This paper describes the empirical literature on consultation effectiveness from 1961-1989, based on a review of journals and a computer search of online databases that identified 119 journal articles, book chapters, and monographs and 59 dissertation abstracts. The median number of data-based publications exploring consultation effectiveness was less than five per year. Psychology journals published more than double the number of such studies than did special education journals. Two-thirds of the investigations used group rather than single-case designs. Behavioral consultation was four times more likely to be investigated than mental health models. In nearly two-thirds of the studies, student or teacher behavior was used alone or in combination with another criterion to judge consultation success, whereas student achievement was a criterion in only one-quarter of the studies. Findings are discussed in terms of how consultation may be perceived in a political climate that is placing increasing emphasis on outcomes and accountability. An appendix contains a list of the identified journal articles, book chapters, monographs, and dissertation abstracts. (27 references) (Author/JDD)
Where is the Research on Consultation Effectiveness?

Douglas Fuchs and Lynn S. Fuchs
George Peabody College of Vanderbilt University

Jeannette Dulan
Oakwood College

Holley Roberts
George Peabody College of Vanderbilt University

Pamela Fernstrom
University of North Alabama

Preparation of this article was supported in part by Grant No. G008730082-88 from the Office of Special Education Programs in the U.S. Department of Education and by the National Institute of Child Health and Human Development (Core Grant HD15052). The article does not necessarily reflect the position or policy of the funding agencies and no official endorsement should be inferred.

We are grateful to Ann Nevin for sharing her published work and that of her colleagues on the Vermont Consulting Teacher Program; to Walt Pryzwansky for sending a list of dissertations published from 1978 to 1985; and to Marilyn Friend for sharing the Bibliography on Consultation for Special Education.

Address inquiries to Douglas Fuchs, Department of Special Education, Box 328, George Peabody College, Vanderbilt University, Nashville, TN 37203.

Running head: CONSULTATION EFFECTIVENESS
This article describes the empirical literature on consultation effectiveness from 1961-1989. A review of journals and a computer search of on-line databases identified 119 journal articles, book chapters, and monographs and 59 dissertation abstracts. During this 29-year period, the median number of data-based publications exploring consultation effectiveness was less than 5 per year. Psychology journals published more than double the number of such studies than did special education journals. Authors of two-thirds of the investigations used group, rather than single-case designs. Behavioral consultation was four times more likely to be investigated than mental health models. In nearly two-thirds of the studies, student or teacher behavior was used alone or in combination with another criterion to judge consultation success, whereas student achievement was a criterion in only one-quarter of the studies. These and other findings are discussed in terms of how consultation may be perceived in a political climate that is placing increasing emphasis on outcomes and accountability.
Where is the Research on Consultation Effectiveness?

As we write, the nation is two months away from receiving its first report card from the National Education Goals Panel, headed by Governor Roy Romer of Colorado ("The First Report Card," 1991). Students are to be graded in math, science, and literacy in an effort to monitor America's progress toward President Bush's ambitious educational goals for the year 2000. The National Educational Goals Panel, as well as numerous reports of blue ribbon panels, commissions, and think tanks, signal greater concern about and expectations for public schools than a decade ago. Moreover, the Panel's work seems in step with public sentiment. The annual Gallup-Phi Delta Kappan Poll of public attitudes toward the public schools found that 80% of those interviewed favor national achievement tests and goals for local schools (Henry, 1991). With "outcomes" and "accountability" as buzz words -- not just in the capitol, but across the nation -- we asked, To what extent has the school consultation literature been oriented toward outcomes and how might we characterize this important segment of scholarly work? Although some who write about consultation argue that many more studies on effectiveness are needed (e.g., Snow, 1988; Witt & Martens, 1988), there have been few efforts in recent years to document the number of such studies.

