle, and high school (as indicated by academic achievement measures and lower rates of truancy, retention in grade, suspensions from school, dropping out, or placement in special education).
- Youngsters avoiding problematic behaviors (as indicated by lower rates of school-age pregnancy, substance abuse, sexually transmitted diseases, or involvement in violence either as victim or perpetrator).
- Young adults who are self-sufficient
- Children in families with incomes over the poverty line.

The items in this list are only a selection from among the many possible outcomes that may be relevant to a given comprehensive, collaborative services arrangement. To be sure, a more discrete subset of these outcomes might be the focus of inquiry, as in one evaluation (Philliber Research Associates, 1994) of a school-community partnership aiming at children's (a) persistence and safety in the home (indicated by rates of abuse, children's removal from the home by social services), (b) noninvolvement in the juvenile justice system, and (c) persistence and performance in school (indicated by students' absences, work habits, social-emotional growth, and academic performance). But studies are just as likely to attend to diverse facets of child and family welfare, as in a current study one state's

- Basic family needs: access to food, clothing, transportation, and child care.
- Employment: jobs for parents and young adults (for those seeking employment).
- Health care use: participation in public health services, incidence of injury or illness, and access to medical and dental care.
- Emotional health: self-reported depression, suicidal thoughts, and problems with hostility and anger.
- Youth behaviors: rates of sexual activity and teen pregnancy.
- School performance: students' grades and classroom behaviors.

The outcome puzzle is especially troubling when a collaborative services initiative encourages different arrangements across sites, as in the case just cited, or when services are individualized for each consumer. In such cases, researchers must attend to a wide range of possible outcomes, though not necessarily for whole populations. Furthermore, if the outcomes represent a developmental progression over time, as in the first list presented previously, then later outcomes are dependent on earlier ones, and the ultimate impact of the collaborative services arrangement will have to be tracked over long periods.

The outcomes described previously apply to individuals and groups, and despite some difficulties in measurement, are relatively discrete and clear. System outcomes are generally not as discrete or clear as those that apply to individuals. Take for example, the challenges involved in capturing the following kinds of system outcomes: penetration of services into communities or the "community embeddedness" of service systems (Bruner, 1994), agency restructuring and shared authority needed to realize more integrated forms of service organization and funding (General Accounting Office, 1992), "consumer-guided and consumer-driven schooling" (Hooper-Briar & Lawson, 1994), or "deep-structure systems changes" related to professional behavior, administrative "scripts," and transaction costs (Crowson & Boyd, 1994). As these varied discussions of system outcomes hint, the more human service systems are organized and operate in fully collaborative and integrative ways, the more complex and elusive the outcomes become. For example, it is one thing to capture change in referral rates or utilization of existing services when these are colocated to render them more accessible, because referral and utilization are relatively easy to measure. It is another to capture the slowly emerging views of collaborative practice that might come about as the colocated professionals have greater access to one another and more immediate reasons for interacting with one another.

So the researcher confronts a fundamental question of ends for which the integration of services is presumably the means. What ends (including, but not limited to, stated program goals) might come about as a result of the integration of services? How many can be meaningfully considered and at what level (individual, system)? What outcomes conceptually represent steps taken towards more ultimate ends?

The complexity and mutability of the independent variable(s), combined with the large number and interrelation among dependent variables, generate an attribution task of the first magnitude. To what do we attribute the state of children and families who participate in integrated forms of education and human service? Put another way, how do we account for the level of any of the child and youth variables noted previously?

The director of a study evaluating a statewide integrated services initiative described the problem this way:

> If we measure benefits for kids and families, what do we say contributed to it? Their individual services? The "program" as it existed in the 3 months they were involved? ... The program in one month is not the same as the next (a new partner joins the collaborative, changing the mix of staff, the number of services, the level of trust or conflict in the collaborative, etc.)... .There is so much going on out there, so much flux that even if we can document change or improvement, we have little idea what to attribute it to. We have one school in the study that is an integrated early childhood program site, site of a Blue Cross managed-care experiment, a new charter school, the recipient of a state restructuring grant, and in a neighborhood that is the recipient of family preservation funds. If we measure improvement in health indicators for the children, is it the early childhood program or the managed-care experiment? If there are educational benefits, is it the peer tutoring program in the early childhood program or the restructuring grant or the "charter schoolness"? (M. Wagner, personal communication, September 1994)

Researchers who study complex social interventions are used to this problem in one sense: They commonly acknowledge that many outcomes worth studying have multiple causes. But the number and elusiveness of the relevant independent variables make this situation especially challenging. Add to the attribution task the difficulty of explaining why certain children fail at school, or why disenfranchised families experience multiple problems with which they are not able to cope (e.g., see Dym, 1988, regarding ecological views of the family; Knapp & Woolverton, 1995, for a discussion of the role of social class in schooling; or Doner-Briar & Lawson, 1994, regarding "root causes" of children's problems in learning and development). Clearly, the availability or nature of human services—attempts at ameliorating school failure or families' problems—is only one contributor to such predicaments. Pervasive social conditions place individuals and groups in a disenfranchised position; human services have only a limited capacity to address questions of social position.

In other words, researchers studying comprehensive, collaborative services face a familiar challenge, that of constructing conceptual maps that link one thing to another. But as they do so, they know the many influences that might bring about results may not "stand still" long enough to permit confident claims about a particular initiative.

Studying Sensitive Processes and Outcomes: How Do We See What's Really Happening?

A final research issue has to do with measurement, intrusion, and the relationship between researchers or evaluators and participants in a program under study. Many of the programmatic events that lie at the core of comprehensive, collaborative services are private matters—for example, between a social worker and a troubled family, or a member of the family and a substance abuse counselor—and as such are not readily open to inspection by someone gathering data. This creates a double problem for those who would study the integration of education and human services: Not only is the nature of service blocked from view, but also the connections among services. The issue is compounded by the potential addition of data gatherers to the cadre of professionals with whom troubled families must interact outside of the context of direct service provision; understandably, many consumers are unwilling to cooperate with research requests, as are the social workers, counselors, or others who work most closely with them.

Though not insurmountable, this matter makes an already difficult task more so, precisely because the target of integrated services is likely to find research an intrusion and because the nature of integration is likely to involve more than one hard-to-inspect service. The situation is among the ones that lead the call for a more participatory conception of research or evaluation (Weiss & Greene, 1992).

Ways to Address the Issues

If the preceding analysis captures essential problems in examining comprehensive, collaborative services, then how can researchers and evaluators proceed to address these issues? What forms of research and evaluation are likely to yield the most useful understandings regarding this broad class of interventions at this stage of public investment in the integration of services?

Resolving these issues in particular instances is too dependent on context, and there are too many such instances for a short article such as this to offer specific advice about research questions or study designs. But it is possible to characterize, at a more global level, attributes of research that appropriately take into account the matters just raised. It is also possible to suggest particular kinds of studies that are more and less likely to yield useful knowledge.

Desirable Attributes for Research and Evaluation on Comprehensive, Collaborative Services

To be most helpful in making sense of integrated services, studies need to be strongly conceptualized, descriptive, comparative, constructively skeptical, positioned from the bottom up, and collaborative (when appropriate).

Strongly conceptualized. The elusiveness of independent or dependent variables and the relationships among them are in part a matter of conceptualization. More than one kind of conceptual framework is relevant to comprehensive, collaborative services, and these need to be worked out with some care, both to clarify what is being studied and to illuminate assumptions on which programmatic initiatives are founded. At a minimum, researchers and evaluators need to make explicit—before and after they have collected data—the conceptual dimensions that underlie the initiatives under study. One useful framework for studying collaborations notes the following dimensions (Merrill et al., 1994):

- Composition of target groups.
- Service scope (e.g., education, health, social service, etc.).

- Location of services.
- Sponsorship and service-provider participation.
- Commitment of participating service organizations.
- Parental and community participation.

Many other ways of identifying dimensions are possible, and some promising ones have been suggested (see, for example, the references noted in discussing the elusive independent variable). The important thing is not that any one framework be selected by everyone who studies collaborative services, but rather that researchers clarify in conceptual terms what is being studied. In this way, research will begin to answer the all-important question: Of what is any instance of comprehensive, collaborative services a (conceptual) case?

But the conceptual work doesn't end there. Two further kinds of conceptual models operate within a given comprehensive, collaborative services initiative, and it is up to the researcher to make them explicit, and hence open to inspection, or to put more powerful conceptualizations in place of the ones held by participants. First, implicitly or explicitly, comprehensive, collaborative services efforts rest on assumptions about those whom integrated services are intended to serve and about the conditions that generate their need for service. Programs operating on a deficit model, for example, tend to locate the problem in the high-risk child and his or her family. There are good reasons to view such models as insufficient and unhelpful. More successful conceptualizations of the problem addressed by integrated services will consider the joint roles played by individual characteristics, family and community conditions, and the expectations or routines of serving institutions (see, for example, Dym, 1988, regarding ecological views of families; Richardson, Casanova, Placier, & Guilfoyle, 1989, regarding relational views of at-risk learners in schools).

Second, the program's "theory of action" (Patton, 1978) is involved. Given some conception of a problem to be addressed, program designers and implementors fashion an intervention strategy that directs effort at key points of leverage. To take a simple example: Consider a collaborative services arrangement that colocates a health worker and social worker in high schools. Their presence is intended to provide additional counsel to youth who are likely to become pregnant, contract sexually transmitted diseases, and engage in other destructive behaviors. The arrangement operates on the premise that the presence of these individuals will increase access to good advice and, when needed, treatment; in addition, the presence of these individuals, it is assumed, will stimulate referrals from teachers and others in the building. By spelling out this strategy, researchers (and participants) have the chance to examine the logic of intervention. Is it reasonable to presume that presence will increase access? Are there other mediating factors that influence whether the presumed relationship would hold? Ultimately, data can be gathered to test the assumptions on which this logic rests.

In constructing useful frameworks for looking at comprehensive, collaborative services, researchers will need to draw on the concepts of different social science disciplines.

forts and the sources of organizational behavior); professional work (to illuminate the presumptions of, and constraints on, professional roles); multicultural interaction (to make sense of the interface between high-risk clients and professionals, and among professionals with different professional cultures); power and influence (to make sense of the sharing of control over services); policy process (to consider the power and limits of programmatic efforts in context); human development (to attend to stages and conditions affecting normal and abnormal growth); family dynamics (to understand families as systems); and group process (to appreciate the evolution of collaborative groups).

Some provocative and helpful conceptual work has begun to appear in the literature. Treatments of the phenomena within overarching ecological frameworks (e.g., Dym, 1988; Mawhinney, 1993) provide a compelling account of how child development, family welfare, and family service interventions operate in community context. Discussions of professional and institutional norms (Mitchell & Scott, 1993) and the way these work at the individual level provide further theoretical grounds for understanding what supposedly collaborative professionals do and do not do. This work complements recent attempts to view the situation at an institutional level drawing on the constructs and tenets of the "new institutionalism"; see, for example, work by Crowson and Boyd (1994) that focuses on the deep-structure of norms, rules, routines, and administrative scripts operating within service-providing institutions.

Descriptive. If service integration can be so many different things, and the collaborating professional disciplines are still not sure what form(s) such integration should optimally take, then it makes sense to put a great deal of emphasis on the description of particular cases of comprehensive, collaborative services. Ideally, such descriptions should be guided by (and should inform) the strong conceptual frameworks called for previously; endless narrative and detail will not serve any useful purpose.

