Journal of Research in Education

Fall2002

An International Journal Published
by the Eastern Educational Research Association

CONTENTS

Cooperative Learning Methods: A Meta-Analysis
David W. Johnson and Roger T. Johnson

Teaching Students to be Peacemakers: A Meta-Analysis
Roger T. Johnson & David W. Johnson

Ethnicity, 'Role Models' and Teaching
Bruce Carrin

Developing Authentic Instruction in the Social Studies
Patricia G. Avery, Carol Freeman, and Dana Carmichael-Tanaka

Perception of Care Among Black Female High School Students
Erin McNamara & Kristine S. Lewis

Reflective Racial Tolerance and its Development in Children, Adolescents and Young Adults: Age Related Difference and Context Effects
Rivka T. Witenberg

Promoting the Development of Talent in Technical Areas: Obstacles to Females Pursuing Technical Careers in Europe, Scandinavia Countries, Asia, and the United States
James Reed Campbell
Guest Editor

Introduction: Gender Inequity Among Academic Olympians Across the Globe: Theoretical Paradigms
James Reed Campbell

Developing Females' Talent: Case Studies of Finnish Olympians
Kirsi Tirri

German Olympiad Studies: Findings from a Retrospective Evaluation and from In-Depth Interviews. Where Have all the Gifted Females Gone?
Angelika Longfelder & Kurt A. Heller

Understanding Gender Inequity in America: Interviews with Academic Olympians
Annie Xuemei Feng, James R. Campbell, & Marilyn A, Vema

American Chemistry Olympians Achieve the Highest Level of Equity
Marilyn A. Vema & Annie Xuemei Feng

Korean Gifted Girls and Boys: What Influenced Them to be Olympians and non-0/ympians?
Seokhee Cho & Hye-Joo Lee

How Do the Theoretical Paradigms Fit the Findings of the International Olympiad Studies?
Sharon O'Connor

Teacher Research
Barbara Beyerbach Editor

Introduction
Barbara Beyerbach

A Review of the Work of Lawrence Stenhouse: Questions, Ambiguities, and Possibilities
Cher C. Hendricks

Discipline Referrals: A Wealth of Information
Carlo Cuccaro

Improving Novice Teachers' Communication Skills and Urban Families: Professor-Action Research
Peter McDermott and Julia Johnson Rothenberg

Methodological Issues
David Shannon and Isadore Newman Editors
Introduction
David Shannon & Isadore Newman

Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers
Andrea D. Clements & Pat W. Myrick

Research Practices that Produce Type VI Errors
Isadore Newman, John W. Fraas, Carole Newman, & Russell Brown

Cooperative Learning Methods: A Meta-Analysis

David W. Johnson
and
Roger T. Johnson
University of Minnesota

Cooperative learning is one of the most widespread and fruitful areas of theory, research, and practice in education. Reviews of the research, however, have focused either on the entire literature which includes research conducted in noneducational settings or have included only a partial set of studies that may or may not validly represent the whole literature. There has never been a comprehensive review of the research on the effectiveness in increasing achievement of the methods of cooperative learning used in schools. An extensive search found 164 studies investigating eight cooperative learning methods. The studies yielded 194 independent effect sizes representing academic achievement. All eight cooperative learning methods had a significant positive impact on student achievement. When the impact of cooperative learning was compared with competitive and individualistic learning, Learning Together (LT), Group Investigation (GI), and Academic Controversy (AC) tended to promote the greatest effect on achievement, followed by Student-Team-Achievement-Divisions (STAD), Teams-Assisted-Individualization (TAI), Jigsaw, Cooperative Integrated Reading and Composition (CIRC), and finally Teams-Games-Tournaments (TGT). The consistency of the results and the diversity of the cooperative learning methods provide strong validation for its effectiveness.

Cooperative learning is one of the most fertile areas of theory, research, and practice in education. *Cooperative learning* exists when students work together to accomplish shared learning goals (Johnson & Johnson, 1999). Each student can then achieve his or her learning goal if and only if the other group members achieve theirs (Deutsch, 1962). In the past three decades, modern cooperative learning has become a widely used instructional procedure in preschool through graduate school levels, in all subject areas, in all aspects of instruction and learning, in nontraditional as well as traditional learning situations, and even in after-school and non-school educational programs. There is broad dissemination of cooperative learning through teacher preparation programs, inservice professional development, and practitioner publications. One of the strengths of cooperative learning is that a wide variety of different ways of operationalizing it have been implemented in schools and colleges. The diversity of these methods begs the question as to whether they are all equally effective. There has been no comprehensive review of the research evidence examining this issue. The purpose of this review, therefore, is to examine the empirical support validating the effectiveness of the different methods of cooperative learning. In order to do so, it is first helpful to discuss why cooperative learning is so widely used.

The widespread use of cooperative learning may be largely due to three factors—it is based on theory which has been validated by research and operationalized into procedures educators can use. First, cooperative learning is based solidly on a variety of theories in anthropology (Mead, 1936), sociology (Coleman, 1961), economics (Von Mises, 1949), political science (Smith, 1759), psychology (Deutsch, 1949), and other social sciences. In psychology, where cooperation has received the most intense study, cooperative learning has its roots in social interdependence (Deutsch, 1949, 1962; Johnson & Johnson, 1989), cognitive-developmental (Johnson & Johnson, 1979; Piaget, 1950; Vygotsky, 1978), and behavioral learning theories (Bandura, 1977, 2000; Skinner, 1968). It is rare that an instructional procedure is central to such a wide range of social science theories.

Second, the amount, generalizability, breadth, and applicability of the research on cooperative, competitive, and individualistic efforts provides considerable validation of the use of cooperative learning, perhaps more than most other instructional methods (Cohen, 1994a; Johnson, 1970; Johnson & Johnson, 1974, 1978, 1989, 1999a; Kohn, 1992; Sharan, 1980; Slavin, 1977, 1991). There are over 900 research studies validating the effectiveness of cooperative over competitive and individualistic efforts. This body of research

Table 1

Modern Methods Of Cooperative Learning

Researcher-Developer	Date	Method
Johnson & Johnson	Mid 1960s	Learning Together & Alone
DeVries & Edwards	Early 1970s	Teams-Games-Tournaments (TGT)
Sharan & Sharan	Mid 1970s	Group Investigation
Johnson & Johnson	Mid 1970s	Constructive Controversy
Aronson & Associates	Late 1970s	Jigsaw Procedure
Slavin & Associates	Late 1970s	Student Teams Achievement Divisions (STAD)
Cohen	Early 1980s	Complex Instruction
Slavin & Associates	Early 1980s	Team Accelerated Instruction (TAI)
Kagan	Mid 1980s	Cooperative Learning Structures
Stevens, Slavin, & Associates	Late 1980s	Cooperative Integrated Reading & Composition (CIRC)

has considerable generalizability since the research has been conducted by many different researchers with markedly different orientations working in different settings and countries and in eleven different decades, since research participants have varied widely as to cultural background, economic class, age, and gender, and since a wide variety of research tasks and measures of the dependent variables have been used (Johnson & Johnson, 1989).

The research on cooperative efforts, furthermore, has unusual breadth, that is, it has focused on a wide variety of diverse outcomes. Over the past 100 years researchers have focused on such outcomes as achievement, higher-level reasoning, retention, time on task, transfer of learning, achievement motivation, intrinsic motivation, continuing motivation, social and cognitive development, moral reasoning, perspective-taking, interpersonal attraction, social support, friendships, reduction of stereotypes and prejudice, valuing differences, psychological health, self-esteem, social competencies, internalization of values, the quality of the learning environment, and many others. Many of these studies have focused on preventing and treating a wide variety of social problems such as diversity (racism, sexism, inclusion of handicapped), antisocial behavior (delinquency, drug abuse, bullying, violence, incivility), lack of prosocial values and egocentrism, alienation and loneliness, psychological pathology, and low self-esteem (see reviews by Cohen, 1994a; Johnson & Johnson, 1974, 1989, 1999a; Johnson, Johnson, & Maruyama, 1983; Kohn, 1992; Sharan, 1980; Slavin, 1991). A strength of cooperative learning is that it simultaneously achieves such diverse outcomes.

The third factor contributing to the widespread use of cooperative learning is the variety of cooperative learning methods available for teacher use, ranging from very concrete and prescribed to very conceptual and flexible. Cooperative learning is actually a generic term that refers to numerous methods for organizing and conducting classroom instruction. Almost any teacher can find a way to use cooperative learning that is congruent with his or her philosophies and practices. So many teachers use cooperative learning in so many different ways that the operationalizations cannot all be listed here. In assessing the effectiveness of specific cooperative learning methods, however, there are a number of "researcher-developers" who have developed cooperative learning procedures, conducted programs of research and evaluation of their method, and then involved themselves in teacher-training programs that are commonly credited as the creators of modern-day cooperative learning. The following ten have received the most attention (see Table 1): Complex Instruction (CI) (Cohen, 1994b), Constructive Controversy (CC) (Johnson & Johnson, 1979), Cooperative Integrated Reading and Composition (CIRC) (Stevens, Madden, Slavin, & Farnish, 1987), Cooperative Structures (CS) (Kagan, 1985), Group Investigation (GI) (Sharan & Sharan, 1976, 1992), Jigsaw (Aronson, et al., 1978), Learning Together (LT) (Johnson & Johnson, 1975/1999), Student Teams Achievement Divisions (STAD) (Slavin, 1978), Teams-Games-Tournaments (TGT) (DeVries & Edwards, 1974), and Team Assisted Individualization (TAI) (Slavin, Leavey, & Madden, 1982).