Toward this end, we have taken a longitudinal look at the number of efficacy studies published in a 29-year period (see Appendix). We also have analyzed the amount published as dissertations in contrast to journal articles and, among the journal articles, how many were published in special education versus psychology versus counseling journals. Moreover, we have attempted a multidimensional characterization of this literature, including (a) how
consultation effectiveness has been measured; (b) how often different consultation models have been studied; (c) what proportion of investigations have used single-case versus group designs; and (d) how many were implemented in elementary schools versus middle schools and junior- and senior-high schools. In short, our analysis is like a topographical map; a detailed delineation of salient characteristics of a body of research to show "elevations" and relative positions, thereby rendering a configuration of sorts.

Method

Inclusion and Exclusion Criteria

Articles targeted by this literature review described data-based studies that explored consultation effectiveness in elementary (K-8) or secondary (9-12) schools. Consultation typically was defined as involving a triadic network of consultant, teacher, and student and indirect service. However, articles judged appropriate also described situations in which teachers functioned as consultants as in "teacher assistance teams" (e.g., Chalfant & Pysh, 1989). Effectiveness was measured in terms of consultation's influence on teacher attitude or behavior or student attitude, behavior, or academic performance. Thus, for example, an article by Chalfant, Pysh, and Moultrie (1979) was not included because the datum was the percentage of students referred to a teacher assistance team that was helped within a building; there were no data on teacher change or student change.

The search also excluded survey (e.g., West & Cannon, 1983) and simulation (e.g., Conoley & Conoley, 1982) studies and evaluations of direct service (e.g., Nevin, Paolucci-Whitcomb, Duncan, & Thibodeau, 1982). It eliminated process-oriented analyses like those of consultant-consultee interviews (e.g., Bergan & Neumann, 1980) as well as research demonstrating
changes in consultants, rather than in teachers and students (e.g., McDougall, Reschly, & Corkery, 1988). Also excluded were interventions of which consultation was a part, but not a major component, as in reports of some inservice training (e.g., Anderson, Kratochwill, & Bergan, 1986) and applied behavior analytic research (e.g., Ayllon & Roberts, 1974).

Search Procedures

The literature search subsumed five strategies pursued in the following order: (a) a hand search of selected professional journals; (b) a computer search of the on-line databases of Educational Resources Information Center (ERIC) and Dissertation Abstracts International (DAI); (c) analysis of references in previously published integrative reviews of consultation research; (d) an "ancestral" search of references in pertinent articles and dissertations identified by the first three strategies; and (e) inquiries to colleagues with interest in consultation. Below is an elaboration of procedures followed in connection with each strategy.


The hand search was conducted in two steps. First, titles and abstracts of every article published in the 11 journals during the years mentioned were scanned visually. A total of 292 articles were identified as "promising." In
the second step, these articles were read carefully to determine whether they met the inclusion criteria. Forty-six were judged appropriate.

Interrater agreement was defined as agreements divided by agreements plus disagreements and multiplied by 100. It was determined (a) by asking two graduate students to evaluate 16 articles from the 1974-1983 issues, inclusive, of the Journal of School Psychology and (b) by giving two other graduate students 15 articles from the same journal published between 1986-1989, inclusive. These 31 Journal of School Psychology articles represented 11% of the 292 "promising" articles identified during step 1 of the search. The first pair of reviewers agreed on the appropriateness of 13 of 16 (81%) articles; the second pair concurred on 13 of 15 (87%). Thus, interrater agreement was obtained for the second, but not the first, step of the 2-step hand search.

Computer search. A computer-assisted search was made of two on-line databases: ERIC and DAI. ERIC was reviewed from 1969-1985, inclusive, using the descriptors "teacher consultation," "behavioral consultation," and "school consultation." They generated 132, 51, and 82 abstracts, respectively, of which only 7 articles were judged appropriate for inclusion. ERIC was reviewed again for the years 1986-1989, inclusive, with additional updated "keywords." These keywords, and in parentheses the corresponding number of abstracts each generated, follow: "consultant and behavior" (25), "school-psychologists and teacher" (39), "organizational-development and consultation," (6), "counselor-teacher-cooperation" (35), and "consultation-programs and behavior" (29). Of these 134 abstracts, 6 were deemed appropriate for inclusion. Thus, an ERIC search for the years 1969 through 1989 yielded 13 appropriate articles in addition to the 46 identified by the hand search.
DAI was searched from 1861-1989. Search descriptors included "school consultation," "teacher consultation," "mental health consultation," "behavior consultation," "consultation with teacher," and "consulting with teacher." A total of 141 abstracts were generated and reviewed. Thirty were judged appropriate.