Qualitative "thick" descriptions (in the sense originally proposed by Geertz) are especially appropriate, though they are not the only kind of useful descriptive account. Qualitative techniques are especially helpful in illuminating what collaborative arrangements mean to participants, how such efforts differ from service-as-usual, and what the nature of collaboration is. The sensitivity of the research topic makes good qualitative description difficult in many instances, but there are ways to gain access to even the most difficult research situations. In this regard, nonevaluative research may make a greater contribution than explicitly evaluative studies, in which the stakes are higher and scrutiny by researchers can take on more negative meanings.

Careful descriptions are needed of at least the following: organizational arrangements; the interface between the consumer and service providers purporting to coordinate their efforts; the sharing of resources, ideas, and professional work; the experience of collaboration; and the extent and nature of change in the consumer's behavior, attitudes, or life circumstances.

Comparative. Given strong conceptualization (which permits cross-case comparison) and good description, the

of initiatives currently under way. Such studies are un-
likely to offer the kind of comparisons presumed by exper-
imental research or planned variation studies, but they can
be instructive regarding the range of conditions that sup-
port or frustrate collaborative work, as well as the possible
variations on the theme that make up promising practice.
Whenever contrasting cases can be chosen with particular
variations on key dimensions in mind, studies can offer
more powerful comparative insights.

Constructively skeptical. Research needs to help audiences
see through the hype, prescription, and program rhetoric,
while remaining sympathetic with overall programmatic
aims. A constructively skeptical stance is thus highly ap-
propriate at this stage in our understanding of integrated
services. Too few discussions in this arena acknowledge, as
does Golden (1991), that

> it is not obvious that collaboration always has good rather
> than bad effects on services for families and children.
> Collaboration might lead agencies to carry out their dif-
> ferentiated, precollaboration mission less well . . . Collab-
> oration might lead a program that has been effective on
> Schorr's criteria to become less so, if it collaborates with a
> more rigid bureaucratic program and its mission and cul-
> ture are diluted. For example, staff in a teenage program
> I visited for recent research on welfare and children's ser-
> vices were very nervous about the emphasis on rules that
> their (ultimately unsuccessful) collaboration with a local
> welfare agency was, they thought, imposing on their ser-
> vices. (p. 85)

Skepticism is called for regarding many kinds of claims
made on behalf of comprehensive, collaborative services,
for example, regarding cost savings, mutually reinforcing
effects, attribution to programmatic efforts, stability of col-
laborative arrangements, incentives for collaboration, and
changes in approach to service, to mention only a few of
the candidates. But even at the stage of conceptualizing
studies, researchers and evaluators would do well to con-
sider, as do some scholars (see Crowson & Boyd, 1994), the
possibility of organizational self-protection in the face of
collaborative pressures or the chance that professionals
work in a less intentional and purposive way than collab-
oration theories seem to imply.

There are obvious complications in keeping a skeptical
stance constructive in evaluative situations. Program op-
ponents are likely to pounce on any negative evidence as
ammunition in future debates about program continua-
tion, whereas program promoters will wax defensive at the
hint of criticism. In addition, the unrealistically high ex-
pectations and short attention span of most policy commu-
nities makes skepticism problematic. At the same time,
there is no great virtue in prolonging the lifetime of inter-
ventions that rest on shaky logic and little evidence.

Skepticism may not be constructive if research ques-
tions, designs, and measures are inappropriately applied
to the program in question, for example, by prematurely
searching for impacts at a relatively early stage in the de-
velopment of a complex program (regardless of pressures
from certain stakeholders to do so). One extended case
study of evaluation in the integrated services arena (Stake,
1986) offers a cautionary tale in this regard: In that in-
stance, the single-minded focus of the evaluation study on

sites of the Cities-in-Schools program generated negative
evidence about the program in a way that may not have
represented subtle benefits to the youths involved. To
guard against this possibility, researchers and evaluators
need to maintain a constructively skeptical stance regard-
ing their own capacity to ask the right questions, employ
sufficiently sensitive measures, and interpret what they
find appropriately.

Positioned from the bottom up. Collaborative services are
ultimately integrated as they converge on individuals,
groups, or populations they serve. Research and evalua-
tion that trace backward from the experiences, behavior,
perceptions, and status of service recipients will be more
likely to show if and how the integration occurred and
whether it achieved valuable ends. Such studies focus on
the consumer and the consumption of services, but need
not be restricted to activity at the street level. Many useful
studies of interagency dynamics, the orchestration of re-
sources, and other features at higher levels in the system
are possible from this vantage point, but by anchoring the
investigation to the consumer, researchers are less likely to
be distracted by studying means and thereby lose sight of
ends (following the notion advanced by Golden, 1991).

Approaching the research problem from the bottom up
needs not be solely concerned with the consumer's-eye
view of comprehensive, collaborative services. For exam-
ple, in sketching out its evaluation strategy for the New
Futures initiative, the Annie E. Casey Foundation envi-
sioned three components to its evaluation, one of which
would feature individual qualitative profiles of youth un-
dergoing change, another assembling quantitative data re-
lated to aggregate impact on youth, and the third
examining institutional effects (Center for the Study of So-
cial Policy, 1987). The important point is that such designs
prominently feature the nature and meaning of service and
system benefits at the ground level, and that the evidence
for such benefits reflects the specifics of particular cases in
their local settings.

Collaborative (when appropriate). Because it is essential to
engage divergent perspectives in studies of comprehen-
sive, collaborative services, it is tempting to expect
research to be itself collaboratively designed and imple-
mented, either by researchers of different disciplines or by
researchers and participants (service providers, con-
sumers) in the collaborative arrangements under study.
Discussions that call for a "partnership" between evalua-
tors and program people contribute to the call for more col-
laborative research on collaborative services (e.g., Weiss &
Greene, 1992).

There are obvious advantages of putting heads together
in such a way. Collaborative approaches to research can
help to draw attention to conceptual elements that one re-
search tradition pays close attention to while others do not,
identify the assumptions and perspectives with which dif-
ferent types of professions approach collaborative work,
develop appropriate measures, and find multiple mean-
ings in results. Collaborative research that encourages dia-
logue with service recipients regarding research goals,
approaches, or findings can probably help researchers stay
tuned to consumers' perspectives (which are often forgot-
ten as professionals try to develop better ways of serving
clients).

applicable to research than to service delivery. Good collaboration is difficult and time-consuming (a challenge to research that must be done on a tight budget), requires a sharing of control (while the logic of many research designs calls for tight control), and may involve unproductive wrangling over paradigms.

Some Promising Kinds of Studies

What follows is not an exhaustive list, but rather several examples of kinds of investigations (or components of large-scale investigations) that embody the attributes discussed previously and are likely to yield more insight at this stage in the understanding of integrated services. Five kinds of studies spring to mind: (a) profiles of individual participation and change, (b) multiple-case, thick descriptions of collaborative service arrangements at the point of service delivery, (c) analyses of cost in both quantitative and qualitative terms, (d) single-subject (and single-system) time-series research to demonstrate impact (at both the individual and organizational level), and (e) investigations of exemplary and typical practice. A sixth type of study—analyses based on management information systems that track comprehensive, collaborative services—has promise, as well, though there are difficulties in developing and maintaining such systems.

Profiles of individual participation and change. This kind of study answers the questions: How does the individual child or family participate in collaborative services? What does participation involve? In what ways do these individual participants change? By treating the individual's participation and experience as the primary unit of analysis, investigations of this sort bypass the problem of treating the whole program as a meaningful treatment. Qualitative and quantitative data can both be part of the profile. Sampling of individuals to study (and gaining access to these individuals) becomes a major issue; depending on the size and complexity of the sample, such investigations could present a picture of comprehensive, collaborative services across the range of people within a community, or even multiple communities.

Multiple-case, thick descriptions of collaborative services arrangements at the point of service delivery. This sort of study represents a kind of programmatic counterpart to the preceding one; rather than focusing primarily on the consumer, it examines the nature of professional work and the contexts in which this happens. This kind of investigation answers the questions: What do professionals do to integrate their efforts at the point of service delivery? What forces and conditions impinge on their attempts to address social needs through collaborative effort? The rationale for this kind of study presumes that collaborative work involves subtle shifts in professionals' conceptions of their craft, work routines, and approach to particular consumers; qualitative approaches are especially suitable for capturing such phenomena. This kind of research is needed to characterize, both conceptually and empirically, the elusive independent variable in comprehensive, collaborative arrangements. Comparative attempts to describe and contrast different kinds of integrated arrangements, chosen to vary along key dimensions, would be particularly helpful.

complex in collaborative services arrangements, and deserves constructively skeptical research to understand whether the costs of these services are simply prohibitive for all but a few children and families. The deceptively simple questions to be answered are: What do comprehensive, collaborative services arrangements cost the human service system, the public, and the consumer? Are these costs "worth it," in terms of definable benefits or effects? Because most such arrangements are new and experimental, there are major start-up costs, as professionals develop new roles and working routines. These costs need to be disentangled from ongoing costs in time, energy, complexity, burden on service providers or consumers, foregone opportunities for less labor-intensive ways of addressing human needs, and long-range failure to address the needs of high-risk children and families. Costs need to be examined in perspective, with "full-service" arrangements contrasted with less comprehensive ones. Reliable numbers will be difficult to obtain but important to pursue, accompanied by some attempt to characterize qualitatively the nature of "cost." To date there is little work that examines costs responsibly, though some have begun to argue the importance of doing so and to identify the key considerations involved (e.g., White, 1988).

Single-subject (and single-system) time-series research to demonstrate impact on individuals or service systems. It is essential to understand impact on children, families, and systems in context. The bottom-line question—What do comprehensive, collaborative services initiatives do for children, families, and human services systems?—must be answered (the constructive skeptic does not assume that such arrangements accomplish what they purport to do). But getting at this matter through group comparative designs, the most common approach to ascertaining impact, may be fruitless when "treatments" are so individualized, meaningful control groups hard to construct, and attribution of result to cause so complicated to trace. In such instances, the individual unit's behavior over time may well be its own best control, as argued by the tradition of single-subject time-series research. Such designs call for some baseline of repeated measures over time prior to participation in collaborative services, a fully described treatment, and a follow-up pattern of repeated measures that can demonstrate change in trajectory associated with exposure to treatment. Though complicated to apply in its conventional form (e.g., as practiced in special education research) to many comprehensive, collaborative services, this design can be adapted to the purposes of studying such initiatives. An analogous design logic pertains at the organizational level to get at the impacts of systems reforms (see Knapp, 1979).

Investigations of exemplary and typical practice. In this kind of study, the researcher or evaluator works backward from instances of presumably effective or "average" practice to explanations for the apparent success. This study answers the questions: What do apparently successful arrangements for comprehensive, collaborative services accomplish and how do they accomplish it? What forces and conditions enable these services to do what they do? Presuming that through some combination of reputation and rough outcome indicators, one can identify instances of

good things for children and families, careful study of these instances using either qualitative or quantitative means (but ideally with some kind of quantitative outcome indicators) should be especially instructive. By including sites that represent more typical practices in the scope of the study, the researcher can cast the accomplishments of exemplary sites (and the conditions that support these accomplishments) in perspective.