This combination of theory, research, and practice provides considerable support for the use of cooperative learning. Knowing that the research indicates that cooperative learning can have powerful effects when properly implemented does not mean, however, that all cooperative learning procedures will be effective or equally effective in maximizing achievement. While many different cooperative learning methods are being advocated and used, educators have very little guidance as to which specific

cooperative learning methods will be most effective in their situation. There is a need to examine the empirical support validating the effectiveness of the different methods of cooperative learning in maximizing achievement. More specifically, four issues will be investigated.

The first issue is how much research has been conducted to validate specific cooperative learning procedures. While the voluminous research on cooperation has been summarized in a various books and articles (Cohen, 1994a; Johnson & Johnson, 1989; Sharan, 1980; Slavin, 1977), the majority of the research studies on cooperation do not directly test the effectiveness of specific cooperative learning procedures. Many of the research studies that have been conducted may be classified as *efficacy studies* (i.e., laboratory studies of short-term effects) as opposed to *effectiveness studies* (i.e., real-world studies of how cooperative learning is actually delivered and what the outcomes are like). Effectiveness studies can be divided into studies aimed at testing theory as well as the effectiveness of a cooperative learning method and curriculum evaluation case studies that have little theoretical relevance but demonstrate that a cooperative learning method worked in a specific situation. The two types of studies complement each other. While a number of people have reviewed the research supporting their cooperative learning methods (e.g., Cohen & Lotan, 1997; Sharan & Sharan, 1992, Slavin, 1991), there has never been a comprehensive and complete review of the effectiveness studies on all the different cooperative learning methods. It is unknown, therefore, how much of the existing research specifically focuses on cooperative learning methods and achievement.

The second issue investigated is how many different cooperative learning methods have been evaluated. As noted earlier, cooperative learning is a generic term referring to numerous methods for organizing and conducting classroom learning. It is used in many different variations, most of which have never been evaluated. There has never been a comprehensive assessment of how many of cooperative learning methods have been empirically tested. The methods most frequently referred to in the research and educational methods literatures are listed in Table 1.

The third issue investigated is how effective are the different cooperative learning methods in maximizing achievement. Once it is known how much research has been conducted on how many of the cooperative learning methods, the next issue is the strength of the empirical support for each method. In order to determine the size of the effect of each cooperative learning method on student achievement, a meta-analysis must be conducted. *Meta-analysis* is a method of statistically combining the results of a set of independent studies that test the same hypothesis and using inferential statistics to draw conclusions about the overall result of the studies (Cohen, 1987; Cooper, 1989). The meta-analysis process basically consists of a literature search and the calculation of effect sizes.

The fourth issue investigated is what are the characteristics of the more effective cooperative learning methods. Methods of cooperative learning may be placed on a continuum from direct to conceptual. More *direct cooperative learning methods* consist of very specific and well-defined techniques, scripts, structures, lessons, or curricula that are used in a lock-step, prescribed manner. Direct methods can be learned in a few minutes, can be applied immediately, and are used in the same way in all situations. More *conceptual cooperative learning methods* consist of conceptual frameworks (e.g., expert systems) teachers learn and use as templates to create cooperative lessons and activities and to restructure current lessons and activities into cooperative ones. Teachers are trained to create unique adaptations to their specific students, circumstances, and needs. Direct methods may initially be more appealing and seem more user friendly, while conceptual methods (once they are mastered) may be integrated into teachers' teaching repertoires and used throughout their career (Antil, et al., 1998; Berman, 1980; Berman & McLaughlin, 1976; Fullan, 1981; Griffin & Barnes, 1984; Johnson, 1970, 1979; Johnson, Druckman, & Dansereau, 1994; Johnson & Johnson, 1994a, 1994b; Joyce & Showers, 1980, 1982; Smith & Keith, 1982). More specifically, more *direct methods* tend to be easy to learn (and require less training time), tend to be easily implemented, are often focused on specific subject areas and grade levels (i.e., nonrobust), are easy to discontinue as interest wanes, and are not easily adapted to changing conditions. More *conceptual methods* tend to be difficult to learn and use initially, may be used in lessons in any subject area for any age student (i.e., robust), and become internalized and routinely used and thus difficult to discontinue, and are highly adaptable to changing conditions. While the considerable research on direct and conceptual innovations documents their strengths and weaknesses affecting implementation and institutionalization, there is almost no research on the important issue of the relative impact of direct and conceptual innovations on achievement and productivity. In this review, therefore, methods of cooperative learning will be classified on a direct-conceptual continuum and correlated with the size of each method's effect on student individual achievement.

Methods

Literature Search

The studies included in this meta-analysis were identified through a thorough search for relevant published and unpublished studies. Methods included conducting computer searches (Educational Resources Information Center [ERIC], Psychological Abstracts [PA], Dissertation Abstracts International [DAI], and the Social Sciences Citation Index [SSCI]), examining relevant bibliographies, searching reference sections of the studies included in the meta-analysis to identify further relevant studies, and contacting relevant researchers and organizations. We also

examined the bibliographies of previous relevant meta-analyses. Over 900 studies on social interdependence were located. The criterion for inclusion in the meta-analysis was that the study evaluated the impact of a specific method of cooperative learning on student achievement. A total of 164 studies met the criteria. Since some reports contained multiple studies, the total number of reports was 158. Since studies that compared multiple cooperative learning methods or had more than one control condition are listed more than once, the tables present 194 separate comparisons of cooperative learning and control methods.

Independent Variables

The first independent variable is method of cooperative learning. *Method of cooperative learning* was defined by the author(s) of each article. If the author stated that the method used was STAD or Jigsaw it was noted as such. In addition, the operationalization of the method had to include positive interdependence. Examples are positive goal interdependence (mutual goals), positive reward interdependence (joint rewards), resource interdependence (each group member has different resources that must be combined to complete the assignment), and role interdependence (each group member is assigned a specific role). Studies that included intergroup competition as part of operationalizing cooperation were included among the cooperative conditions.

Cooperative learning is compared with competitive or individualistic learning. *Competition* was operationally defined as the presence of negative goal or reward interdependence. Participants worked alone or with a minimum of interaction and rewards were given on a norm-referenced basis or by ranking participants from best to worst. All studies in this analysis focused on competition among group members, not competition between groups. *Individualistic efforts* were operationally defined as the lack of social interdependence between participants. Participants worked alone or with a minimum of interaction and rewards were given according to set criteria so there was little opportunity for social comparison. When the control condition was labeled as *traditional instruction*, the condition was coded as either competitive or individualistic depending on the description of the condition.

The second independent variable was the classification of cooperative learning methods on a continuum of direct to conceptual. More *direct cooperative learning methods* consist of well-defined procedures that teachers are supposed to follow in an exact, lock-step way while more *conceptual cooperative learning methods* consist of conceptual frameworks teachers use as a template to overlay lessons and activities they structure to fit their specific circumstances. Each cooperative learning method was rated by two psychology professors on five criteria, each of which was defined as a five-point scale. The ratings on the scales were added together to get a total score. Ease of learning (how quickly the method can be learned) was rated on a five-point scale from "a simple procedure easy to understand and remember" to "a conceptual system difficult to understand and apply." Ease of initial use (the effort required to implement the method initially) was rated on a five-point scale from "a simple procedure easy to do perfectly the first time" to "general, conceptual guidelines that are applied to specific lessons and activities." Ease of maintaining its use over time (once implemented, how difficult it is to discontinue) was rated on a five-point scale from "a procedure that is not integrated into basic teaching patterns" to "conceptual framework that is integrated into basic teaching patterns." Robustness (applicable to specific subject area and age level) was rated on a five-point scale from "aimed specifically at a subject area and grade levels" to "can be applied to any subject area and grade level." Adaptability (how difficult it is to modify cooperative learning to ensure its effectiveness in changing conditions) was rated on a five-point scale from "lock-step, specific procedures that have to be done the same way every time" to "conceptual system that can be modified and changed to meet changing conditions." A direct cooperative learning method, for example, may be easy to learn, easy to use initially, can be performed without integrating framework into basic teaching patterns, aimed at a specific subject area and grade level, and difficult to adapt to changing conditions. Conceptual method, on the other hand, may be hard to learn, difficult to implement initially, integrated into basic teaching patterns and thus maintained long-term, applicable to all subject areas and grade levels, and easy to adapt to changing conditions.

The method of cooperative learning, the control conditions, and the direct-conceptual nature of the cooperative learning methods were coded by two or more analysts, all psychology professors, with extensive experience coding and analyzing research on social interdependence. Interrater reliability was calculated using the kappa coefficient (Cohen, 1960). The interrater reliability kappa was 0.82. The occasional differences in coding were discussed and resolved through consensus.

Dependent Variable

The dependent variable was student achievement. *Achievement* was defined as an outcome measure for some type of performance (standardized and teacher-made tests, grades, quality of performances such as compositions and presentations, quality of products such as reports, and so forth). A variety of experimental settings and tasks were used in the studies yielding effect sizes for the dependent variable of achievement.