**Integrative reviews of consultation research.** References in eight previously published reviews of consultation research were explored. The reviews were conducted by Alpert and Yammer (1983), Carrow (1988), Duncan and Pryzwansky (1988), Mannino and Shore (1975), Medway (1979, 1982), Medway and Updyke (1985), and Sibley (1986). This strategy generated an additional 68 appropriate articles.

**Ancestral search.** References in appropriate articles identified by our hand and computer searches were reviewed. These references generated 8 additional appropriate articles, and the references of these references produced 3 more for a total of 11.

**Contacting colleagues.** Ann Nevin provided published literature on the Vermont Consulting Teacher Program, which included eight appropriate papers. Two additional manuscripts were obtained from colleagues.

**Analysis of Databased Studies**

The empirical literature describing consultation effects on teacher attitude and behavior and student attitude, behavior, and academic performance was analyzed in the following ways: (1) the number of published studies (journal articles, book chapters, and monographs) and dissertations per year; (2) the number of journal articles by journal type (psychology vs. special education vs. counseling); (3) the number of journal articles reported both by journal type and by year; and (4) the number of published articles: (a) by type of study design (group vs. single case with \( N = 1 \) vs. single case with
Consultation

more than one study participant); (b) by model (behavioral vs. mental health vs. organizational); (c) by measure; (d) by outcome; (e) by client (teacher vs. student); and (f) by consultation setting. These last six analyses were conducted on 119 published papers (113 journal articles, 3 book chapters, and 3 monographs).

Interrater agreement for the last six analyses was again defined as agreements divided by agreements plus disagreements multiplied by 100, and was determined as follows: Two reviewers independently coded 19 (19%) of the 98 studies published between 1961 and 1985, with the following levels of agreement per analysis: study design, 100%; consultation model, 92%; dependent measure, 100%; outcome, 97%; targeted clients, 95%; consultation setting, 100%. Another pair of reviewers independently coded 10 (48%) of the 21 studies published between 1986-1989, with the following levels of agreement: experimental design, 90%; consultation model, 93%; dependent measure, 92%; outcome, 80%; targeted clients, 80%; consultation setting, 100%. Altogether, 29 (24%) of the 119 published studies included in the review were used to calculate interrater agreement, with overall agreement of 93%.

Results

A total of 119 published studies and 59 dissertation abstracts are represented in Figure 1, which displays the number of data-based investigations of consultation effectiveness produced during 1961-1989, inclusive. Each vertical bar in the Figure represents production for 2 years, not 1.

Insert Figure 1 about here

As indicated by the Figure, it was not until the early 1970s that the
number of pertinent published articles and dissertation abstracts began to increase, and continued to do so until the late 1970s when the numbers dropped off. Throughout much of the 1980s, the number of articles and abstracts was not dramatically higher than that produced through the 1960s. The median total number of articles and abstracts across the 15, 2-year time intervals displayed in Figure 1 is 9.5. This suggests that, on average, less than 5 new investigations of consultation effectiveness were produced per year from 1961 through 1989.

Figure 2 is based on 113 published journal articles identified by our search (119 published studies minus 6 non-journal studies), and shows the number of such articles by 2-year intervals and journal type (i.e., special education, psychology, or counseling and other). As indicated, psychology journals published the most data-based articles on consultation effectiveness with 62, primarily between 1971 and 1982. Across the 29 years represented by the figure, this amounts to a mean of 2.1 published articles per year. Special education journals published 28 articles. In 1981-1982, and again in 1985-1986, there were 6, the highest 2-year total for the years searched. Contrastingly, in 1961-1962, 1963-1964, 1965-1966, and 1983-1984, there were no data-based studies of consultation found in special education journals. From 1961-1962 through 1989, the average number of special education data-based articles exploring consultation effectiveness was just less than one per year.