Analyses of data from management information systems that routinely track consumers' access to, and use of, multiple services. A sixth kind of research also has promise though it faces significant obstacles in practice. Researchers and program designers alike have noted the importance, as well as the difficulties, of getting succinct data that track how people interact with services, especially where these services are separately housed and governed by restrictions on the flow of information. Experiments have been undertaken to put information systems in place that gather and record the presumably comprehensive provision of service (see Family Impact Seminar, 1993a, for a summary of the work in this area). In theory, such tools may be useful for answering questions such as: What services have X, Y, and Z used, when? What did service providers do in attempting to meet the needs of children A, B, or C and their families? What changes in indicators are associated with which patterns of service use? Such systems are only as good as the data put into them, however, and it is not easy to ensure that high-quality data are entered *and* updated on a regular basis. Often, more data are collected than are needed for researchers' or any one user's purposes, and this can quickly feel burdensome to participants at the "street-level," especially if imposed from the top down. Systems that provide service providers at the operating level with information they want and can use are more likely to get better quality data; when such information corresponds with what researchers and evaluators need to know *and* where their access to such information is politically and organizationally feasible, this device has considerable potential.[2]

None of these six types of studies constitutes a comprehensive investigation analogous to what is called for in most of the major evaluative studies now under way. Obviously these kinds of studies and others like them can be viewed as components of a large investigation. The payoff to such elaborate studies is not always assured, but for large system initiatives and elaborate demonstration projects, more complex investigations are hard to avoid.

Types of Studies That May Be Less Useful

Once again, with no attempt at completeness, some approaches to research seem less likely to yield useful insights, given what we now understand about comprehensive, collaborative services for children and families. Competently executed, these forms of research may contribute to an understanding of collaborative services, but there are serious questions regarding the degree of payoff. One of these forms—group-comparative experimental studies—has wide popularity among researchers and evaluators, as well as their audiences.

Group-comparative experimental studies contrasting recipients with nonrecipients. The obvious advantages of such designs include the compelling logic and apparent rigor of experimental contrasts (where the assumptions on which

this form of knowledge generation among many audiences. But the drawbacks are many, as some discussions in the literature on comprehensive, collaborative services have suggested (see Bruner, 1994; Family Impact Seminar 1993a; Weiss & Greene, 1992). The burden of proof is on the experimentally inclined researcher or evaluator to demonstrate that key assumptions are viable (e.g., Is there an identifiable and uniform treatment? Are recipients and nonrecipients sufficiently comparable?). As has been learned from years of social experiments, including studies of programs that are much more easily specified and applied to groups (e.g., academic programs in school settings), group comparative studies are harder to realize in practice than on paper, and the logic often breaks down. A great danger exists that the requirements of the research design will force evaluable situations to be constructed that compromise or limit what comprehensive, collaborative services are attempting to do. A similar danger is that the press for experimental results will force a premature search for evidence of widespread impact—just the thing that new and ambitious programs are least able to provide, regardless of their merits.

Several other categories of research and evaluative activity are less commonly called for, but appear at first glance to offer insight into the implementation and impact of comprehensive, collaborative services. On closer examination, these approaches may be less useful.

Factor analytic studies that search for empirical clustering of large numbers of programmatic and nonprogrammatic variables. Given the many pieces of the independent variable, researchers may believe that correlational investigations have the most to offer. Unless they are very strongly conceptualized, such attempts risk identifying statistical clusters that are nearly impossible to name meaningfully—precisely the kind of results that will do little to advance an understanding of collaborative services. The circumstances surrounding most collaborations and the design of these services invite too many ways for variables to be spuriously correlated.

Meta-analyses of service integration studies. As large numbers of studies emerge, it is tempting to undertake an apparently rigorous way of aggregating what has been learned from all of them. Meta-analyses may appear to offer this possibility, and some have argued for them with regard to family support programs (Hauser-Cram, 1988). But in the absence of a commonly defined independent variable (or even comparable outcome measures), this technique seems either premature or altogether unsuited to this portion of the social intervention research terrain (Bangert-Drowns, 1986). Other approaches to aggregating results (e.g., Schorr & Both, 1991; Wang, Haertel, & Walberg, 1994) appear to offer more at this point, even though their reliance on reviewers' judgments appear to weaken the conclusions that can be drawn. In addition, meta-analysis cannot be applied to qualitative findings, and these, too, beg for some kind of meaningful aggregation as they accumulate.

Conclusions: A Call for Appropriate Research on and Evaluation of Comprehensive, Collaborative Services

In one sense, there is little need to call for studies of comprehensive, collaborative services for children and fami-

...—miging researchers and evaluators in droves. Numerous studies are under way and more are on the drawing board. Big investments beget big evaluation studies, and many are in process at this writing (see Behrman, 1992, and Family Impact Seminar, 1993b, for a list of evaluative activities now in progress).

The plea with which this article concludes is for researchers and evaluators, and those who sponsor studies (or demand they be done), to consider what is *appropriate* to ask and answer at the current stage of development, experimentation, and understanding. These are generally not mature programs that have developed a relatively stable modus operandi; in most cases we are witness to (and participants in) rapidly evolving experimentation within turbulent reform contexts. We are observing a class of intervention that is hard to name, no less describe. And we have yet to answer a critical question: What do the many instances of collaboration represent conceptually?

In this context, it is debatable what we should be studying—that is, what makes a study appropriate or not. To do so lies beyond the scope of this article, and there are other efforts under way that are attempting to suggest a more specific research agenda in this area.[3] Nonetheless, several observations can be made. In such circumstances there are compelling reasons to engage in research and evaluation of many kinds. The early program rhetoric, filled with visions and promises, may be taken as gospel (and already is in some quarters) long before we know whether anyone is helped or whether we can afford it; just as likely, impatient audiences will lose faith in collaborations because no evidence appears of instant impact. There is a sufficient number and variety of investments in comprehensive, collaborative services initiatives to afford numerous opportunities for learning and various forms of "natural experiments." And the children and families who are the recipients of integrated services are too needy and too numerous to ignore.

But there are big dangers in overinvesting in unproductive kinds of research. For one thing, we may end up studying only what we know how to study, and not engaging in the kind of methodological learning that new forms of social intervention require. For another, we may prematurely declare the experimentation a failure, neglecting to be clear about what failed. Or, we may proclaim and describe programmatic victory, only to find that multiply served children continue to fail in school or their families continue to confront health and social challenges with which they cannot cope.

There are more subtle dangers as well, some of them arising when we follow our own advice too well. Eager to detect combinations of services that are more potent, we may neglect to note the way these services categorize and demean the people they serve—if it is not careful, research on comprehensive, collaborative services may help reify a new deficit model of the "truly disadvantaged" or "superneedy." Or, mindful of the fuzziness regarding the independent variable, we may unwittingly become preoccupied with the intricacies of collaboration or the different forms of interprofessional work and lose sight of the ends (e.g., children's health, education, and welfare) for which this is only one means. Finally, in an attempt to engage all

...of comprehensive, collaborative services with the result that service providers and the consumers they are trying to help feel besieged. If these things come to pass, the segment of society for whom comprehensive, collaborative services are being devised will not be well served by the research and evaluation community. We can and must do better.

Notes

1I am indebted to Mary Wagner of SRI International for this way of describing one of the essential dilemmas regarding the independent variable in research on comprehensive, collaborative services.

2This point is based on an observation made by Bill Morrill of Mathtec, Inc., who has carried out informal analyses contrasting "top-down" information systems in integrated services arrangements with other information systems that are more directly responsive to practitioners' needs at the service delivery level.

3For example, the results of the AERA/OERI conference referenced earlier in the notes are currently being assembled in monograph form, as one outline of a comprehensive research agenda related to collaborative services.

References

Agranoff, R. (1991). Human services integration: Past and present challenges in public administration. *Public Administration Review*, 51(6), 533-542.

American Public Welfare Association. (1992). *Status report on JOBS case management practices.* Washington, DC: Author.

Bangert-Drowns, P. L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin, 39,* 388-399.

Behrman, R. E. (Ed.). (1992). School-linked services. *The Future of Children, 2*(1).

Bruner, C. (1994). A framework for measuring the potential of comprehensive services strategies. In N. Young, S. Gardner, S. Colley, L. Schorr, & C. Bruner (Authors), *Making a difference: Moving to outcome-based accountability for comprehensive service reforms* (pp. 29-40). Falls Church, VA: National Center for Services Integration.

Center for the Study of Social Policy. (1987). *New Futures evaluation strategy.* Unpublished document. Washington, DC: Author.

Chang, H. N., Gardner, S. L., Watahara, A., Brown, C. G., & Robles, R. (1991). *Fighting fragmentation: Collaborative efforts to serve children and families in California's counties.* San Francisco: California Tomorrow and Children and Youth Policy Project.

Chaudry, A., Maurer, K. E, Oshinsky, C. J., & Mackie, J. (1992). *Service integration: An annotated bibliography.* Falls Church, VA: Mathtec Inc.

Chynoweth, J. K., & Dyer, B. (1991). *Strengthening families: A guide for state policymaking.* Washington, DC: Council of Governors' Policy Advisors.

Coulton, C. J. (1992). *Framework for evaluating family centers.* Cleveland, OH: Center for Urban Poverty and Social Change, Case Western Reserve University.

Cronbach, L. J., & Associates (1980). *Towards the reform of program evaluation.* San Francisco: Jossey-Bass.

Crowson, R. L., & Boyd, W. L. (1993). Coordinated services for children: Designing arks for storms and seas unknown. *American Journal of Education, 101,* 140-179.

Crowson, R. L., & Boyd, W. L. (1994, September). *Achieving coordinated school-linked services: Facilitating utilization of the emerging knowledge base.* Paper presented at the Invitational Conference on Improving Urban Schools: Better Strategies for the Dissemination and Utilization of Knowledge, Arlington, VA.

...... Euismington, IX.: The Finance Project.

Dryfoos, J. (1994). *Full-service schools*. San Francisco: Jossey-Bass.

Dym, B. (1988). Ecological perspectives on change in families. In H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs* (pp. 477–495). New York: A. de Gruyter.

Family Impact Seminar. (1993a). *Data integration and evaluation: Essential components of family-centered systems reform—Background briefing report*. Washington, DC: American Association of Marital and Family Therapy (AAMFT) Research and Education Foundation.

Family Impact Seminar. (1993b). *Data integration and evaluation: Essential components of family-centered systems reform—Highlights of seminar meeting*. Washington, DC: AAMFT Research and Education Foundation.

Farrow, F., & Joe, T. (1992). Financing school-based, integrated services. *The Future of Children*, 2(1), 56–67.

Garvin, J. R., & Young, A. H. (1993). Resource issues: A case study from New Orleans. In L. Adler & S. Gardner (Eds.), *The politics of linking schools and social services* (pp. 93–106). London: The Falmer Press.

General Accounting Office. (1992). *Integrating human services*. Washington, DC: Author.

Golden, O. (1991). Collaboration as a means not an end: Serving disadvantaged families and children. In L. B. Schorr, D. Both, & C. Copple (Eds.), *Effective services for young children—Report of a workshop* (pp. 84–104). Washington, DC: National Academy Press.

Gomby, D. (1992). *School-linked services programs: Lessons from the past*. Unpublished manuscript. Los Altos, CA: Center for the Future of Children/the David and Lucille Packard Foundation.

Gomby, D. S., & Larson, C. S. (1992). Evaluation of school-linked services. *The Future of Children*, 2(1), 68–84.

Gray, S. T. (1993). *A vision of evaluation—Report of learnings from the Independent Sector's work on evaluation*. Washington, DC: Independent Sector.

Hauser-Cram, P. (1988). The possibilities and limitations of meta-analysis for understanding family program impact. In H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs* (pp. 445–460). New York: A. de Gruyter.

Hooper-Briar, K., & Lawson, H. (1994). *Serving children, youth, & families through interprofessional collaboration and services integration: A framework for action*. Oxford, OH: The Institute for Educational Renewal, Miami University.

James, W. H., Smith, A., & Mann, R. (1991). Educating homeless children: Interprofessional case management. *Childhood Education*, 67(5), 305–308.