Effect Size

The statistical methods and terminology for meta-analysis are from Cohen (1987), Hedges and Olkin (1985), Cooper (1989), Hunter, Schmidt, and Jackson (1982), and Glass, McGaw, and Smith (1981). The *effect size d* was the difference between treatment divided by the pooled standard deviation of the two groups (Cohen, 1989). When means

were not given, but significance task results were, the *F*, *T*, or X^2 was converted to *d* (Cooper, 1989). All effect sizes were adjusted to control for small sample bias (Hedges & Olkin, 1985). Within studies where there were multiple achievement measures, the *average effect size* was found by averaging the multiple measures to derive one effect size for each treatment contrast. The *mean weighted effect size* was found by multiplying each independent effect size by the inverse of its variance and then the sum of these products was divided by the sum of the inverses. The resulting weighted mean effect size is referred to as "d+." Confidence intervals (95 percent) were calculated to determine the statistical significance of each weighted mean effect size (Cooper, 1989). Tests for homogeneity of variance (Q) of effect sizes were calculated.

Participants

The participants in the studies included in this study include included kindergarten, elementary, middle school, secondary, college, other post-secondary, and adult education students. The participants were from urban, suburban, and rural schools and were from a variety of ethnic groups. While most of the participants were residents of the United States, studies were conducted in Canada, Asia, Europe, Africa, and the Middle East.

Results

Characteristics Of The Studies

A total of 158 studies on specific cooperative learning studies met the criteria for inclusion in this meta-analysis. The characteristics of the studies are found in Table 2. All studies have been conducted since 1970 with 28 percent conducted since 1990. Thirty percent did not randomly assign participants to conditions, 45 percent randomly assigned participants to conditions, and 25 percent randomly assigned groups to conditions. Forty-six percent were conducted in elementary schools, 20 percent were conducted in middle schools, 11 percent were conducted in high schools, and 24 percent were conducted in post-secondary and adult settings. Sixty-six percent of the studies were published in journals. Fifty-two percent lasted for 2 to 29 sessions (a session was defined as 60 minutes or less), and 46 percent lasted for 30 sessions or more. Ninety-four percent of the studies involved mixed gender groups. Four studies were conducted in Southeast Asia, 3 studies were conducted in the Middle East, 3 studies were conducted in Europe, four studies were conducted in Africa, and several of the North American studies contained minority group students.

Results For Different Cooperative Learning Methods

The results for the different methods of cooperative learning appear in Table 3. While both the averaged and the weighted effect sizes appear in Table 3, only the weighted effect sizes are discussed here. For *Learning Together*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.70 and 0.91 respectively). There tends to be no meaningful difference between competitive and individualistic efforts, effect size = 0.08. For *Teams-Games-Tournaments* (TGT), cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.47 and 0.55 respectively). For *Group Investigation*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.86 and 0.62 respectively). For *Academic Controversy*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.61 and 0.86 respectively). For *Jigsaw*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.41 and 0.09 respectively). For *Student-Teams-Achievement-Divisions*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.46 and 0.28 respectively). For *Team Assisted Individualization*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.19 and 0.19 respectively) and individualistic efforts promoted higher achievement than competition, effect size = -0.32. Finally, for *Cooperative Integrated Reading and Composition*, cooperation promotes higher achievement than do competitive or individualistic efforts (effect sizes = 0.20 and 0.18 respectively). The results for the averaged effect sizes were very similar, somewhat higher for Learning Together and TAI and somewhat lower for Group Investigation and Jigsaw. No studies were found for *Cooperative Learning Structures*. There are studies of *Complex Instruction* (see Cohen & Lotan, 1997), but the studies compared complex instruction with other group instructional procedures (rather than comparing them with competitive or individualistic instruction) and, therefore, relevant effect-sizes could not be derived.

The cooperative learning methods may be ranked by the size of the effect they have on achievement and by the number of comparisons available (see Table 4). When the impact of cooperative lessons is compared with competitive learning, Group Investigation promoted the greatest effect, followed by Learning Together, Constructive Controversy, STAD, Jigsaw, TAI, CIRC, and finally TGT. When the impact of cooperative lessons is compared with individualistic learning, Learning Together promotes the greatest effect, followed by Constructive Controversy, Group Investigation, STAD, TAI, CIRC, Jigsaw, and TGT. There are reasons, however, why these rankings should be suggestive only. The few number of studies conducted on several of the methods makes the effect sizes very tentative. The fewer the number of studies, the more likely the same investigator conducted the studies, the less the variety of achievement measures used, the more homogeneous the participants, and the more similar the subject matter studied. In addition, different measures of achievement were used in the different studies. The confidence educators can have in the effect

Table 2

Characteristics Of Studies Of Cooperative Learning Methods

Characteristic	Number	Percent
1970–1979	26	16
1980–1989	88	56
1990–1999	44	28
No Random Assignment	48	30
Randomly Assigned Groups, Subject Unit of Analysis	30	19
Randomly Assigned Groups, Group Unit of Analysis	9	6
Randomly Assigned Subjects	71	45
Primary (K–3)	22	14
Intermediate (4–6)	43	27
Primary & Intermediate	8	5
Middle School (7–9)	32	20
High School (10–12)	17	11
Post Secondary	33	21
Adult	4	3
Journal Article	105	66
Book	2	1
Theses	28	18
Technical Report	17	11
Unpublished	6	4
1 Session	3	2
2–9 Sessions	38	24
10–29 Sessions	45	28
30+ Sessions	72	46
Same Gender Groups	10	6
Mixed Gender Groups	148	94
North America	145	91
Southeast Asia	3	2
Africa	4	3
Europe	3	2
Middle East	3	2
Total	158	100

Total Number of Reports = 158; Total Number of Studies = 164 (some reports gave results for multiple studies)

sizes, furthermore, is inversely related to the number of studies that have been conducted on the method. When the methods are ranked by the number of effects associated with each findings, for the cooperative versus competitive comparison, the ranking of the methods is Learning Together, Academic Controversy, STAD, Jigsaw, TGT, TAI, CIRC, and Group Investigation. For the cooperative versus individualistic comparison, the ranking is Learning Together, Academic Controversy, STAD, TAI, TGT, Group Investigation, and CIRC.

There are five dimensions on which the methods of cooperative learning may be evaluated (see Table 5): (a) ease of learning the method, (b) ease of initial use in the classroom, (c) ease of long-term maintenance of use of the method, and (d) robustness of the method (applicability to a wide variety of subject areas and grade levels), and (e)

adaptability of method to changing conditions. Each cooperative learning method may be classified on a five-point scale (easy to difficult) on these dimensions. When the resulting score is correlated with the effect-sizes for each method, the results indicate that the more conceptual the cooperative learning method, the higher the achievement of cooperative compared with competitive, $r(197) = 0.32, p < 0.001$, and individualistic learning, $r(197) = 0.46, p < 0.001$. Finally, the direct-conceptual rating and the characteristics of the studies (whether participants were randomly assigned to conditions, age of participants, type of cooperative learning, whether cooperation is contrasted with competition or individualistic efforts, duration of the study, mixed vs. homogeneous groups in terms of gender, and year published) were placed in a linear regression analysis using weighted effect sizes. The results appear in Table 6. The only significant predictor of the impact of cooperative learning on achievement was the direct-conceptual rating, $t = 6.78, p < .0001$.

File Drawer Problem

A potential source of bias in reviewing a set of studies may be that only studies that tend to find significant differences are published and available for review. There may be numerous unpublished works that might change the overall findings. Orwin (1983) presented a procedure for determining how many studies would have to be unpublished to change the results found. He makes an assumption that the effect sizes from unretrieved findings are equal to zero (which is very conservative). His statistic then determines how many studies in file drawers with an average effect size of zero would be needed to shift the obtained weighted mean effect size to a criterion level such as 0.20 (which is small as defined by Cohen [1987, p. 25-26]). The results from Table 3 indicate that for TAI, CIRC, the cooperative vs. individualistic comparison for Jigsaw, the competitive vs. individualistic comparison for LT, and the cooperative vs. individualistic comparison for GI, even one more study could significantly change the results. For the other cooperative learning methods, it would take from 2 to 200 additional studies to change the results significantly.

Discussion

Cooperative learning has been around a long time (Deutsch, 1949; Johnson, 1970; Johnson & Johnson, 1989, 1999). It will probably never go away due to its rich history of theory, research, and actual use in the classroom. Markedly different theoretical perspectives (social interdependence, cognitive-developmental, and behavioral learning) provia a a clear rationale as to why cooperative efforts are essential for maximizing learning and ensuring healthy cognitive and social development as well as many other important instructional outcomes. Hundreds of research studies demonstrate that cooperative efforts result in higher individual achievement than do competitive or individualistic efforts. Educators use cooperative learning throughout North America, Europe, and many other parts of the world. This combination of theory, research, and practice makes cooperative learning one of the most distinguished of all instructional practices.

Knowing that cooperative learning can significantly increase student achievement (compared with competitive and individualistic learning) when properly implemented does not mean, however, that all operationalizations of cooperative learning will be effective or that all operationalizations will be equally effective. Without reviewing the research on the different cooperative learning methods, it is difficult to recommend specific cooperative learning procedures to educators. This meta-analysis, therefore, focuses on four issues: (a) determining how much research has been conducted on cooperative learning methods, (b) determining how many different cooperative learning methods have been evaluated, (c) determining how effective each method evaluated is in maximizing student achievement, and (d) determining the characteristics of the more effective cooperative learning methods.