------------------------
Insert Figure 2 about here
------------------------

Analysis of 119 published studies by type of experimental design revealed that the authors of 79 studies (66%) used group design. Single-case design
was employed in the remaining 40 articles, with 22 of these including more than one study participant.

Five categories were applied to the 119 published studies to determine the frequency with which the effectiveness of specific consultation models was studied. Three of the five categories correspond to well-known consultation models: Behavioral, mental health, and organizational development. Of the two remaining categories, "joint" signifies a study of two or more of the just-stated models, whereas "other" refers to a consultation process that was eclectic or insufficiently clear. Behavioral consultation was implemented in 59 studies (50%), nearly four times more than the number that explored the efficacy of mental health (n = 16, or 13%). Organizational development's effectiveness was explored in only 9 (8%) of the investigations. Eight articles (7%) reported studies of two or more consultation approaches, and 27 (23%) were classified in the "other" category. (The sum of percentages does not equal 100 due to rounding.)

There were four dependent measures used alone or in combination in all but 7 of the 119 published studies: teacher ratings of students (TR), systematic observations of teacher or students (SO), questionnaires or interviews (Q/I), and tests of academic achievement (T). Figure 3 shows that systematic observation of teacher or students was the most frequently used dependent measure, with 33 (28%) studies relying on it as a single index of consultation effectiveness and 60 (50%) studies employing it alone or in combination with additional measures. Questionnaires or interviews were used almost as frequently, with 28 (24%) and 53 (45%) investigations using the measure alone and in combination, respectively. In comparison, only 5 (4%) studies used tests of academic achievement alone, and 25 (21%) used them in combination. A mere 2 (2%) investigations employed teacher ratings of
Consultation

students as a sole measure of consultation effects; 17 (14%) studies used the
measure alone or in combination with other dependent measures.

----------------------------------

Insert Figure 3 about here

----------------------------------

The number of studies focusing on student (n = 47, 39%) versus teacher (n
= 42, 35%) was nearly evenly divided. Thirty focused on both the teacher and
student.

How were consultation outcomes operationalized? Or, what domain was
chosen to index consultation effects? We found four: Behavior (B), academic
achievement (AC), attitudes (AT), and other (O). As shown in Figure 4,
student or teacher behavior was the single criterion of consultation
effectiveness in 46 (39%) studies. In 77 (65%) studies, it was one of two or
more criteria used to judge consultation success. Change in student or
teacher attitude was the litmus test for consultation effectiveness in 23
(19%) investigations. In 37 (31%) studies, it was used alone or in
combination with other criteria. Only 8 (7%) and 32 (27%) studies used
student achievement alone and in combination, respectively, to gauge the
efficacy of consultation. And 8 (7%) studies concentrated singly on outcomes
other than behavior, academic achievement, and attitudes; 12 (10%) used
"other" outcomes in combination with one or more of these criteria.

----------------------------------

Insert Figure 4 about here

----------------------------------

An overwhelming majority of consultation effectiveness studies were
implemented in general kindergarten through eighth-grade classrooms.
Specifically, 77 of 119 (65%) took place in these settings in comparison to
only 9 (8%) that were implemented in grades 9 to 12. Such studies in special education are a rarity; 6 and 3 were conducted, respectively, in kindergarten through eighth grade and grades 9 to 12. Remaining studies (n = 24, 20%) either did not specify a setting or were implemented in combined special and general education, or combined elementary and secondary, locations.

Discussion

The search of data-based studies on consultation effectiveness included a hand search of 11 professional journals, a computer search of two on-line databases, an analysis of references in previously published integrative reviews, an "ancestral" search of references in pertinent articles and dissertations, and inquiries of colleagues with interest in consultation. Given the comprehensive nature of this search, we found surprisingly few pertinent investigations. From 1961 through 1989, less than five dissertations and published articles were produced on average per year (see Figure 1). Across these 29 years, 62 data-based articles on consultation effectiveness were published in the psychology journals reviewed, translating into a mean of just 2.1 per year. Special education journals published 28 articles in the same time span, an average of less than 1 article per year (see Figure 2).