Kagan, S. L. (1991). *United we stand: Collaboration for child care and early education services*. New York: Teachers College Press.

Kagan, S. L., Rivera, A. M., & Lamb-Parker, F. (1990). *Collaborations in action: Reshaping services for young children and their families*. New Haven, CT: Bush Center for Child Development and Social Policy.

Kirst, M. (1994, September). *School-linked services: An appraisal and future directions*. Paper presented at the Invitational Working Conference on Comprehensive School-linked Services for Children and Families, Leesburg, VA.

Knapp, M. S. (1979). Applying time series strategies: An underutilized solution. In L-E. Datta & R. Perloff (Eds.), *Improving evaluations* (pp. 101–130). Beverly Hills, CA: Sage.

Knapp, M. S., & Woolverton, S. (1995). Social class and schooling. In J. Banks & C. M. Banks (Eds.), *Handbook of research on multicultural education* (pp. 548–569). New York: Macmillan.

Levy, J. E., Kagan, S. L., & Copple, C. (1992). *Are we ready? Collaboration to support young children and their families*. Washington, DC: Council of Chief State School Officers.

Marzke, C. H., Chimerine, C. B., Morrill, W. A., & Marks, E. L. (1992). *Service integration programs in community settings*. Falls Church, VA: Mathtec, Inc.

Mawhinney, H. (1993). Discovering shared values: Ecological models to support interagency collaboration. In L. Adler & S. Gardner (Eds.), *The politics of linking schools and social services* (pp. 33–52). London: The Falmer Press.

Mitchell, D. E., & Scott, L. D. (1993). Professional and institutional perspectives on interagency collaboration. In L. Adler & S. Gardner (Eds.), *The politics of linking schools and social services* (pp. 93–106). London: The Falmer Press.

Melaville, A. I., Blank, M. J., & Asayesh, G. (1993). *Together we can: A guide to crafting community-based, family-centered strategies for integrating education and human services*. Washington, DC: Office of Ed-

ucational Research and Improvement, U. S. Department of Education.

Morrill, W. A., Reisner, E. R., Chimerine, C. B., & Marks, E. L. (1991). *Collaborations that integrate services for children and families*. Washington, DC: U. S. Department of Education.

National Association of State Boards of Education. (1991). *Caring communities: Supporting young children and families*. Alexandria, VA: Author.

National Commission to Prevent Infant Mortality. (1991). *One-stop shopping: The road to healthy mothers and infants*. Washington DC: Author.

New Beginnings Team. (1990). *New beginnings: A feasibility study of integrated services for children and families*. San Diego: San Diego City Schools.

Nucci, L. P., & Smylie, M. A. (1991). University-community partnerships: Addressing problems of children and youth through institutional collaboration. *Metropolitan Universities*, 2(1), 63–91.

Patton, M. Q. (1978). *Utilization-focused evaluation*. Beverly Hills, CA: Sage.

Philliber Research Associates. (1994). *An evaluation of the Caring Communities Program at Walbridge Elementary School*. Accord, NY: Author.

Richardson, V., Casanova, U., Placier, L., & Guilfoyle, K. (1989). *School children at-risk*. London: The Falmer Press.

Robison, E. (1993). *An interim evaluation report concerning a collaboration between the Children's Aid Society and the New York City Board of Education Community School District 6 and the I. S. 218 Salome Urena de Henriquez School*. New York: The Children's Aid Society.

Rossi, P. (1992). Strategies for evaluation. *Children and Youth Services Review*, 14(1&2), 167–191.

Schorr, L. B. (1988). *Within our reach: Breaking the cycle of disadvantage*. New York: Anchor.

Schorr, L. B., & Both, D. (1991). Attributes of effective services for young children: A brief survey of current knowledge and its implications for program and policy development. In L. B. Schorr, D. Both, & C. Copple (Eds.), *Effective services for young children—Report of a workshop* (pp. 23–47). Washington, DC: National Academy Press.

Schorr, L. B. (1994). The case for shifting to outcome-based accountability. In N. Young, S. Gardner, S. Coley, L. Schorr, & C. Bruner (Eds.), *Making a difference: Moving to outcome-based accountability for comprehensive service reforms* (pp. 29–40). Falls Church, VA: National Center for Service Integration.

Scriven, M. (1974). Goal-free evaluation. In E. House (Ed.), *School evaluation: The politics and process*. Berkeley, CA: McCutchan Publishers.

Stake, R. (1986). *Quieting reform: Social science and social action in an urban youth program*. Urbana, IL: University of Illinois Press.

State Reorganization Commission. (1989). *Services integration pilot projects: An evaluative report from Arizona, Florida, Maine, Oklahoma, South Carolina*. Columbia, SC: Author.

Wagner, M., Golan, S., Shaver, D., Newman, L., Wechsler, M., & Kelley, F. (1994). *A healthy start for California's children and families: Early findings from a statewide evaluation of school-linked services*. Menlo Park, CA: SRI International.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1994, September). *School-linked services: A research synthesis*. Paper presented at the Invitational Working Conference on Comprehensive School-Linked Services for Children and Families, Leesburg, VA.

Wehlage, G., Smith, G., & Lipman, P. (1992). Restructuring urban schools: The New Futures experience. *American Educational Research Journal*, 29, 51–93.

Weiss, H. B., & Greene, J. G. (1992). An empowerment partnership for family support and education programs and evaluation. *Family Science Review*, 5(1&2), 131–148.

White, K. R. (1988). Cost analyses in family support programs. In H. B. Weiss & F. H. Jacobs (Eds.), *Evaluating family programs* (pp. 429–443). New York: A. de Gruyter.

White, W. A. (1993). California's state partnership for school-linked services. In L. Adler & S. Gardner (Eds.), *The politics of linking schools and social services* (pp. 171–178). London: The Falmer Press.

Wilson, W., Karasoff, P., & Nolan, B. (1993). The integrated services specialist: An interdisciplinary preservice training model. In L. Adler & S. Gardner (Eds.), *The politics of linking schools and social services* (pp. 161–164). London: The Falmer Press.

MlkeJnto^I^L

N?

watered down

specific italization very good

working reports rather than reports

1. expectation that people know what this is
about — tell about the study and definitions in
the first

2. As progrep take out the insert and put iin page
1. Otherwise don't know what

3. 11-13 and others. what is many in-service?
put in context

4. the recommendations are not sharp enough..
Except for 4.

5. does it reflect the findings as I see them ((tb
Q) ?

6. started as teachers,. not satisfied with money
but willing to live,. want to make career.

MAYBE RE-FOCUSED ON TURNOVER ((PART OF THE PROBLEM
OR OF THE SOLUTION?

BECAUSE OF THE STRUCTURE OF SOCIETY BUILD SYSTEM
THAT CAN LIVE WITH TURNOVER ((ADAPT IT).

Recommendation of meeting more: what does it
mean?

Images of meeting.

█████ FOR POLICY RESEARCH IN EDUCATION
Rutgers, The State University of New Jersey
University of Wisconsin-Madison • Harvard University
Stanford University • University of Michigan

THE POLICY CENTER • THE FINANCE CENTER

June 7, 1995

Annette,

The enclosed paper will be presented at the Jewish ed. research conference on June 11. I'd be grateful for any comments you, Mike, or Seymour might have.

Would July 5 be a good target date for comments?

In keeping with our policy, I am asking for approval of our advisory subcommittee on publications (now consisting of you and Mike) before submitting the paper for external publication. If you can let me know about what in July, that would be fine or if you want to wait for another draft of the paper, that would be fine too.

Adam

UNIVERSITY of
WISCONSIN-MADISON

Wisconsin Center for Education Research • University of Wisconsin-Madison
1025 West Johnson Street, Room 753, Madison, WI 53706-1796
608-263-4260 • Fax: 608-263-5448

BACKGROUND AND TRAINING OF TEACHERS IN JEWISH SCHOOLS: CURRENT STATUS AND LEVERS FOR CHANGE

Adam Gamoran
Ellen Goldring
Bill Robinson
Roberta Louis Goodman
Julie Tammivaara

Council for Initiatives in Jewish Education

BACKGROUND AND TRAINING OF TEACHERS IN JEWISH SCHOOLS: CURRENT STATUS AND LEVERS FOR CHANGE

ABSTRACT

A survey of teachers in day schools, supplementary schools, and pre-schools in three communities shows that only 19% of teachers have professional training in both Jewish content areas and in the field of education. Despite incomplete professional backgrounds, teachers in Jewish schools engage in relatively few professional development activities: pre-school teachers reported attending an average of 6.2 workshops over a two-year period, while supplementary teachers attended an average of 4.4 and day school teachers attended 3.8 workshops over the two year period. What can be done to enhance and expand professional growth activities for teachers in Jewish schools? This paper examines three possible "levers" for changing standards for professional growth: state licensing requirements for pre-schools, state requirements for continuing education among professionally-trained teachers, and federation-led standards for training of supplementary teachers. Results indicate that pre-school teachers in state-licensed pre-schools and supplementary school teachers who were paid for meeting a professional growth standard reported that they were required to attend more in-service workshops, compared to other teachers who were not faced with these standards.

BACKGROUND AND TRAINING OF TEACHERS IN JEWISH SCHOOLS: CURRENT STATUS AND LEVERS FOR CHANGE

"A new two-year study of Jewish educators in three North American communities offers a striking assessment of teachers' preparation and professional development in day schools, supplementary schools, and pre-schools."   — CIJE Policy Brief

Recent research at the Council for Initiatives in Jewish Education (CIJE) shows that only a small proportion of teachers in Jewish schools in three communities are formally prepared in both Jewish studies and in the field of education. This paper presents and extends selected findings from the CIJE research. In addition, it moves beyond findings that have been made public thus far by exploring mechanisms that may raise standards for in-service teacher training in Jewish schools. These levers include state licensing requirements for pre-schools, state requirements for continuing education among professionally-trained teachers, and federation-led standards for training of supplementary teachers.

Background

In 1991 the Commission on Jewish Education in North America released *A Time to Act*, a report on the status and prospects of Jewish education. The report concluded that building the profession of Jewish education (along with mobilizing community support for education) is essential for the improvement of teaching and learning in Jewish schools. This conclusion rested on the best available assessment of the field at that time: "well-trained and dedicated educators are needed for every area of Jewish education....to motivate and engage children and their parents [and] to create the necessary educational materials and methods" (1991, p.49). In response, the Commission created the CIJE, whose mandate includes

establishing three Lead Communities in North America, and working with these communities to serve as demonstration sites for improving Jewish education.

What is the current state of the profession of Jewish education in these communities? What mechanisms are available to improve it, and how will we know whether improvement in the profession training of teachers fosters better teaching and learning? These questions cannot be addressed fully —in particular, no data are available on the links between training, teaching, and learning —but this paper begins to address the issues by examining the current professional backgrounds of teachers in Jewish schools as well as considering potential levers for increasing teacher's professional development activities.

Professional Preparation and Development in Jewish Education

Modern conceptions of teaching emphasize formal, specialized preparation (e.g., Sedlak, 1987). This preparation typically involves training in both pedagogy and subject matter, as well as in the links between the two (Shulman, 1987). Moreover, teachers are expected to maintain their subject matter and pedagogical skills through continuous professional development. As Aron (1990, p. 6) explained, teachers need "to keep pace with new developments in their field. The knowledge base of teaching has grown and changed....Therefore, it would be imperative for veteran teachers to have mastery of this new body of information, skills, and techniques." In Jewish education, where many teachers lack formal preparation for their work, professional development is not a matter of keeping pace, but of getting up to speed.