The first issue was to determine the amount of research that has been conducted on cooperative learning methods. One-hundred-sixty-four studies on specific cooperative learning methods were found. This is a substantial number of studies, especially considering that 28 percent of them have been conducted since 1990 and 100 percent have been conducted since 1970. The studies have been conducted at all levels of schooling (46 percent were conducted in elementary schools, 20 percent were conducted in middle schools, 11 percent were conducted in high schools, and 24 percent were conducted in post-secondary and adult settings) and the majority lasted for considerable time (46 percent lasted for 30 sessions or more, 52 percent lasted for 2 to 29 sessions, and 2 percent of the studies lasted only for one session). Most of the studies used good to excellent methodology (45 percent randomly assigned participants to conditions, 25 percent randomly assigned groups to conditions, and only 30 percent did not randomly assign participants or groups to conditions). The research has been conducted in North America, Europe, the Middle East, Asia, and Africa and has involved minority as well as majority populations. Thus, there is considerable research on specific cooperative learning methods and the research has considerable validity and generalizability. As with the overall research, educators can have a great deal of confidence in the effectiveness of cooperative learning.

The second issue investigated was to determine how many different cooperative learning methods have been evaluated. Of all the numerous ways that cooperative learning is used, only eight methods have been subjected to empirical validation in a way that a relevant effect size could be computed. Of these methods, some have more empirical support than others. The more research studies conducted on any method, the more valid and reliable the results can be expected to be. There are 118 independent effects in the studies on Learning Together and Constructive Controversy,

Table 3

Meta-Analysis Results For Cooperative Learning Methods

	Average Effect Sizes			Weighted Effect Sizes							
	Effect	*SD*	*k*	Effect	*SE*	*k*	*Cld* 95%	*fsn*	*Qw*	*p* value	*df*
Learning Together											
Cooperation vs. Competition	0.82	0.50	25	0.70	0.06	25	±0.12	62	54.23	0.00	24
Cooperative vs. Individualistic	1.03	0.69	56	0.91	0.04	56	±0.08	200	188.66	0.00	55
Competitive vs. Individualistic	0.06	0.47	10	0.08	0.10	10	±0.19	0	15.76	0.07	9
TGT											
Intergroup Comp vs. Competition	0.48	0.69	9	0.47	0.05	9	±0.10	12	141.30	0.00	8
Intergroup Comp vs. Individualistic	0.58	0.43	5	0.55	0.11	5	±0.22	9	10.60	0.03	4
Group Investigation											
Cooperation vs. Competition	0.37	1.19	2	0.86	0.14	2	±0.27	7	24.73	0.00	1
Cooperation vs. Individualistic	0.62	0.00	1	0.62	0.44	1	±0.86	2	0.00	0.00	0
Academic Controversy											
Cooperative vs. Competition	0.59	0.44	16	0.61	0.07	16	±0.14	32	36.82	0.00	15
Cooperative vs. Individualistic	0.91	0.59	11	0.86	0.10	11	±0.19	36	22.08	0.01	10
Jigsaw											
Cooperation vs. Competition	0.29	0.78	9	0.41	0.05	9	±0.11	9	68.46	0.00	8
Cooperation vs. Individualistic	0.13	0.29	5	0.09	0.11	5	±0.21	0	4.86	0.30	4
STAD											
Intergroup Comp vs. Competition	0.51	0.72	15	0.46	0.05	15	±0.09	19	205.32	0.00	14
Intergroup Comp vs. Individualistic	0.29	0.71	14	0.28	0.07	14	±0.14	6	53.89	0.00	13

Table 3 (continued)

Meta-Analysis Results for Cooperative Learning Metods

	Average Effect Sizes			Weighted Effect Sizes							
	Effect	*SD*	*k*	Effect	*SE*	*k*	*Cld* 95%	*fsn*	*Qw*	*p* value	*df*
TAI											
Cooperative vs. Competitive	0.25	0.14	7	0.19	0.04	7	±0.09	0	4.92	0.55	6
Cooperative vs. Individualistic	0.33	0.26	8	0.19	0.06	8	±0.12	0	11.57	0.12	7
Competitive vs. Individualistic	-0.08	0.52	2	-0.32	0.13	2	±0.25	0	5.00	0.03	1
CIRC											
Cooperation vs. Competition	0.18	0.23	7	0.20	0.04	7	±0.07	0	13.43	0.04	6
Cooperation vs. Individualistic	0.18	0.00	1	0.18	0.00	1	±0.22	0	0.00		0

Note: *SD* = standard deviation; *k* = the number of averaged effect sizes in the meta-analysis; *SE* = standard error; *Cld* 95% = the value of the 95% confidence interval around the weighted effect size; *fsn* = fail safe *N* (the number of additional studies needed to change the significance of the results below 0.20).

Table 4

Ranking Of Cooperative Learning Methods: Weighted Effect Sizes

Method	Coop v Comp	*n*	Method	Coop v Ind	*n*
GI	0.86	2	LT	0.91	56
LT	0.70	25	AC	0.86	11
AC	0.61	16	GI	0.44	1
STAD	0.46	15	STAD	0.28	14
Jigsaw	0.41	9	TAI	0.19	8
TAI	0.25	7	CIRC	0.18	1
CIRC	0.20	7	Jigsaw	0.11	5
TGT	0.19	9	TGT	0.11	5

Note: LT = Learning Together, AC = Academic Controversy, STAD = Student Teams Achievement Divisions, TGT = Teams-Games-Tournament, GI = Group Investigation, TAI = Team-Assisted-Instruction, CIRC = Cooperative Integrated Reading and Composition.

Table 5

Rating Of Direct-Conceptual Nature Of Cooperative Learning Methods

Method	Learn	Initial Use	Maintain	Robust	Adaptability	Total
Learning Together	5	5	5	5	5	25
TGT	3	3	1	2	2	11
Group Investigation	5	5	3	2	2	17
Academic Controversy	5	5	5	4	4	23
Jigsaw	2	2	3	3	3	13
STAD	2	2	1	2	2	9
TAI	2	2	1	1	1	7
CIRC	2	2	1	1	1	7
Complex Instruction	5	5	3	3	3	19
Cooperative Structures	1	1	1	1	5	9

Note. Methods of cooperative learning were evaluated on five dimensions: (a) ease of learning the method, (b) ease of initial use in the classroom, (c) ease of long-term maintenance of use of the method, (d) robustness of the method (applicability to a wide variety of subject areas and grade levels), and (e) ease of method's adapting to changing conditions.

Table 6

Linear Regression Coefficients: Predictors of Impact of Cooperative Learning

Model		Unstandardized Coefficients		Standardized Coefficients	*t*	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.919	11.202		.439	.661
	Random Assignment	-1.389E-02	.057	-.020	-.245	.807
	Age	-5.071E-02	.029	-.138	-1.765	.079
	Type Cooperation	1.407E-02	.017	.062	.812	.418
	Condition	-6.329E-03	.058	-.008	-.109	.913
	Duration	6.056E-02	.067	.088	.910	.364
	Gender	-1.052E-02	.172	-.004	-.061	.951
	Year Published	-2.579E-03	.006	-.034	-.457	.648
	Direct-Conceptual	4.077E-02	.006	.583	6.780	.000

a Dependent Variable: Effect Size
b Weighted Least Squares Regression - Weighted by WEIGHT

68 independent effects in the studies on the cooperative learning methods developed at Johns Hopkins University, 14 independent effects in the studies on the Jigsaw Procedure, and 3 independent effects in the studies on the Group Investigation Method. It is somewhat surprising that so few methods have been evaluated. While any teacher may develop a version of cooperative learning that is very effective, without research studies it is unknown whether other teachers can expect reliable results when the method is used. The unevaluated cooperative learning methods, therefore, should be used with some caution. In addition, there is a need for a new generation of researcher-developers who formulate new operationalizations of cooperation for classroom and school use and who subject their formulations to rigorous empirical evaluation.

The third issue investigated is the effectiveness of the different cooperative learning methods researched. There is no reason to expect that all operationalizations of cooperation will be effective. While the largest effect sizes were found for the Learning Together, Group Investigation, and Constructive Controversy Methods, all of the methods have substantial effect sizes and all of the methods have been found to produce significantly higher achievement than did competitive or individualistic learning. Any teacher should feel quite comfortable using any of these eight cooperative learning methods.

The diversity of the eight cooperative learning methods provides additional validation of the effectiveness of cooperative learning. The methods range from specific procedures (such as Jigsaw and CIRC) to conceptual frameworks educators use to build their own cooperative lessons (such as Learning Together and Group Investigation) to curriculum packages in which cooperative learning is a central part (such as TAI and STAD), to rather complex procedures that require some sophistication to use (such as Constructive Controversy). That all of these methods are effective in increasing achievement is a tribute to the power of cooperation.

The fourth issue investigated was the characteristics of the different cooperative learning methods. Among the researcher-developers of cooperative learning, there are those who believe that the best way to ensure implementation of cooperative learning is to devise very specific techniques that teachers can learn in a few minutes and apply immediately (direct approach) and those who believe that teachers must learn a conceptual system and use it to adapt current lessons and activities into cooperative ones (conceptual approach). Previous research indicates that direct methods may be easier to learn and implement than are conceptual methods, but once implemented, conceptual methods are more robust and are more frequently maintained over time and easier to adapt to changing conditions and circumstances (Antil, et al., 1998; Berman, 1980; Berman & McLaughlin, 1976; Fullan, 1981; Griffin & Barnes, 1984; Johnson, 1970, 1979; Johnson, Druckman, & Dansereau, 1994; Johnson & Johnson, 1994a, 1994b; Joyce & Showers, 1980, 1982; Smith & Keith, 1982). There is very little research, however, on whether direct or conceptual methods differentially affect achievement and productivity. The results of this meta-analysis indicate that the more conceptual the method of cooperative learning, the

greater its impact on student achievement tends to be. This is an important addition to the literature on implementation and institutionalization of innovations. Differences in the way achievement was measured, however, make these findings tentative. Further research is needed to corroborate this finding.