Such numbers may underestimate the quantity of extant studies because while Figures 1 and 2 encompass the period from 1961 through 1989, the hand search of journals began in 1974 and the on-line search of the ERIC database started in 1969. That is, our search was more rigorous from the mid-1970s onward than it was before the early 1970s. Nevertheless, we do not believe a more comprehensive search of the earlier years would have generated enough additional articles to alter the conclusion that few data-based studies of consultation effectiveness have been reported in the fugitive (e.g., ERIC) and
published literature.

At this point, some readers may be wondering, "So what? Haven't integrative reviews determined that consultation is effective? How many studies do the fields of special education and school psychology need before all can be satisfied?" This line of questioning has a commonsense ring, but it is misleading. The issue is not whether consultation is (or isn't) effective. To claim it is effective is as correct and useful as stating that phonics (or language experience) is an effective reading strategy. Just as phonics is effective sometimes with some teachers and students, so too is consultation. Neither works all the time, perhaps not even most of the time, which brings us to the business of science.

An important purpose of consultation research is specifying the right consultation strategy for a given situation. How collaborative versus prescriptive should it be, for example, and for whom and when? Should consultation be the same or different in non-departmentalized elementary schools versus middle schools and junior- and senior-high schools? And given that many school psychologists, special education teachers, and guidance counselors -- professionals likely to have interest and training in consultation -- often lack administrative support and opportunity to provide consultation, how streamlined may we make the process and still have confidence in its effectiveness?

These "situational" questions highlight the importance of additional findings from our topographical analysis. For example, only 13% of the published studies addressed the effectiveness of mental health consultation, and a mere 8% explored the efficacy of an organizational-development approach. Moreover, from a previous review, we know of just one study that compares systematically alternative models of consultation within the same experimental
Consultation 1

design. Similarly, whereas 77 studies were confined to mainstream kindergarten through grade 8 settings, only 9 concentrated on grades 9 to 12. And to our knowledge, we know of no efforts, aside from our own (e.g., Fuchs & Fuchs, 1989; Fuchs, Fuchs, Bahr, Fernstrom, & Stecker, 1990), to conduct component analyses of well-known consultation models.

Although the extant database may indicate that, overall, consultation tends to be effective, it is too small to provide much guidance in a particular situation. Gresham and Kendell (1987) make the same point in a different way: "To say that there are 'experts' in consultation is an oxymoron because expertise denotes that an individual has special knowledge in a particular field. We simply do not know enough about consultation, how it works, under what conditions it works, or the most important variables in predicting successful consultation outcomes" (p. 314). The problem is compounded by the possibility that much of the database may be inaccurate and misleading because many studies have been determined to be poorly conceptualized and executed (see, for example, Alpert & Yammer, 1983; Gresham & Kendell, 1987; Medway, 1979; Meyers, Pitt, Gaughan, & Freidman, 1978; Pryzwansky, 1986). Only a small handful of group-design studies have been experimental in nature. Moreover, much of the research using single-case designs involve only one or two subjects. Whereas some of this work can lay claim to internal validity and importance (Kratochwill, 1985), its external validity is questionable (Medway, 1982).

The foregoing prompts the question, Why are there so few data-based studies of consultation effectiveness? We offer two reasons. First, as Pryzwansky (1986) and others have noted, intervention-oriented, consultation research is very difficult to do well. It demands careful planning, attention to detail, interpersonal skills, flexibility, positive relations with school
Consultation

folks, and, needless to say, knowledge about research. Its unique challenge, we believe, has to do with the "links-in-the-chain" problem. The researcher first must train participating consultants who subsequently must interact effectively with teachers who in turn must work with one or more students. If the research involves many consultants, teachers, and students, the researcher may rely on assistants whom he or she must train before they train consultants and so forth. Each link in the process -- from researcher to assistant, assistant to consultant, consultant to teacher, and teacher to student -- involves issues concerning materials, training, fidelity of treatment, summative evaluation, and more. As anyone knows who has played the child's game of "telephone," the more links in a chain the greater the difficulty in communicating clearly and maintaining consistency of purpose and effort. In short, Murphy's Law is more likely affirmed by conducting intervention-oriented consultation research than by implementing research on other topics.