In public education, the profession of teaching is regulated by certification at the state level. Although exceptions are made, generally states require formal preparation in the field

of education, including study of content knowledge and pedagogy, for teacher licensing. In addition, many states require a set amount of professional development over a fixed period of time for the renewal of one's teaching license. In Jewish schools, because of a shortage of certified teachers, it is often not possible to hire only teachers who are formally prepared in their fields. Hence, the question of professional development becomes especially salient.

What circumstances lead to more in-service workshops for teachers? On the one hand, schools with teachers who are more professionally oriented may be able to place greater demands for professional growth of teachers. A staff that is trained for Jewish education, holding degrees in education and in Jewish content areas, and viewing Jewish education as a career, may create the kind of community that allows professional norms to flourish, including more extensive professional development.

On the other hand, even without a highly professional staff, there may be conditions that can increase the amount of professional development activity. In this paper we examine three possible mechanisms, or levers for change, which may lead to more in-service workshops. The particular mechanisms we explore were not chosen on theoretical grounds; rather, they are the mechanisms we encountered in a study of three Jewish communities. We found that communities and schools varied in their policies and in the conditions associated with policies about staff development. This type of "natural experiment" can yield important information about the prospects for increasing professional growth activities in Jewish education.

The possible levers we encountered were as follows:

(1) **State certification for pre-schools.** Most of the pre-schools in our study are licensed or certified by the state, and certification requires a set amount of staff development for teachers. For example, in one state teachers had to take 18 hours of in-service per year for a school to maintain its certification. Other states had different requirements but all demanded some level of in-service among teachers to maintain certification. Consequently, one may expect to find higher rates of in-service training among pre-school teachers compared to other teachers, and we reported this pattern in our earlier work (Gamoran et al., 1994). Here we test this interpretation by comparing in-service training in the pre-schools that are not certified to those that are. We expect to find higher rates of in-service required in state-certified pre-schools.

(2) **State in-service requirements for re-licensing.** The communities we studied are located in three different states. One state requires that licensed K-12 teachers engage in 180 hours of workshop training over a five-year period in order to be re-licensed. Another state requires 100 hours of in-service over the same period. The third state has no such mandate. Are Judaica teachers in Jewish schools responsive to these mandates? Even if teachers on average are not affected by these requirements, one may expect that teachers who are professionally trained would keep up with licensing requirements.

(3) **Federation incentives for supplementary teachers.** In one community, the federation provides an extra incentive to encourage in-service attendance among supplementary school teachers. Teachers who attend at least 4 workshops in a year (3 for those who teach only on Sundays) receive a special stipend. In addition,

supplementary schools in which at least three-quarters of the teachers meet the in-service standards receive funds from the federation. Thus, the incentive program encourages not just individual but school-wide professional growth. If these incentives are effective, we would expect to find that supplementary school teachers reported more workshops in this community than in the other two.

Data and Methods

Data from this paper are drawn from two data sources: A survey of teachers, and intensive interviews with a sample of teachers and other educators. The surveys and interviews were conducted in the three CIJE Lead Communities: Atlanta, Baltimore, and Milwaukee, in 1992 and 1993. All Judaica teachers in day schools, supplementary schools, and pre-schools were asked to respond to the survey, and a response rate of 82% (983/1192 teachers in total) was obtained. Formal in-depth interviews were carried out with 125 educators, including teachers and education directors of day schools, supplementary schools, and pre-schools, as well as central agency staff and Jewish educators in higher education. The survey and interviews covered a wide variety of issues, such as teachers' background and training, earnings and benefits, and careers of Jewish educators. Only matters of background and formal training are addressed in this paper.

Statistical Methods

For the most part, we combine data from all three communities for our survey analyses. Despite some differences between communities, on the whole the results were far more similar than they were different. Also, our results are largely consistent with surveys carried out in other communities, where comparable data are available. Moreover, in this

paper we will explicitly examine some of the more salient differences across communities. Finally, whereas the data will mainly be aggregated across communities, we will generally break down the data by setting: day school, supplementary school, and pre-school.

We present both descriptive and analytic results. The descriptive results are cross-tabulations of background and training variables by setting. The analytic results derive from ordinary least squares regressions aimed at sorting out predictors of the extent of in-service training.

The analyses rely primarily on survey responses. Information from interviews helped us frame our analytic questions -- in particular, they allowed us to discern the levers for change examined in the regressions -- and they helped us understand the survey findings more thoroughly.

Variables

Most variables indicate aspects of teachers' backgrounds and experiences. These were drawn from surveys. Others provide information about the settings in which teachers work. These came from survey administration records.

Workshop attendance. The dependent variable for this study derives from teachers' responses to the questions, "Were you required to attend in-service workshops during the past two years? If so, how many?" Only teachers who were required to attend at least one workshop are included in the analyses, and first year teachers are excluded because of the two-year time frame implied by the question. This resulted in an effective sample size of 726 teachers. About 15% of teachers who were required to attend workshops failed to indicate how many, and these are treated as missing and excluded from the analyses,

resulting in a sample of 574 teachers, or 85% of the eligible cases. On average, teachers in our sample said they were required to attend 4.75 workshops over a two-year period. (Means and standard deviations of all variables are listed in the appendix.)

Ideally one would like to know how many workshops teachers actually attended, whether required or not, in addition to how many were required. Unfortunately this was not asked in the Lead Community surveys. Future versions of the survey will include an additional question that addresses this distinction (Gamoran, et al., 1995).

Background variables. We employed several measures to take account of differences among teachers in their professional backgrounds. Teachers indicated their years of experience in Jewish education. To allow for possible non-linear effects, we divided experience into four categories: 5 years or less, 6-10 years, 11-20 years, and 21 years or more. An additional category indicates persons with missing data on experience. (We used this strategy of dummy categories for missing data for all independent variables in the regression analyses.)

Teachers also responded to questions about how much schooling they had, what their majors were, and whether they were certified in Jewish education. For this study, we defined "training in education" as a university or teachers' institute degree in education. We defined "training in Jewish studies" as a college or seminary degree in Jewish studies, or as certification in Jewish education.

We used two measures to indicate teachers' professional orientation. First, we asked whether teachers think of their work in Jewish education as a career. Second, we asked teachers about their plans for the future, and from this item we constructed a single indicator

for teachers who said they plan to leave Jewish education in the near future. Presumably it would be possible to demand more in-service work from teachers who are oriented to Jewish education as a career, and are not planning on leaving the field.

Finally, teachers reported their sex, and this is indicated by a dummy variable with 1 = male and 0 = female.

Context and policy variables. Dummy variables are used to distinguish among teachers in day schools, supplementary schools, and pre-schools. Teachers who taught in more than one setting (about 20% of all respondents) are counted in the setting in which they taught the most hours.

For pre-school teachers only, we created an indicator to distinguish among schools that are accredited by the state and those that are not (certified = 1, not certified = 0). For supplementary school teachers only, we created an indicator for the one community with an incentives program for in-service workshops (incentives program = 1, others = 0). For all teachers, we created indicators of the amount of in-service required for re-licensing: 180 hours and 100 hours are compared to the reference category of no in-service requirement.

Results

First we present descriptive information on teachers' professional backgrounds in education and Judaica. Then we examine possible mechanisms for raising levels of in-service training in Jewish education.

Descriptive Results

What sort of professional training in Jewish education characterizes teachers in the three communities? Overall, Table 1 shows that only 19% of teachers in Jewish schools are

formally trained in both education and in Jewish studies. Thirty-five percent were trained in education but not Jewish studies, and another 12% were trained in Jewish studies but not education. This leaves a significant minority —34% —with no formal preparation in either field. Table 1 further shows, not surprisingly, that day school teachers more often have training in Jewish studies than teachers in other schools, and that day school and pre-school teachers more often have professional backgrounds in education than teachers in supplementary schools (combine rows 1 and 2 in Table 1). However, the greater proportion of teachers trained in education in day and pre-schools reflects one- and two-year degrees from teacher training programs as well as university degrees in education. If non-university programs were excluded, day school and pre-school teachers would have formal backgrounds in education similar to that of supplementary teachers.

Further analysis shows that the dearth of formal training is not compensated by extensive in-service education. Table 2 shows that (excluding first-year teachers) day school teachers were required to attend an average of 3.8 workshops during the two-year period, supplementary teachers averaged 4.4, and pre-school teachers were required on average to attend just 6.2 workshops over a two-year period.

Clearly, the infrequency of in-service training is not adequate to make up for deficiencies, nor even to maintain an adequate level of professional growth among teachers who are already professionally trained. What can be done to increase the level of in-service training?

Analytic Results

Table 3 explores background differences in workshop attendance. The first column shows a trend for experience that is roughly linear, with teachers who are more experienced reporting more workshops. In addition, one can see in the first column that controlling for sex and experience, pre-school teachers still reported 2.36 more workshops than day school teachers (the reference category), and supplementary teachers reported .66 more workshops on average. Thus, the pattern that emerged in Table 2 is maintained in multivariate analyses.

The second column presents results for the same model with the additional effects of pre-service training. Teachers with formal preparation in education did not report more in-service workshops, but teachers who are trained in Jewish studies reported that they were required to attend 1.02 workshops more than teachers without such training. The third column of Table 3 shows that teachers who think of Jewish education as their career reported more workshops and teachers who plan to leave the field reported fewer workshops than other teachers. Note also that the initial effects of experience appear to diminish in the second and third columns of Table 3. This pattern suggests that more experienced teachers reported more workshops because they tend to be better trained in Jewish studies and more oriented to a career in Jewish education, two conditions that are obviously connected to longevity in the profession and apparently related to in-service standards as well.

Does the higher rate of reported workshops among pre-school teachers reflect state licensing requirements, as the interviews led us to conclude? To further probe this interpretation, we present in Table 4 the results of a regression that is restricted to pre-school teachers, and which includes an indicator of state-certified pre-schools. As Table 4 shows,

teachers in certified schools reported 3.35 more workshops, a substantial difference considering that the average for pre-school teachers was 6.2 (see Table 2). As in the full-sample analysis, career-oriented pre-school teachers reported more workshops, and those planning to leave reported fewer, although the latter coefficient is not statistically significant due to the smaller number of cases when the sample is restricted to pre-school teachers. (Sex is excluded from the pre-school analysis because all but one of the pre-school teachers are female.)

Do state requirements for re-licensing of trained teachers encourage higher levels of required workshops? Table 5 indicates the answer is no. This analysis, restricted to day school teachers, shows that teachers in states requiring 180 hours or 100 hours of workshop training for re-licensing did not report more workshops than teachers in the state without a fixed workshop requirement. The second column of Table 5 shows that even day school teachers who are formally trained in the field of education did not report more workshops when they worked in states that required many hours of workshops for re-licensing. These results may indicate that day school Judaica teachers do not see themselves as bound by the norms of the general teaching force in the state.

Finally, did the federation-sponsored incentives program encourage higher rates of required workshops? The regression reported in Table 6, restricted to supplementary teachers, shows that teachers who encountered the incentives program reported an average of 2.52 more workshops than supplementary schools in the other two communities, where such federation programs are not in place.

Discussion

This study shows that teachers in three Jewish communities have relatively little formal preparation for their work in Jewish schools. Moreover, they are not typically held to high standards for professional development. However, it appears there are policies that may raise the quantity of in-service. Teachers who are trained in Jewish studies and who are oriented towards a career in Jewish education reported more required workshops. This finding suggests that standards for professional development could be raised by recruiting teachers who are committed to the profession. Better recruitment is an appropriate goal, but it remains a major challenge in light of the relatively small number of opportunities to obtain formal preparation for teaching in Jewish education (Davidson, 1990).