It seems reasonable to hypothesize that the effectiveness of a cooperative learning method will tend to increase the more that cooperation is the foundation on which classroom and school life is based. If cooperative learning is used within a primarily competitive or individualistic school, for example, its effectiveness may be dampened by the overall culture of the school. Two of the cooperative learning methods have been extended to the overall organizational structure of the school. The Learning Together method has been adapted to include faculty interactions as well as student interactions and is known as the *Cooperative School* (Johnson & Johnson, 1994). School leaders are trained to implement a cooperative structure in colleagial teaching teams, faculty study groups, task forces, site-based decision making, and cooperative faculty meetings. The procedures have been used in elementary, middle, secondary schools and institutions of higher education. The Johns Hopkins cooperative learning methods have been extended into a schoolwide program for elementary schools known as *Success For All* (Slavin, et al., 1996). The extension of cooperation to the overall school structure is a promising area for future research.

There is no reason to expect the different methods of cooperative learning to interfere with each other's effectiveness. All the methods may be used in the same classroom and school. A teacher, for example, may use TAI in math, Learning Together in science and language arts, and Group Investigation in social studies and expect that the different methods will enhance and enrich each other's effectiveness. There is currently, however, no research on the ways in which the different methods of cooperative learning may enhance or interfere with each other's effectiveness.

The results of this meta-analysis provide evidence that (a) considerable research has been conducted on cooperative learning methods, (b) eight diverse methods have been researched, (c) all methods have produced higher achievement than competitive and individualistic learning, and (d) the more conceptual approaches to cooperative learning may produce higher achievement than the direct methods. These conclusions are all the stronger due to the diversity of the research on which they are based, ranging from controlled field experimental studies to evaluational case studies.

Based on the amount and diversity of the research, several conclusions about the effectiveness of the cooperative learning methods may be made. First, there is an obvious need for more research on the various cooperative learning methods. Research has been conducted only on eight of the cooperative learning methods and on several of these methods only a few studies have conducted. There are many, many cooperative learning methods on which no research has been conducted. Hopefully this review will spark more interest in assessing the effectiveness of the wide variety of cooperative learning methods.

Second, the differences among the methods need to be clarified and research conducted to identify the characteristics that are related to effectiveness. The characteristic focused on in this review was how direct or conceptual a cooperative learning method is. This is not the only way to characterize the differences among the cooperative learning methods. Further research needs to be conducted that identifies other important differences among methods and assesses their impact on effectiveness.

Third, it is reasonable to assume that no one cooperative learning method is effective under all conditions. Considerable more research is needed to identify the conditions under which the various methods are effective. More direct methods, for example, tend to be easy to learn and initially use, which make them appealing for brief staff development sessions. More conceptual methods, on the other hand, once implemented tend to be easier to maintain long-term and adapt to changing circumstances, which make them appealing for long-term staff development programs. Conclusions such as these, however, are actually based on research on innovation in education that is somewhat inconsistent and of widely varying quality. More precise understanding of the conditions under which the different methods are effective can only come from additional research.

Despite the amount and diversity of the research, several limitations and cautions about the research on the effectiveness of the cooperative learning methods may be made. First, the conclusions are limited by the small number of validating studies on several of the methods. The more studies conducted on a method, the more accurate the effect size may be. Conclusions about methods that have only a few validating studies may be misleading.

Second, while future research is needed, conducting research to compare directly the effectiveness of different cooperative learning methods is not very helpful. Studies in which two or more methods of cooperative learning are directly compared are difficult to interpret, especially if they are conducted by a researcher-developer who has a vested interest in one of the methods. It is virtually impossible to implement different methods at exactly the same strength. If one method is strongly implemented and another method is weakly implemented the resulting differences would be due to the strength of the implementation, not the differences in effectiveness between the methods.

Third, the differences in effect sizes for the different cooperative learning methods should be interpreted cautiously. The measures of academic achievement in various studies may not be equivalent. Lower effect sizes, for example, would be expected on standardized tests than on nonstandardized tests. Methods of cooperative learning aimed at lower-level tasks may produce high effect sizes on simple recognition level tests than methods of cooperative learning aimed at higher-level reasoning and critical thinking.

Thus, a lower effect size may be due to the type of measure of academic achievement or the match between the method and the dependent measure, not the overall effectiveness of the method.

Fourth, most of the validating studies on methods of cooperative learning have been conducted by the researcher-developer who originated the method. This introduces potential bias into the results. Ancient Romans advised individuals to ask, "cui bono" (who benefits) and the researcher-developer often has interests at stake that may bias his or her results toward confirming the effectiveness of his or her program. More studies conducted by independent investigators are needed.

Finally, many of the studies conducted on the impact of cooperative learning methods on achievement have methodological shortcomings and, therefore, any differences found could be the result of methodological flaws rather than the cooperative learning method. In the future, researchers should concentrate on conducting highly controlled studies that add to the confidence with which their conclusions will be received.

References

Antil, L., Jenkins, J., Wayne, S., & Vadasy, P. (1998). Cooperative learning: Prevalence, conceptualizations, and the relation between research and practice. *American Educational Research Journal, 35*(3), 419-454.

Aronson, E., Blaney, N., Sikes, J., Stephan, C., & Snapp, M. (1978). *The jigsaw classroom.* Beverly Hills, CA: Sage.

Aronson, E., & Patnoe, S. (1997). *The jigsaw classroom* (2nd ed.). New York: Longman.

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations.* Santa Monica, CA: Rand Corporation.

Bandura, A. (1977 of). *Principles behavioral modification.* New York: Holt, Rinehart, & Winston.

Cohen, E. (1994a). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*, 1-35.

Cohen, E. (1986/1994b). *Designing groupwork: Strategies for the heterogeneous classroom.* New York: Teachers College Press.

Cohen, E., & Lotan, R., (Eds.). (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice.* New York: Teachers College Press.

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence-Erlbaum.

Coleman, J. (1961). The adolescent society. New York: Macmillan.

Cooper, H. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Newbury Park, CA: Sage.

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations, 2*, 129-152.

Deutsch, M. (1962). Cooperation and trust: Some theoretical notes. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 275-319). Lincoln, NE: University of Nebraska Press.

Deutsch, M. (1973). *The resolution of conflict.* New Haven, CT: Yale University Press.

DeVries, D., & Edwards, K. (1974). Student teams and learning games: Their effects on cross-race and cross-sex interaction. *Journal of Educational Psychology, 66*, 741-749.

Fullan, M. (1990). Staff development innovation and institutionalization development. *Changing school culture through staff development: The 1990 ASCD Yearbook.* Alexandria, VA: Association for Supervision and Curriculum Development.

Fullan, M. (1993). *Change forces: Probing the depths of educational reform.* New York: Falmer Press.

Glass, G., McGaw, B., & Smith, M. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage Publications.

Gillies, R., & Ashman, A. (1998). Behavior and interactions of children in cooperative groups in lower and middle elementary grades. *Journal of Educational Psychology, 90*(4), 746-757.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hunter, J., Schmidt, F., & Jackson, G. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Johnson, D. W. (1970). *Social psychology of education.* New York: Holt, Rinehart, & Winston.

Johnson, D. W. (1979). *Educational psychology.* Englewood Cliffs,. NJ: Prentice Hall.

Johnson, D. W., Druckman, D., & Dansereau, D. (1994). Training in teams. In D. Druckman, R. Bjork, et al., (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 140-170). Washington, D. C.: National Academy Press.

Johnson, D. W., & Johnson, R. (1974). Instructional goal structure: Cooperative, competition, or individualistic. *Review of Educational Research, 44*, 213-240.

Johnson, D. W., & Johnson, R. (Eds.). (1978). Social interdependence within instruction. *Journal of Research and Development in Education, 12*(1).

Johnson, D. W., & Johnson, R. (1979). Conflict in the classroom: Controversy and learning. *Review of Educational Research, 49*, 51-70.

Johnson, D. W., & Johnson, R. (1989). *Cooperation and competition: theory and research.* Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R. (1994a). Professional development in cooperative learning: Short-term popularity vs. long-term effectiveness. *Cooperative Learning, 14*(2), 52-54.

Johnson, D. W., & Johnson, R. (1994b). Successful professional development sessions: Laying the groundwork first. *Cooperative Learning, 14*(4), 53-54.

Johnson, D. W., & Johnson, R. (1998). Cooperative learning and social interdependence theory. In R. Tindale, L. Heath, J. Edwards, E. Posavac, F. Bryant, Y. Suzrez-

Balcazar, E. Henderson-King, & J. Myers (eds.), *Theory and research on small groups* (pp. 9-36). New York: Plenum. Social Psychological Applications To Social Issues, Volume 4.

Johnson, D. W., & Johnson, R. (1975/1999a). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Boston: Allyn & Bacon. First Edition, 1975.

Johnson, D. W., & Johnson, R. (1999b). *Human relations: Valuing diversity*. Edina, MN: Interaction Book Company.

Johnson, D. W., Johnson, R., & Holubec, E. (1998). *Cooperation in the classroom*, (7th ed.). Edina, MN: Interaction Book Company.

Johnson, D. W., Johnson, R., & Maruyama, G. (1983). Interdependence and interpersonal attraction among heterogeneous and homogeneous individuals: A theoretical formation and a meta-analysis of the research. *Review of Educational Research, 53*, 5-54.

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages f research. *Educational Leadership, 37*(5), 379-385.