A second reason for the infrequency of research on consultation effectiveness is associated with "teacher empowerment." Many school-based consultants and researchers are right, we believe, to be interested in how consultation may enhance teachers' autonomy, self-esteem, and problem ownership. This concern, however, has led at least some practitioners and researchers to rely exclusively on teacher reaction as a bellwether of their success or failure. At professional (psychology and special education) conferences it is not uncommon for speakers to describe teachers' "involvement," "enjoyment," and "satisfaction" as necessary and sufficient proof of consultation effectiveness; as if the teacher, not the student, were the client. Consonant with this observation are data in our Figure 4, which indicate that only 32 of 119 (27%) published studies used student academic
achievement as an outcome. When, as in most cases, the student is the client, the consultant should demonstrate his or her effectiveness in terms of student performance. This is not to gainsay teachers' thoughts and feelings; rather, it is to argue that such information must be supplemented by data on students.

Researchers must generate new knowledge about which type of situation calls for what type of consultation, and how consultation may be made more effective, efficient, and attractive to teachers. Practitioners can contribute by regularly collecting meaningful data. With researchers' and practitioners' combined efforts, consultants are more likely to achieve success, be perceived by their colleagues and administrative superiors as valuable, and be positioned strategically in a political environment placing increasingly high premiums on accountability. Without greater emphasis on outcomes, we predict consultants will be pushed further to the margins of school life, where they and their expertise will be viewed as expendable.
References


Figure Captions

Figure 1. Number of dissertations and published studies by 2-year periods.

Figure 2. Number of special education, psychology, and counseling journal articles by 2-year period.

Figure 3. Number of published studies using: teacher ratings (TR), systematic observations (SO), questionnaires/interviews (Q/I), tests (T), and other (O).

Figure 4. Number of published studies operationalizing consultation effectiveness in terms of behavior (B), academic achievement (AC), attitudes (AT), and other (O).
NUMBER OF PUBLISHED ARTICLES

TR  SO  Q/I  T  O

TYPE OF MEASURE

Used this measure in combination
Used this measure alone
NUMBER OF PUBLISHED STUDIES

Measured this outcome in combination
Measured this outcome alone

TYPE OF OUTCOME

B
AC
AT
O
Appendix:

Data-based Journal Articles, Book Chapters, Monographs, and Dissertation Abstracts on Consultation Effectiveness


Graden, J. L., Casey, A., & Bonstrom, O. (1985). Implementing a prereferral intervention system:


Hasazi, S. E. (1976). The consultant teacher. In J. Jordan (Ed.), *Teacher, please don't close the door* (pp. 50-59). Reston, VA: Council for Exceptional Children.


McCoy-Simandle, L. (1989). The effect of consultation as a follow-up to assessment on the achievement of students in kindergarten (Doctoral dissertation, University of Kentucky,


program to increase behaviors and attitudes that promote learning. *Journal of Community Psychology, 5*, 246-256.


Trione, V. (1967). The school psychologist, teacher change and fourth grade reading achievement.


process variable-collaboration (Doctoral dissertation, Rutgers University, The State
University of New Jersey). *Dissertation Abstracts International*, 37, 3531A-3532A.


White, G. W., & Pryzwansky, W. B. (1982). Consultation outcome as a result of in-service resource


Whitley, A. D. (1971). Counselor-teacher consultations including video analysis to reduce
*Dissertation Abstracts International*, 31, 5142A-5143A.


sensory extinction for reducing the disruptive classroom behavior of a profoundly retarded
child. *School Psychology Review*, 12, 205-211.

Wixson, S. E. (1980). Two resource room models for serving learning and behavior disordered

Yaryan, C. M. (1975). Cognitive and affective consultation with teachers: Effect on interaction and
International*, 35, 7074A.

Young, C. E. (1971). The effects of three consultation models on individual and group decision
Abstracts International*, 32, 2413B.

35, 278A.