Teachers in certified pre-schools reported substantially more workshops than teachers in other pre-schools. Could this type of policy be implemented in supplementary schools, and in the Judaica divisions of day schools? Where would certification standards come from? One answer is from the community level —the federation or central agency might certify schools whose teachers engage in specified levels of professional growth. For this certification to be meaningful, however, it must be accompanied by some sort of rewards. Parents of pre-school children take certification into account when choosing a school, but this logic does not hold when one is choosing a supplementary school. However, it may be possible to raise parents' expectations so that they seek out supplementary schools and day schools with higher standards for professional growth. In addition, other incentives such as financial support might induce school to seek communal certification.

Although certification of pre-schools made a difference, re-licensing requirements for K-12 teachers did not. In one sense these results may reflect the particular question we asked on the survey, which concerned required workshops instead of any workshops teachers may have attended. Teachers who are meeting individual re-licensing requirements may not have indicated that such workshops are required by their schools. Another interpretation of the results is that rewards and sanctions aimed at individuals are ineffective, but that incentives for schools, as in the case of pre-schools, have more impact.

Finally, supplementary teachers reported more workshops in the community that had an incentives program. This finding suggests that incentives for both individuals and schools affect teachers' professional growth in a positive way. Hence, we conclude that incentives for individuals can be effective if the incentives are meaningful (for example a cash stipend as in this case).

This paper addresses only the quantity of in-service education. The question of quality is at least as important, if not more so. It is essential to consider recent ideas about creating more effective opportunities for professional growth (e.g., Sparks, 1995), at the same time as one thinks about raising the amount of in-service to which teachers are held.

The CIJE's ultimate hypothesis is that building Jewish education as a profession is critical for improving teaching and learning in Jewish education. This paper does not answer that question, but it addresses two crucial concerns along the way: What is the state of the profession? What can be done to improve it? By exploring three potential avenues for reform, we are furthering the broader endeavor. The results of this study suggest two mechanisms ---community incentives and certification of schools -- that can increase the professional growth activities of teachers in Jewish schools.

References

Aron, I. 1990. Toward the Professionalization of Jewish Teaching. Background paper prepared for the Commission on Jewish Education in North America, New York.

Commission on Jewish Education in North America. 1990. A Time to Act. Lanham, MD: University Press of America.

Davidson, A. 1990. The preparation of Jewish Educators in North America: A Research Study. Background paper prepared for the Commission on Jewish Education in North America, New York.

Gamoran, A., Goldring, E., Goodman, R. L., Robinson, B., and Tammivaara, J. 1994. Policy Brief: Background and Training of Teachers in Jewish Schools. New York: Council for Initiatives in Jewish Education.

Gamoran, A., Goldring, E., Goodman, R. L., Robinson, B., and Tammivaara, J. 1995. Manual for the CIJE Study of Educators. New York: Council for Initiatives in Jewish Education.

Sedlak, M. 1987. "Tomorrow's Teachers: The Essential Arguments of the Holmes Group Report." Teachers College Record, 88, 314-325.

Shulman, L. 1987. "Knowledge and Teaching: Foundations of the New Reform." Harvard Educational Review, 57, 1-22.

Sparks, D. 1995. "A Paradigm Shift in Staff Development." The ERIC Review, 3, 2-4.

Table 1. Professional Training of Teachers in Jewish Schools

	Day School	Supplementary School	Pre- School	All Schools
Trained in Education and Jewish Studies	35%	13%	9%	19%
Trained in Education Only	24%	32%	50%	35%
Trained in Jewish Studies Only	25%	11%	3%	12%
Trained in Neither Education Nor Jewish Studies	16%	44%	38%	34%

Table 2.   **Average Number of Workshops Teachers in Jewish Schools Were Required to Attend**

                          Average Number of Workshops
                            in the Past Two Years

Day Schools	3.8
Supplementary Schools	4.4
Pre-Schools	6.2
All Schools	**4.8**

Note: Figures include only those teachers who said they were required to attend workshops, and exclude first-year teachers.

Table 3. Differences among individuals and settings in number of workshops teachers reported they were required to attend.

<u>Independent Variable</u>

Independent Variable			
Sex (Male=1)	-.61	-.74	-.86*
	(.39)	(.39)	(.39)
Experience 6-10 years	.48	.45	.16
	(.35)	(.35)	(.35)
Experience 11-20 years	.81*	.67	.26
	(.37)	(.38)	(.39)
Experience 21+ years	1.02*	.69	.34
	(.43)	(.45)	(.45)
Trained in Education		-.02	-.11
		(.29)	(.29)
Trained in Jewish Studies		1.02**	.60
		(.33)	(.34)
Jewish Education is a Career			1.30**
			(.94)
Will Leave Jewish Education			-1.00*
			(.50)
Pre-school	2.36**	2.76**	2.65**
	(.36)	(.39)	(.38)
Supplementary School	.66*	.98**	1.19**
	(.33)	(.35)	(.35)
Constant	3.37**	2.89**	2.54**
	(.37)	(.43)	(.44)
R^2	.09	.10	.13

*$p < .05$ **$p < .01$

Notes: Metric regression coefficients, with standard errors in parentheses. N=574 teachers. Equation also includes controls for missing data on sex, experience, training in education, training in Jewish studies, career, and plan to leave Jewish education.

Table 4.   Differences between certified and uncertified pre-schools in the number of
                workshops teachers reported they were required to attend.

Independent Variable

Experience 6-10 years	-.81
	(.82)
Experience 11-20 years	-.84
	(.94)
Experience 21+ years	-.74
	(1.18)
Trained in Education	.09
	(.67)
Trained in Jewish Studies	.59
	(.95)
Jewish Education is a Career	1.53*
	(.75)
Will Leave Jewish Education	-1.76
	(1.18)
Certified Pre-school	3.34**
	(1.00)
Constant	2.74*
	(1.17)
Adjusted R^2	.08

*p < .05 **p < .01

Notes: Metric regression coefficients, with standard errors in parentheses. N = 169 teachers.
        Equation also includes controls for missing data on experience, training in education, training
        in Jewish studies, career, and plan to leave Jewish education.

Table 5. Differences in the number of workshops day school teachers were required to attend in states with different professional growth requirements for re-licensing.

Independent Variable		
Sex (Male=1)	-1.07*	-1.05**
	(.45)	(.46)
Experience 6-10 years	1.62*	1.61*
	(.64)	(.64)
Experience 11-20 years	1.12	1.11
	(.62)	(.62)
Experience 21+ years	1.61*	1.62**
	(.67)	(.67)
Trained in Education	-.32	.21
	(.42)	(.49)
Trained in Jewish Studies	.23	-.20
	(.49)	(.53)
Jewish Education is a Career	-.25	-.24
	(.57)	(.58)
Will Leave Jewish Education	-.65	-.60
	(.94)	(.95)
180 Hours Required for Re-License	-.08	-.11
	(.54)	(.92)
100 Hours Required for Re-License	-.36	-.03
	(.48)	(.76)
180 Hours X Trained in Education		.03
		(1.14)
100 Hours X Trained in Education		-.51
		.93
Constant	3.26**	3.19**
	(.66)	(.68)
Adjusted R²	.05	.04

*p < .05   **p < .01

Notes: Metric regression coefficients, with standard errors in parentheses. N = 176 day school teachers. Equation also includes controls for missing data on sex, experience, training in education, training in Jewish studies, career, and plan to leave Jewish education.

Table 6.   Number of workshops supplementary school teachers were required to attend in a community that offered incentives for attendance, compared to other communities.

Independent Variable

Sex (Male= 1)	-.13
	(.46)
Experience 6-10 years	.58
	(.42)
Experience 11-20 years	1.11*
	(.49)
Experience 21 + years	.84
	(.57)
Trained in Education	-.06
	(.37)
Trained in Jewish Studies	.81
	(.44)
Jewish Education is a Career	1.19**
	(.38)
Will Leave Jewish Education	-.53
	(.57)
Community Incentives for Workshops	2.52**
	(.35)
Constant	2.17**
	(.35)
Adjusted R^2	.30

*p < .05 **p < .01

Notes: Metric regression coefficients, with standard errors in parentheses. N=229 supplementary school teachers. Equation also includes controls for missing data on sex, experience, training in education, training in Jewish studies, career, and plan to leave Jewish education.

APPENDIX

Means and Standard Deviations of Variables

	Mean	Standard Deviation
Number of Workshops	4.75	3.31
Sex (Male= 1)	.15	.36
Experience 2-5 years	.27	.44
Experience 6-10 years	.31	.46
Experience 11-20 years	.25	.43
Experience 21+ years	.15	.36
Trained in Education	.54	.50
Trained in Jewish Studies	.32	.47
Jewish Education is a Career	.62	.49
Will Leave Jewish Education	.07	.26
Day School	.31	.46
Supplementary School	.40	.49
Pre-school	.29	.45
Accredited Pre-school	.26	.44
Missing Sex	.01	.11
Missing Experience	.02	.15
Missing Trained in Education	.04	.19
Missing Trained in Jewish Studies	.04	.20
Missing Career	.02	.14
Missing Plans to Leave	.05	.22

Note: N = 574 teachers.

h8 GAMORAN@WISCSSC => ANNETTE@HUJIVMS; 04/05/94, 20:42:04; M GAMORAN MAIL
     EBCDIC (<GAMORAN@WISCSSC>)
MIME type: text/plain

Received: by HUJIVMS via NJE (HUyMail-V61); Wed, 04 May 94 20:42:03 +0300
Date: Wed, 4 May 1994 10:58 CDT
From: <GAMORAN@WISCSSC>
Subject: Notes from our meeting of May 1
To: annette@hujivms
Original To: ALHOFUS, ANNETTE
Original_cc: ELLEN

May 4, 1994

To: Alan Hoffman
From: Adam Gamoran and Ellen Goldring
CC: Annette Hochstein and Steve Hoffman
Re: notes from our meeting with you on 5/1/94

Hit <CR> for next page, : to skip to next part....
BMAIL>
As we decided at the conclusion of our meeting, here are (a) notes from
our m .ing and (b) a list of potential tasks for us, tentatively prioritized,
and a corresponding list of support needed to carry them out.

AGENDA
We were able to discuss four major items on our agenda:

(1) The board subcommitte on research and evaluation
(2) the MEF work plan
(3) the MEF advisory committee
(4) dissemination beyond Lead Community reports

BOARD SUBCOMMITTEE
We observed two problems with the recent meeting of the board subcommittee:

(a) Members of the subcommittee were not familiar with MEF, and the linkage
    between MEF in Lead Communities and CIJE's research mission was ambiguous.

(b) Members of the subcommittee seemed unaware of CIJE's overall program
    of promoting Jewish continuity by improving Jewish education; some
    questioned whether why we were studying personel (how did we know that

Hit <CR> for next page, : to skip to next part....
BMAIL
    would make a difference?) and others, stimulated by Barry Kosmin's
    presentation, asked whether we should perhaps study identity instead
    of education.

As you explained, this is a problem of educating the lay board. At the next
meeting, we need a serious discussion of what it means to set out a research
agenda for Jewish education. This may require a panel of experts. Is there
a potential for research on Jewish education in America? If yes, what would
be the role of the Jewish community, and what would be the role of the
secular educational research community?

Your view was that the October meeting must be carefully thought through
and planned well in advance. You also noted that alternate staffing of
Ellen and Adam is problematic in this context.

There are three main tasks to working with a board committee: (1) Working
with the chair; (2) Working with other committee members; (3) Working on
the content. Of these, the third is the real work.

MEF WORK PLAN

In responding to our work plan of 4/1/94, you raised four concerns at the
outset:

(1) There is not enough attention to informal education.