Joyce, B., Weil, M., & Showers, B. (1992). *Models of teaching*. Boston: Allyn & Bacon.

Kagan, S. (1985). *Cooperative learning resources for teachers*. Riverside, CA: University of California at Riverside.

Kohn, A. (1992). *No contest* (2nd ed.). Boston: Houghton Mifflin.

Mead, M. (Ed.). (1936/1961). *Cooperation and competition among primitive peoples*. Boston: Beacon.

Piaget, J. (1950). *The psychology of intelligence*. New York: Harcourt.

Sharan, S. (1980). Cooperative learning in teams: Recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research, 50*, 241-272.

Sharan, S., & Sharan, Y. (1976). *Small group teaching. Englewood Cliffs*, NJ: Educational Technology Publications.

Sharan, S., & Sharan, Y. (1992). *Group investigation: Expanding cooperative learning*. New York: Teacher's College Press.

Slavin, R. (1977). Classroom reward structure: An analytical and practical review. *Review of Educational Research, 47*, 633-650.

Slavin, R. (1978). Student teams and achievement divisions. *Journal of Research and Development in Education, 12*, 39-49.

Slavin, R. (1986). *Using student team learning* (3rd Edition). Baltimore: Johns Hopkins University.

Slavin, R. (1991). Group rewards make groupwork work. *Educational Leadership, 5*, 89-91.

Slavin, R., Leavey, M., & Madden, N. (1986). *Team accelerated instruction: Mathematics*. Watertown, MA: Charlesbridge.

Smith, A. (1759). *The theory of moral sentiments*. Reprint. Edited by D. Raphael & A. Maclie. (1976). Oxford: Clarendon.

Stevens, R., Madden, N., Slavin, R., & Farnish, A. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly, 22*, 433-454.

Von Mises, L. (1949). *Human action: A treatise on economics*. New Haven, CN: Yale University Press.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

David W. Johnson is a Professor of Educational Psychology at the University of Minnesota. He specializes in cooperation and competition, conflict resolution, and organizational change.

Roger T. Johnson is a Professor of Education at the University of Minnesota. He specializes in science education.

They are co-directors of the Cooperative Learning Center.

Appendix A: Complete List Of Studies Used In Meta-Analysis

Allen, W., & VanSickle, R. (1984). Learning teams and low achievers. *Social Education, 48*, 60-64.

Ames, G., & Murray, F. (1982). When two wrongs make a right: Promoting cognitive change by social conflict. *Developmental Psychology, 18*(6), 894-897.

Armstrong, B., Johnson, D. W., & Balow, B. (1981). Effects of cooperative vs individualistic learning experiences on interpersonal attraction between learning-disabled and normal-progress elementary school students. *Contemporary Educational Psychology, 6*, 102-109.

Armstrong, D. (1998). The effect of student team achievement divisions cooperative learning technique on upper secondary social studies students' academic achievement and attitude towards social studies class (Doctoral dissertation, University of Southern Mississippi, 1997). *Dissertation Abstracts International, 59*/02, 405-A.

Artzt, A. (1983). The comparative effects of the student-team method of instruction and the traditional teacher-centered method of instruction upon student achievement, attitude, and social interaction in high school mathematics course (Doctoral dissertation, New York University). *Dissertation Abstracts International, 44*(12), 3619-A.

Baker, L.J. (1995). The effect of cooperative study groups on achievement of college-level computer science programming students. (Doctoral dissertation, University of Texas, Austin). *Dissertation Abstracts International, 56*(06), 2185.

Bilbao, P.P. & Dominguez, E.D. (1997). Comparative effectiveness of small group discussion and individualized written report as post laboratory task in high school physics learning. In Aranador, L.C., Valencia, I.N. & Vui, T. (Eds.) *Proceedings: 1997 International Conference on Cooperative Learning and Constructivism* (pp. 3-1 - 3-13). Penang, Malaysia: SEAMEO Regional Centre for Education in Science and Mathematics.

Bonaparte, E.C. (1990). The effects of cooperative versus competitive classroom organization for mastery learning on the mathematical achievement and self esteem of urban second-grade pupils (Doctoral dissertation, Morgan State University, 1989). *Dissertation Abstracts International, 50*/07, 1911-A.

Botvin, G. J., & Murray, F. B. (1975). The efficacy of peer modeling and social conflict in the acquisition of conservation. *Child Development, 46*, 796-799.

Bryant, R. R. (1982). Effects of team-assisted individualization on the attitudes and achievement of third, fourth, and fifth grade students of mathematics (Doctoral dissertation, University of Maryland, College Park, 1981). *Dissertation Abstracts International, 43*/01, 70-A.

Carroll, E. J. R. (1990). A comparative study of achievement in college-level written business communication using lecture and cooperative learning teaching methods. (Doctoral dissertation, Oklahoma State University, 1989). *Dissertation Abstracts International, 50*/12, 3832A.

Cavanagh, B. (1985). Effects of interdependent group contingencies on the achievement of elementary school children (Doctoral dissertation, University of Maryland, 1984). *Dissertation Abstracts International, 46*/06, 1558-A.

Dees, R. (1991). The role of cooperative learning in increasing problem-solving ability in a college remedial course. *Journal for Research in Mathematics Education, 22(5)*, 409-421.

DeVries, D., Edwards, K., & Wells, E. (1974a). *Team competition effects on classroom group process* (Tech. Rep. No. 174). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

DeVries, D. L., Lucasse, P. R., & Shackman, S. L. (1980). *Small group versus individualized instruction: A field test of their relative effectiveness* (Tech. Report No. 293). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

DeVries, D. L., & Mescon, I. T. (1975). *Teams-games-tournament: An effective task and reward structure in the elementary grades* (Tech. Rep. No. 189). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

DeVries, D. L., Mescon, I. T., & Shackman, S. L. (1975). *Teams-Games- Tournament in the elementary classroom: A replication* (Tech. Rep. No. 190). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

DeVries, D. L., Mescon, I. T., & Shackman, S. L. (1976). *Student teams can improve basic skills: TGT applied to reading*. Paper presented at the American Psychological Association Convention, Washington, DC.

Doise, W. & Mugny, G. (1979). Individual and collective conflicts of centrations in cognitive development. *European Journal of Psychology, 9*, 105-108.

Dudley, B., Johnson, D.W., & Johnson, R. (1996). Conflict resolution training and middle school students' integrative negotiation behavior. *Journal of Applied Social Psychology, 26*(22), 2038-2052.

Edwards, K., & DeVries, D. (1972). *Learning games and student teams: their effects on student attitudes and achievement* (Tech.Report No. 147). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Edwards, K. J., DeVries, D. L., & Snyder, J. P. (1972). Games and teams: A winning combination. *Simulation and Games, 3*(3),247-269.

Emley, W. P. (1987). The effectiveness of cooperative learning versus individualized instruction in a college level remedial mathematics course, with relation to attitudes toward mathematics and Myers-Briggs personality type (Doctoral dissertation, University of Maryland, 1986). *Dissertation Abstracts International, 48*/01, 70A

Falk, D. R. & Johnson, D. W. (1977). The effects of perspective-taking and egocentrism on problem solving in heterogeneous and homogeneous groups. *The Journal of Social Psychology, 102*, 63-72.

Frank, M. J. (1984). A comparison between an individual and group goal structure contingency that differed in the behavioral contingency and performance-outcome components (Doctoral dissertation, University of Minnesota). *Dissertation Abstracts International, 45*/05, 1341-A.

Gabbert, B., Johnson, D. W., & Johnson, R. (1986). Cooperative learning, group-to-individual transfer, process gain, and the acquisition of cognitive reasoning strategies. *Journal of Psychology, 120*(3), 265-278.

Garibaldi, A. M. (1979). Affective contributions of cooperative and group goal structures. *Journal of Educational Psychology, 71*(6), 788-794.

Giraud, G. (1996). *Cooperative learning and statistics instruction*. Unpublished manuscript, University of Nebraska at Lincoln.

Hall, J. & Williams, M. (1970). Group dynamics training and improved decision making. *Journal of applied Behavioral Science, 6*(1), 39-68.

Holliday, D.C. (1995). The effects of the cooperative learning strategy Jigsaw II on academic achievement and cross-race relationships in a secondary social studies classroom (Doctoral dissertation, University of Southern Mississippi). *Disseration Abstracts International, 57*(01), 0087-A.

Holubec, E., Johnson, D. W., & Johnson, R. (1993). Impact of cooperative learning on naval air traffic controller

training. *Journal of Social Psychology, 133*, 337-346.

Huber, G. L., Sorrentino, R. M., Davidson, M. A, Epplier, R., et al. (1992). Uncertainty orientation and cooperative learning: Individual differences within and across cultures. *Learning and Individual Differences, 4*(1), 1-24.

Hulten, B. H. & DeVries, D. L. (1976). *Team competition and group practice effects on student achievement and attitudes* (Tech. Rep. No. 212). Baltimore: Johns Hopkins University, Center for Social Organization of Schools. (ERIC Document Reproduction Service No. ED 154 021)

Humphreys, B. , Johnson, R. T., & Johnson, D. W. (1982). Effects of cooperative, competitive and individualistic learning on students' achievement in science class. *Journal of Research in Science Teaching, 19*(5), 351-356.

Hwong, N., Caswell, A., Johnson, D. W., & Johnson, R. T. (1993). Effects of cooperative and individualistic on prospective elementary teachers' music achievements and attitudes. *Journal of Social Psychology, 133*(1), 53-64.

Jenkins, J.R., Jewell, M., Leicester, N., Jenkins, L. & Troutner, N.M. (1991). Development of a school building model for educating students with handicaps and at-risk students with handicaps in general education classrooms. *Journal of Learning Disabilities, 24*(5), 311-320.

Jenkins, J., Jewell, M., Leicester, N., O'Connor, R.E., Jenkins, L.M., Troutner, N.M. (1994). Accommodations for individual differences without classroom ability groups: An experiment in school restructuring. *Exceptional Children, 60*(4), 344-358.

Jensen, M. (1996). Cooperative quizzes in the anatomy and physiology laboratory: A description and evaluation. *Advances in Physiology Education, 16*(1), S48-S54.

Jensen, M., Johnson, D. W., & Johnson, R. (2002). The impact of positive interdependence during electronic quizzes on discourse and achievement. *Journal of Educational Research*, 95 (3), 161-166.

Johnson, D. W., & Johnson, R. T. (1981a). The integration of the handicapped into the regular classroom: Effects of cooperative and individualistic instruction. *Contemporary Educational Psychology, 6*, 344-353.

Johnson, D. W., & Johnson, R. T. (1981b). Effects of cooperative and individualistic learning experiences on inter-ethnic interaction. *Journal of Educational Psychology, 73*(3), 444-449.

Johnson, D. W., & Johnson, R. T. (1982a). The effects of cooperative and individualistic instruction on handicapped and nonhandicapped students. *Journal of Social Psychology, 118*, 257-268.

Johnson, D. W., & Johnson, R. T. (1984b). Building acceptance of differences between handicapped and nonhandicapped students: The effects of cooperative and individualistic instruction. *Journal of Social Psychology, 122*, 257-267.

Johnson, D. W., & Johnson, R. T. (1985a). Classroom conflict: controversy versus debate in learning groups. *American Educational Research Journal, 22*(2), 237-256.

Johnson, D. W., & Johnson, R. T. (1985b). Mainstreaming hearing-impaired students: the effect of effort in communicating on cooperation and interpersonal attraction. *Journal of Psychology, 119*(1), 31-44.

Johnson, D. W., Johnson, R., Johnson, J., & Anderson, D. (1976). Effects of cooperative versus individualized instruction on student prosocial behavior, attitudes toward learning, and achievement. *Journal of Educational Psychology, 68*(4), 446-452.

Johnson, D.W., Johnson, R., Dudley, B. & Acikgoz, K. (1994). Effects of conflict resolution training on elementary school students. *Journal of Social Psychology, 134*(6), 803-817.

Johnson, D.W., Johnson, R., & Dudley, B., & Magnuson, D. (1995). Training elementary school students to manage conflict. *Journal of Social Psychology, 135*(6), 673-686.

Johnson, D.W., Johnson, R., & Dudley, B., Mitchell, J. & Fredrickson, J. (1997). The impact of conflict resolution training on middle school students. *Journal of Social Psychology, 137*(1), 11-21.

Johnson, D. W., Johnson, R., Pierson, W., & Lyons, V. (1985). Controversy versus concurrence seeking in multi-grade and single-grade learning groups. *Journal of Research in Science Teaching, 22*(9), 835-848.

Johnson, D. W., Johnson, R., Roy, P. & Zaidman, B. (1985). Oral interaction in cooperative learning groups: speaking, listening, and the nature of statements made by high-, medium-, and low-achieving students. *Journal of Psychology, 119*(4), 303-321.

Johnson, D. W., Johnson, R., & Scott, L. (1978). The effects of cooperative and individualized instruction on student attitudes and achievement. *Journal of Social Psychology, 104*, 207-216.

Johnson, D. W., Johnson, R., & Skon, L. (1979). Student achievement on different types of tasks under cooperative, competitive, and individualistic conditions. *Contemporary Educational Psychology, 4*, 99-106.

Johnson, D. W., Johnson, R. T., Stanne, M. B. & Garibaldi, A. (1989). Impact of goal and resource interdependence on problem-solving success on a computer assisted task. *Journal of Social Psychology, 129*(5), 621-629.

Johnson, D.W., Johnson, R., & Taylor, B. (1993). Impact of cooperative and individualistic learning on high-ability students' achievement, self esteem, and social acceptance. *Journal of Social Psychology, 133*(6), 839-844.

Johnson, D. W., Johnson, R., & Tiffany, M. (1984). Structuring academic conflicts between majority and minority students: Hindrance or help to integration. *Contemporary Educational Psychology, 9*, 61-73.

Johnson, D. W., Johnson, R. T., Tiffany, M., & Zaidman, B. (1983). Are low achievers disliked in a cooperative situation? A test of rival theories in a mixed ethnic situation. *Contemporary Educational Psychology, 8*, 189-200.

Johnson, D. W., Skon, L., & Johnson, R. (1980). Effects of cooperative, competitive, and individualistic conditions on children's problem-solving performance. *American Educational Research Journal, 17*(1), 83-93.

Johnson, R. T., Bjorkland, R., & Krotee, M. L. (1984). The effects of cooperative, competitive and individualistic student interaction patterns on the achievement and attitudes of the golf skill of putting. *Research Quarterly, 55*(2), 129-134.

Johnson, R. T., Brooker, C., Stutzman, J., Hultman, D., & Johnson, D. W. (1985). The effects of controversy, concurrence seeking, and individualistic learning on achievement and attitude change. *Journal of Research in Science Teaching, 22*(3), 197-205.

Johnson, R. T., & Johnson, D. W. (1979). Type of task, and student achievement and attitudes in interpersonal cooperation, competition, and individualization. *Journal of Social Psychology, 108*, 37-48.

Johnson, R. T., & Johnson, D. W. (1981). Building friendships between handicapped and non-handicapped students: Effects of cooperative and individualistic instruction. *American Educational Research Journal, 18*(4), 415-423.

Johnson, R. T., Johnson, D. W., DeWeerdt, N., Lyons, V., & Zaidman, B. (1983). Integrating severely adaptively handicapped seventh-grade students into constructive relationships with nonhandicapped peers in science class. *American Journal of Mental Deficiency, 87*(6), 611-618.

Johnson, R., Johnson, D. W., Scott, L., & Ramolae, B. (1985). Effects of single-sex and mixed-sex cooperative interaction on science achievement and attitudes and cross-handicap and cross-sex relationships. *Journal of Research in Science Teaching, 22*(3), 207-220.

Johnson, R., Johnson, D. W., & Stanne, M. B. (1985). Effects of cooperative, competitive, and individualistic goal structures on computer-assisted instruction. *Journal of Educational Psychology, 77*(6), 668-677.

Johnson, R., Johnson, D. W., & Stanne, M. B. (1986). Comparison of computer-assisted cooperative, competitive, and individualistic learning. *American Educational Research Journal, 23*(3), 382-392.

Johnson, R., Johnson, D. W., & Tauer, M. (1979). The effects of cooperative, competitive, and individualistic goal structures on students' attitudes and achievement. *Journal of Psychology, 102*, 191-198.

Kambiss, P. (1990). *The effects of cooperative learning on student achievement in a fourth grade classroom* (Research Project Report, Mercer University, Atlanta Georgia). (ERIC Document Reproduction Service No. ED318 563)

Keeler, C. & Anson, R. (1995). An assessment of cooperative learning used for basic computer skills instruction in the college classroom. *Journal of Educational Computing Research, 12*(4), 379-393.

Kinney, J.H. (1989). *A study of the effects of a cooperative learning program on the achievement of ninth grade multi-cultural general biology classes.* Paper presented to the Alexandria City, Virginia School Board.(ERIC Document Reproduction Service No. ED 309 096)

Kosters, A.E. (1990). The effects of cooperative learning in the traditional classroom on student achievement and attitude (Doctoral dissertation, University of South Dakota). *Dissertation Abstracts International, 51*/07, 2255-A.

Lang, N.A. (1983). The effects of a cooperative learning technique, Teams-games-tornaments, on the academic achievement and attitude toward economics of college students enrolled in a principles of microeconomics course (Doctoral dissertation, University of Georgia). *Dissertation Abstracts International, 44*(05), 1517-A.

Lazarowitz, R., Baird, H., Bowlden, B., & Hertz-Lazarowitz, R. (1983, April). *Academic achievements, learning environment, self esteem, and inquiry skills of high school students in biology taught in cooperative-investigative small groups.* Paper presented at the National Association of Research in Science Teaching (NARST) Annual Meeting, The Abbey.

Lazarowitz, R., Hertz, R.L., Baird, J.H., & Bowlden, V. (1988). Academic achievement and on-task behavior of high school biology students instructed in a cooperative small investigative group. *Science Education, 72*(4), 475-487.

Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. (1986a). Positive interdependence, academic and collaborative skills, group contingencies, and isolated students. *American Educational Research Journal, 23*(3), 476-488.

Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. T. (1986b). Components of cooperative learning: Effects of collaborative skills and academic group contingencies on achievement and mainstreaming. *Contemporary Educational Psychology, 11*, 229-239.

Lowry, N., & Johnson, D. W. (1981). Effects of controversy on epistemic curiosity, achievement and attitudes. *Journal of Social Psychology, 115*, 31-43.

Lucker, G. W., Rosenfield, D., Sikes, J., & Aronson, E. (1976). Performance in the interdependent classroom: A field study. *American Educational Research Journal, 13*(2), 115-123.

Lynch, B. (1984). Cooperative learning in interdisciplinary education for the allied health professions. *Journal of Allied Health, 13*(2), 83-93.