(2) You are pleased to see educational leaders addressed, but noted that
only the characteristics of leaders, and not leadership itself, will be
addressed. That is a concern.

(3) It is not clear how the work plan moves beyond three communities, as
CIJE is planning to do.

(4) The timing of writing the cross-community report on educators was
not satisfactory.

Essentially, you said that the pieces of the work plan are fine in themselves,
but the timing and priorities they imply need further discussion.

 Mobilization_
We discussed our ongoing monitoring of community mobilization, and reached

a co nsus that the documents produced by the field researchers, though
rich in detail, are better seen as raw data than as interpretive reports.
We discussed the need for a cross-community report on mobilization with
more interpretation.   This might be useful for the board subcommittee on
mobilization, as well as for CIJE staff. Adam suggested that a comparative
report could be helpful in drawing lessons in anticipation of CIJE's
likely transformation as envisioned in the 10-year plan.

 Institutional Profiles_
In light of the emerging centrality of the goals project, an instrument
to create institutional profiles will definitely be needed. Our task
now should be to study and design an instrument for the profiles, without
necessarily planning to implement them on a community-wide basis next fall.
On the contrary, we should move more towards an instrument useful to
individual institutions (as opposed to an instrument mainly for community-wide
purposes), which would be used by institutions engaged in vision-driven
reform. (Note: We raised, but did not have time to discuss, the question of
what happens to systemic reform when innovation is driven by individual
institutions.) If possible, we should have an instrument ready to be used
by institutions that get "on-board" after the goals seminar this summer.
The purpose of the instrument would be to permit baseline assessment of

the institution(s) so that progress could be assessed over time.

We also discussed the need for deeper profiles that would include data
about teachers' sense of mission, unity of purpose, etc. We did not
set a time frame for addressing such issues. Moving a step further,
we also discussed the need for data on constituencies (and potential
constituencies) -- i.e., parents and students. We did not set a time
frame for this work either.

We concluded that Bill should go ahead with interviews of experts in
Jewish education, with the aim of creating a draft instrument to present
at the August MEF advisory committee meeting. The draft would be accompanied
by a rationale for each indicator.

Bill needs to talk with Dan Pekarsky to discuss the linkage between the
institutional profiles and the goals project. (Probably we'll bring him
to Madison for this. He can also meet with Roberta to get her input on
the indicators. Ellen's participation in the goals seminar will also be
helpful.)

The decision not to try to implement institutional profiles in the Lead

Communities, at least not next fall, is a MAJOR CHANGE in our work plan.

Reports on Educators
In your ideal schedule, we would have a cross-community report on Jewish
educators ready to present at the October board meeting and to release
at the November GA. This is not possible. However, we could make a
presentation at the GA (and previewed at the board meeting) on a fairly
narrow topic -- for example, educational backgrounds and professional
development of teachers -- at the GA, to accompany related presentations
by leading educational figures. We anticipate having a draft of the
full cross-community report to our advisory committee by December 31.

We understand that this project is our TOP PRIORITY.

MEF ADVISORY COMMITTEE
We discussed the meetings and composition of our advisory committee.
It is desirable to add another educational researcher, especially if
Jim Coleman is not able to participate. We discussed a few names but
did not reach any conclusion. One possibility is to elevate our committee
into an advisory committe for research, for which MEF in Lead Communities

is but one component. In that case, we might add David Cohen and Lee
Shulman as committee members. We think this is a promising idea that
warrants further discussion.

We set a tentative date for our next advisory committee meeting of
August 24-25. An alternate would be August 25-26. If you agree, we will
ask Ginny to contact the members of our committee to find out if this
would work for them.

We also discussed the possibility of a meeting in Israel in the first
few days of January, 1995. Adam, Ellen, Alan, Annette, Seymour, and Mike I.
would be invited to this meeting, a sort of pre-advisory meeting. The
discussion would presumably center on the cross-community report on teachers
in Jewish schools, which will just have been drafted. Another topic of
discussion at this meeting would be our work plan for 1995.

****
TASKS FOR ADAM AND ELLEN (in order of priority) (comments follow)
((all dates are for drafts submitted to MEF advisory committee))

TASK	MAIN SUPPORT NEEDED FROM:	ADDITIONAL SUPPORT NEEDED FROM:
Cross-community teacher report (December 31, 1994)	Bill Robinson	Roberta, Julie
Report on Baltimore teaching force (June 30, 1994)	Nancy Hendrix	Julie
Report on Atlanta teaching force (August 31, 1994)	Nancy Hendrix	
"Module" of educator surveys and interviews (May 31, 1995)	Julie, Roberta	
Reports on characteristics of educational leaders in the L.C.s (Fall, 1994)	Bill Robinson	Roberta, Julie

Instrument for institutional profiles (August, 1994, through 1995)	Bill Robinson	Roberta, Julie	A
Cross-community mobilizat. report (June, 1995)	Roberta, Julie	Bill	N.G.
Conceptual paper on Jewish community mobilization (September 30, 1994)	Roberta, Julie	Bill	
Instruments development for study of informal educators (Winter, 1995-1996)	outside experts	Roberta, Julie, Bill	
Participation on the CIJE Steering Committee (Ongoing)			
Staffing the CIJE Board Sub-committee on Research & Eval.	outside experts		

(Ongoing)

Comments:

All these tasks seem doable under the schedule indicated, with one important exception: We cannot see a way of adequately staffing the Board Subcommittee on Research and Evaluation, along with all our other work. This, we recognize, is a serious problem. *what about more staff/this*

An important omission from this list is additional meetings and presentations which are frequently asked of us by CIJE and/or Lead Communities. We continue to be very reluctant to add this extra work, because we are too busy with our main agenda. *W*

The longer we have field researchers on staff, the more we'll be able to say in the cross-community report on mobilization. However, we recognize that this report is not the highest priority.

If we drop the cross-community mobilization report, we could prepare the instr 'nts for studying informal education next year (1995). *DO DROP* *other*

The role of the field researchers in preparing the teaching force reports should not be overlooked. We expect they will make substantial contributions to each LC report, and we also expect them to assist us in preparing the cross-community report on educators.

After January 1995, we will still have great need for a data analyst, and we hope Bill Robinson will prove capable in that role. If he also turns out to be effective in preparing instruments for institutional profiles, CIJE may want to hire him as much as 100%. If his work for us will be restricted to data analysis, it is crucial that we have at least 50% of his time for CIJE; 100% would be better but if an accommodation can be made with Atlanta, perhaps they could have 50% of his time and we could have the other 50%.

Finally, a couple of activities we mentioned but which do not appear on the list: A study of leadership in Jewish education; a study of institutional practices (as opposed to profiles of institutional characteristics); a study of students and/or parents. These items need further discussion.

Facsimile Transmission

To: Mrs. B. Hochstein ((for Micky))   .. May 31,1994
Mayflower Hotel

From: Annette Hochstein   No.. Pages 9 '

Fax number: 972-2-662-837

Dear Miri,,

 It was lovely talking to you before Shabbat. I hope
both of you are well,, and that time flies so you can
soon fly home.
 Could you please give the attached document to Micky..

                    Love,,

                    [signature]

Dear Micky,,

 As promised here is Shira Breuer's document following
a rework with me. She is sending it to the foundations
listed on the last page. It would be great if you
could do your own thing as discussed.   .

 Your suggestions are welcome-we will introduce changes
as you see fit. She would like a cover letter from the
school-I think we should endorse it. If you agree I'll
take care of that.

 Hope you're having a good trip..

                    Love,

                    A══

                  Annette

To Whom It May Concern:

For the past four years, I have served as principal of the
Pelech Religious Experimental High School for Girls, following
five years as a teacher of history there. As a school that
provides a religious humanistic education for women, Pelech is
a unique institution whose innovative programs and alternative
teaching methods have served as models for both religious and
secular schools. In 1993 the school was awarded the Ministry
of Education Prize for curriculum development.

After several years as a practicing educator, I wish to devote
my forthcoming sabbatical year (1994-95) to research and
study, concentrating in the areas of developing leadership
among young women, humanistic religious education, and women's
studies.

The greater New York area, with its plethora of educational
institutions, provides an ideal location for my planned
program of study, as outlined in the attached proposal. It
will allow me to become acquainted with the latest
developments in alternative education and provide me with
state-of-the-art new ideas and approaches for my
administrative work.

To make it possible to carry out this program of study, I am
requesting a one-year fellowship as a living stipend and to
cover the cost of tuition fees and associated study expenses.

Please accept my thanks for considering this request.

Sincerely,

Shira Breuer
Principal

PELECH

Religious Experimental High School for Girls

WHOM DOES PELECH SERVE?

Pelech High School was established in 1967 to fulfill a void in Israel's educational system; namely, to provide a religious humanistic education for girls. Pelech today serves approximately 250 students in grades 9 through 12 from the greater Jerusalem area. They reflect a broad spectrum of ethnic, religious, and socioeconomic backgrounds - new immigrant and native Israeli; Anglo-Saxon, European, and Oriental, disadvantaged and middle class - with academic excellence, intellectual curiosity, and leadership potential the sole criteria for acceptance.

EDUCATING FOR RESPONSIBLE LEADERSHIP

Pelech seeks to prepare its students for leadership roles in a modern, technological society, while working to prepare women for changing roles in the 21st century. As an intrinsic part of the school's basic ideology, students, through representative institutions, actively participate in setting and carrying out school policy. Special emphasis is placed on public responsibility as a fundamental value; upon graduation, virtually all students enter the army, many as officers in elite units, or national service. Because of the leadership roles alumnae play in so many walks of life, their influence is beyond all proportion to their numbers.

WHAT MAKES PELECH SPECIAL?

To foster the values that it espouses and to encourage independent, creative thinking, Pelech constantly develops interdisciplinary curricula, innovative extra-curricular programs, and alternative teaching and study methods. Special units on issues such as "The Arts and the Bible," "Conflict and Conflict Resolution," "Presentation of Women in the Media," "Toward a Responsible Environmental Policy," and "Religious Feminists in Modern Society" are aimed at providing the tools to grapple with contemporary social and political challenges. Innovative learning formats such as a "Beit Midrash" that facilitate discussion and dialogue are designed to enhance comprehension and insight and to develop analytical skills.

Pelech has been officially designated as an "experimental" high school - one of only three in the country. Accordingly, it serves as a model school for the student teachers at the Hebrew University's School of Education. Every year, Pelech also hosts dozens of visiting principals, supervisors, and teachers from Israel and abroad - from religious, secular, and even non-Jewish educational institutions. Pelech staff members participate in a variety of forums for innovative curriculum development. Various curricula developed at Pelech in the humanities, Jewish studies, and sciences have served as prototypes for numerous other schools.

CURRICULUM VITAE

Shira Breuer
1 HaGoren St.
POB 1370
Efrat, Israel

Experience

◆1990-1994 Pelech Religious Experimental High School for Girls,
Principal

◆1985-1990 Pelech Religious Experimental High School for Girls,
Teacher of History

◆1990 Kerem Institute for Jewish and Humanistic Education, Teacher
of course on "Didactics of Teaching History"

◆1983-1990 Department of Jewish Education, University of Tel Aviv
and Ministry of Education. Counsellor of teachers of Judaic
Studies

◆1987 Tali Educational System, Wrote comprehensive study program
for teaching Mishna, grades 1-8

◆1985-1989 Ben Zvi Institute, Teacher of course on "Jerusalem
During the Period of the Second Temple" to high school teachers
on sabbatical

◆1984 Midreshet Jerusalem (Neve Schechter), Taught course on
"4000 Years of Jewish History"

◆1980-1983 Department of Jewish History, Hebrew University,
Teaching Assistant

◆1979-1982 Tali School, Gilo
Built eight-year Jewish studies curriculum for new school

Education

1976-1979 Hebrew University, BA in Jewish History and Philosophy
1980-1983 Hebrew University, MA in Jewish History

Kerem Institute for Jewish and Humanistic Education, Teaching
Degree

Military Service

1972-1974 Lieutenant

Personal

Married + 4 children

PELECH

Religious Experimental High School for Girls

PROPOSAL FOR A ONE-YEAR FELLOWSHIP
Presented May 1994

I) The Need for Enrichment in Educational Administration

Constant enrichment and intellectual cross-fertilization are essential in the field of educational administration. This is especially true for experimental education, which demands constant staff development, creative initiatives, and the encouragement of alternative thinking.

Furthermore, the school administrator's daily responsibilities generally leave little time to explore new approaches and theories. Immersion in the routine tasks of running the school seldom permit the educational overview and comprehensive thinking necessary for long-range planning and development. These reasons underline the importance of a periodic sabbatical for the principal, to allow him or her to take stock, evaluate, seek solutions to practical questions that have arisen, explore new areas, and regain the perspective needed for effective educational aaminstration.

II) The Present Opportunity

Accordingly, I wish to devote my forthcoming sabbatical year (1994-95) to research and study, concentrating on the areas of developing leadership amount young women, humanistic religious education, and women's studies. Expanding my horizons by engaging in new areas of study, while free from ordinary responsibilities and concerns, is crucial to my professional growth.

The abundance of important teaching institutions, progressive Jewish and non-Jewish secondary schools, and leaders in the field of alternative education in and within easy reach of New York, makes it 5 most appropriate place for such study.

III) Objectives

My goal is to become acquainted with the latest developments in the aforementioned areas, at both the theoretical and practical levels, and methods of implementing them in school. The specific purpose of this research is to enhance my work as an administrator upon my return to Israel, by introducing new methodologies and initiating new programs to develop leadership, strengthen students' involvement in decision-making and responsibility, and deepen their religious-humanistic understanding.

My specific objectives include:

a) Learning the latest theories by taking courses and becoming acquainted with professional literature

b) Meeting and networking with key personalities involved in both religious and general experimental education

c) Making site visits - to alternative schools, religious institutions, and principals" centers

IV) Planned Program of Study
I hope to integrate theory and practice through the activities that follow. Upon arrival in New York, I shall be able to explore opportunities in greater depth and consult with knowledgeable individuals in order to finalize details of the program.

a) Taking selected courses at academic institutions (e.g., Columbia University Teachers" College and others)

b) Visiting leading progressive schools (e.g., Fieldston and The Dalton School in New York, the democratic Southberry School in Boston, the Philadelphia Parkway School, and other models of alternative education)

c) Visiting principals" centers (e.g., the Harvard Principals' Centers. etc.)

d) Visiting Jewish and non-Jewish religious schools, to observe how religious beliefs are presented vis a vis the modern world

e) Meeting with noted educators and experts (e.g., Daniel Greenberg, the principal of the Soutnbury School: Theodore Sizer, of Brown University, Chairman of the Coalition of Essential Schools), to observe how their educational ideas are carried out and for advice regarding our educational ideas

f) Participation in educational conferences (e.g. Educational Leadership Conference, American Educational Researchers" Assocation, etc.)

V) Anticipated Results
This experience will expose me to the latest developments in curricular approaches, methodologies, and learning situations, providing new ideas to introduce in my administrative work upon my return to Israel. I would like to write a Women's Studies Program that will help raise consciousness and develop women's leadership in a meaningful way.

V) The Need: A One-Year Fellowship
Estimated living expenses for one year are $20,000, plus an additional $10,000 for research and study expenses, and related costs, as itemized below. I therefore respectfully request a fellowship in that amount. Upon completion of my research, I will present the donor with a detailed report of my activities and conclusions.

VI) Personal Background
I am a native American, with B.A. and M.A. degrees from the Hebrew University and a teaching degree from the Kerem

Institute for Jewish and Humanistic Education. I have been solely or jointly responsible for designing a variety of innovative interdisciplinary programs at Pelech and other schools, which have served as models for both religious and secular schools. I have taught both undergraduates and practicing teachers at various institutions including the University of Tel Aviv, the Hebrew University, the Kerem Institute, and the Ben Zvi Institute.

VII) Expenses – One Year

Living costs	$20,000
Tuition for courses	5,000
Commuting and travel	3,000
Books, periodicals, and conference fees	2,000
Total	$30,000

May 29, 1994

Mr. Henry Everett
The Everett Foundation
3th floor?" F™ nd«i
635 Madison Avenue
New York, NY 10022
USA

Mrs. Erica Jesseslon
The Jesselson Foundation
1221 Avenue of the Americas
New York, NY 10020

Dear Mrs. Jesselson.

May 29, 1994

Mr. Arthur Fried
20 Balfour Street
Jerusalem

Dear Mr. Fried.

For the past four years I have

Ms. Joy Ungerleider-Mayerson
President
The Dorot Foundation
Three Manhattanviile Road
Purchase, NY 10577-2110

Dear Ms. Ungerleider-Mayerson,

PrL'r

University of Wisconsin—Madison

MADISON, WISCONSIN 53705

DEPARTMENT OF SOCIOLOGY
SOCIAL SCIENCE BUILDING
1180 OBSERVATORY DRIVE

TO CALL WRITER DIRECT
?!PHONE (608)

MEMORANDUM

April 29, 1994

To: Annette Hochstein
From: Adam Gamoran
Re: Conference of the Research Network on Jewish Education

Julie Tammivaara and Roberta Goodman have submitted the attached "work-in-progress" for presentation at an informal roundtable at the conference of the Research Network on Jewish Education on June 4-6. The presentation will not make reference to specific communities or individuals. It seems clear from the abstract that neither problems of confidentiality nor problems of conflict with ongoing CIJE or Lead Community efforts are raised by this proposal, so I have given tentative consent to the presentation. Please let me know if you see any problems with this.

It is understood that should a written document be prepared, it would need to go through our review process before it could be published or disseminated.

OK
No problem
tell Adam

Abstract

Teacher Power in Jewish Education

Julie Tammivaara and Roberta Goodman
January 1994

The purpose of this paper is to examine the concept of power in the context of formal Jewish educational settings. We take as our definition of power Heilbrun's statement that "Power is the ability to take one's place in whatever discourse is essential to action and the right to have one's part matter," and as our conceptual perspective Berger and Luckmann's formulation of the social construction of reality. Our data sources include interviews with teachers, principals, and rabbis in two major American cities. We found that, in general, the concept of power outside the immediate classroom setting is relatively alien to most teachers. Within this constraint, we argue that what constitutes important educational discourse and, thus , where teachers can exert power, varies across settings. While Judaic knowledge, meaning, feeling, and practice are important to virtually all educators, they are not equally important to all. Similarly, the primacy of curriculum, teaching methods, educational philosophy, and school administration differs widely across schools and affects the nature of discourse important to action within a setting.

Adam and Ellen,

[I have copied Ellen but have been unsuccesful in reaching her
already once this week. Please ask her to e-mail me to my
Compuserve address so that I can get the right address off the
e-mail]

I thought the report was excellent!! Cogent, powerful and very
damning without it looking as though you were taking a
particular
position.

Some comments:

Hit <CR> for next page, : to skip to next part.....
BMAIL>
-[2J:-[田
1.   Although you mention it in the box on page 5 you use
the
term "Jewish teachers" when you really mean Jewish studies
teachers
except for the pre-school. Maybe "Jewish educator" is a
better
term or maybe it needs a definitional statement or footnote
right
up in thefront.   We also do not quite make clear that this
excludes (and that is the problem with the word "educator1, I
know)
informal and adult educators who are not caught in the net of
schooling.   I think these distinctions need some clarifying
in a
way which does not make your text cumbersome.

2.   Page 1, para.2: I would quibble with the word 'portrait"
reserving that term for thicker description and use profile as
you
used it in para. 3 of the same page.

3.   If you have more info. on deliberate hiring of part-time
to
reduce salary costs, (para. 3, p.3) it would be helpful as I
see this as
a very powerful and damaging finding.   It at least should be
revisited
somewhere in the policy conclusions.

4.   The comparison with administrators and principals is
missing
right throughout.   It is germane with regard to salaries and

-[2J*-[H
benefits ((p.4, para.2) but also later on with regard to Judaic and
general background and attitudes to career. Do we have the data and
can we introduce a comparative perspective in the report?

5.   The issue of salaries as "supplementary" or "insignificant"
((p.4, paras. 2 & 3) is not really clear. There is a difference
between the two.   In general, one of the questions in my mind
is whether upper middle class women who enjoy the few hours in
the classroom and for whom the pay and benefits are not of any
consequence are really stretchable in terms of upgrading and
uptraining. The complacency resistance may be quite high
here.

6.   Is it not worth alluding somewhere to those good people who
maybe COULD BE in the system but have elected to search elsewhere
for work because of compensation, benefits and part-time. In
other words, WHO WOULD A BETTER PACKAGE BRING IN?

7.   On page 5, maybe Milwaukee cannot recruit nationally because
of the conditions.

8.   In the box on page 5 I think you need to say something about

-[2J-[H
administrators and principals.

9.   Bottom of page 6: Do we know whether salaries are closely
correlated with experience.

10. Your use of the term "experienced" on page 7, para.2 and
several times thereafter has a connotation of "good" teachers.
That, of course, we do not know.

11. Top of page 8: Given the socio-economics, we do not know that
we want those who NEED full-time work to be full-time. It may be
that other incentives need to be found for upper middle-class
woment whom we will wish to have full time.

12. The pre-school Jewish background picture on page 10, para.2
is truly horrific and very powerful.

13. The impression is created ((page 12, para.2 see quote) that

the
professional development for early childhood educators is good,
even in Judaica (("I would hope")).   I am not at all sure that is
true.   Do we have any info. on the content of professional

Hit <CR> for next page, : to skip to next part......
BMAIL>
«[2J=JH[H
developemnt for this population.   My sense is that although
early childhood has to fulfill general licensing requirementss,
the
Judaic pro. dev is very, very weak.

14.   The section on Career Perceptions ((p. 17) is striking
particularly para. 3 on that page.   Why do you think that is?
Is
it related to the socio-economic issues referred to above?

15.   On the tables: Table 3 - Do weknow how many of the
teachers
get which benefits together in which packages?
                      Table 9 - Can this be broken down inot
Early Childhood, Day School ?

14.   All the comparisons with the Wisconsin public school
system   .
are excellent and some should appear in boxes.

I hope this is helpful.   We need to talk about the MEF
meeting

Hit <CR> for next page, : to skip to next part.....
BMAIL>
=[2J=[H
in February and about what to do with the report.
I told them this week in Milwaukee that they would have a
draft
within a week.   I shall call you ((either one of you) soon.

alan

Wisconsin Center for Education Research
University of Wisconain–Madison / School of Education
1025 West Johnson Street, Madison, Wisconsin 53706

FAX: (608)263-6448

TEL: (608) 263-3451

DATE: 2/18/94

PROJECT NO.: 1710

Total # of pages in
transmitted document,
(including this page): 4

Facsimile Transmission Form

To: Annette Hochstein   FAX #: (2) 265-5048
                               Tel. #: _____
Organization: Mayflower Hotel

From: Adam Gamoran   i Tel. # _____

Address: _____

Message:
For Annette), revised executive summary.

Thanks for meeting with us on this.