Madden, N. A., & Slavin, R. E. (1983). Effects of cooperative learning on the social acceptance of mainstreamed academically handicapped students. *Journal of Special Education, 17*(2), 171-182.

Maitland, K.A. & Goldman, J.R. (1974). Moral judgment as a function of peer group interaction. *Journal of Personality and Social Psychology, 30*(5), 699-704.

Martino, L., & Johnson, D. W. (1979). Cooperative and individualistic experiences among disabled and normal children. *Journal of Social Psychology, 107*, 177-183.

Mehta, J.I. (1993). Cooperative learning in computer programming at the college level. (Doctoral dissertation, University of Illinois at Chicago, 1993). *Dissertation Abstracts International A-54* (04).

Mei, D. (1982). Grades as rewards: Distributive justice in four high schools (Doctoral dissertation, Columbia University, 1981). *Dissertation Abstracts International, 42*/09, 3875-B.

Merebah, S. A. (1987). Cooperative learning in science: A comparative study in Saudi Arabia ((Doctoral dissertation, Kansas State University, 1987). *Dissertation Abstracts International, 48*/04, 892A.

Mesch, D., Johnson, D. W., & Johnson, R. T. (1988). Impact of positive interdependence and academic group contingencies on achievement *Journal of Social Psychology, 128*(3), 345-352.

Mesch, D., Lew, M., Johnson, D. W., & Johnson, R. (1986). Isolated teenagers, cooperative learning and the training of social skills. *Journal of Psychology, 120*(4), 323-334.

Mitchell, J.M., Johnson, D.W. & Johnson, R.T. (March, 2000). Are all types of cooperation equal? Impact of academic controversy versus concurrence-seeking on AIDS, alcohol and tobacco education. Unpublished manuscript.

Morgan, B. (1988). Cooperative learning: teacher use, classroom life, social integration, and student achievement.(Doctoral dissertation, University of Southern California, 1987). *Dissertation Abstracts International, 48*/02, 3043-A.

Moskowitz, J. M., Malvin, J. H., Schaeffer, G. A., & Schaps, E. (1983). Evaluation of a cooperative learning strategy. *American Educational Research Journal, 20*(4), 687-696.

Moskowitz, J. M., Malvin, J. H., Schaeffer, G. A., & Schaps, E. (1985). Evaluation of jigsaw, a cooperative learning technique. *Contemporary Educational Psychology, 10*, 104-112.

Murray, F .B. (1972). Acquisition of conservation through social interaction. *Developmental Psychology, 6*(1), 1-6.

Nowak, T.A. (1996). Promoting academic achievement and social behaviors of children in kindergarten classes: The effects of cooperative learning (Doctoral dissertation, Lehigh University). *Dissertation Abstracts International, 57*(05), 1964-A.

Okebukola, P. A. (1985). Cooperative and competitive interaction techniques in strengthening students' performance in science classes. *Science Education, 69*(4), 501-509.

Ortiz, A., Johnson, D.W., & Johnson, R.T. (1996). The effect of positive goal and resource interdependence on individual performance. *Journal of Social Psychology, 136*(2), 243-249.

Overlock, T.H. (1994). *Comparison of the effectiveness of collaborative learning methods and traditional methods in physics classes at Northern Maine Technical College.* East Lansing, MI: National Center for Research on Teacher Learning. (ERIC Document Reproduction Service No. ED 367 394).

Peck, G.L. (1991). The effects of cooperative learning on the spelling achievement of intermediate elementary students (Doctoral dissertation, Ball State University, Muncie, Indiana). *Dissertation Abstracts International, 52*/06, 2022-A.

Perreault, R.J. (1984). Cooperative learning: Its effects on academic achievement in suburban junior high industrial arts classes. *Journal of Epsilon-Pi-Tau, 10*(1), 44-49.

Petersen, R. P.,Johnson, D. W., & Johnson, R. T. (1985). The effects of cooperative learning on perceived status of male and female pupils. *Journal of Social Psychology, 131*(5), 717-735.

Randolph, W.M. (1992). The effect of cooperative learning on academic achievement in introductory college biology. (Doctoral dissertation, Washington State University, 1992), *Dissertation Abstracts International, A-53* (08).

Ravid, R. & Shapiro, S .(1992). The use of cooperative learning in Jewish schools. *Journal of Research and Development in Education, 25*(2), 96-102.

Robertson, L. (1983). Integrated goal structuring in the elementary school: Cognitive growth in mathematics (Doctoral dissertation, Rutgers University, 1982). *Dissertation Abstracts International, 43*/08, 2549-A.

Roon, R. J., Van Pilsum, J. F., Harris, I, Rosenberg,, P. Johnson, R., Liaw, C., & Rosenthal, L. (1983). The experimental use of cooperative learning in a biochemistry class for first-year medical students. *Biochemical Education, 11*,(1), 12-16.

Ross, J. A. (1988). Improving social-environmental studies problem solving through cooperative learning. *American Educational Research Journal, 25*(4), 573-591.

Ross, S. T. (1985). *Effects of cooperative and individual goal structures on students achievement and attitudes toward science.* Unpublished masters thesis, University of Minnesota, Minneapolis.

Scott, T. J. (1990). The effects of cooperative learning team vs. traditional classroom/resource room instruction on handicapped student self esteem and academic achievement (Doctoral Dissertation, Boston College, 1989). *Dissertation Abstracts International, 50*/10, 3145-A.

Sengendo, A. B., (1988). The effects of computer-assisted cooperative learning on the science achievement and attitudes of American Indian students (Doctoral Dissertation, University of Kansas, 1987). *Dissertation Abstracts International, 49*:09, 1435-A.

Sharan, S., Kussell, P., Hertz-Lazarowitz, R., Bejarano, Y., Raviv, S., & Sharan, Y. (1985). Cooperative learning effects on ethnic relations and achievement in Israeli junior high school classrooms. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to Cooperate, Cooperating to Learn.* NY:Plenum Press, 313-344.

Sharan, S., Kussell, P., Hertz-Lazarowitz, R., Bejarano, Y., Shulamit, R., & Sharan, Y. (1984). *Cooperative learning in the classroom: Research in desegregated schools.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sharan, S. & Shachar, C. (1986). *Cooperative learning effects on students' academic achievement and verbal behavior in multi-ethnic junior high school classrooms in Israel.* Final Report Submitted to the Israel Trustees Foundation and to the Ford Foundation Trust.

Sherman, L. W., (1986). Cooperative vs competitive educational psychology classrooms: A comparative study. *Teaching & Teacher Education, 2*(3), 283-295).

Sherman, L. W. (1988). A comparative study of cooperative and competitive achievement in two secondary biology classrooms: The group investigation model versus an individually competitive goal structure. *Journal of Research in Science Teaching, 26*(1), 55-64.

Sherman, L. W. & Sosby, S. (1986, November). Cooperative learning in ninth grade remedial mathematics classes. In Sherman, L. W. (Chair), *Three comparative studies contrasting cooperative with competitive and individualistic goal structures.* Symposium conducted at the National Meeting of the School Science and Mathematics Association, Lexington, KY.

Sherman, L. W. & Thomas, M. (1986). Mathematics achievement in cooperative versus individualistic goal-structured high school classrooms. *Journal of Educational Research, 79*(3), 169-172.

Silverman, I. W., & Stone, J. M. (1972). Modifying cognitive functioning through participation in a problem-solving group. *Journal of Educational Psychology, 63*(6), 603-608.

Skon, L., Johnson, D. W., & Johnson, R. T. (1981). Cooperative peer interaction versus individual competition and individualistic efforts: Effects on the acquisition of cognitive reasoning strategies. *Journal of Educational Psychology, 73*(1), 83-92.

Slavin, R. E. (1977c). *Student learning teams and scores adjusted for past achievement: A summary of field experiments* (Tech. Rep. No. 227). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Slavin, R. E. (1978a). Student teams and comparison among equals: Effects on academic performance and student attitudes. *Journal of Educational Psychology, 70*(4), 532-538.

Slavin, R. E. (1978b). Effects of student teams and peer tutoring on academic achievement and time on-task. *Journal of Experimental Education, 48*(4), 252-257.

Slavin, R. E., & Karweit, N. A. (1981). Cognitive and affective outcomes of an intensive student team learning experience. *Journal of Experimental Education, 50*, 29-35.

Slavin, R. E., & Karweit, N. L. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal, 21*(4), 725-736.

Slavin, R. E., & Karweit, N. A. (1985). Effects of whole class, ability grouped, and individualized instruction on mathematics achievement. *American Educational Research Journal, 22*(3), 351-367.

Slavin, R. E., Leavey, M. B., & Madden, N. A. (1984). Combining cooperative learning and individualized instruction: Effects on student mathematics achievement, attitudes, and behaviors. *Elementary School Journal, 84*(4), 409-422.

Slavin, R. E., Madden, N. A., & Leavey, M. (1984a). Effects of cooperative learning and individualized instructions on mainstreamed students *Exceptional Children, 50*(5), 434-443.

Slavin, R. E., Madden, N. A., & Leavey, M. (1984b). Effects of team assisted individualization on the mathematics achievement of academically handicapped and nonhandicapped students. *Journal of Educational Psychology, 76*(5), 813-819.

Slavin, R. E., & Oickle, E. (1981). Effects of cooperative learning teams on student achievement and race relations: Treatment by race interactions. *Sociology of Education, 54*, 174-180.

Smith, K., Johnson, D. W., & Johnson, R. T. (1981). Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology, 73*(5), 651-663.

Smith, K., Johnson, D. W., & Johnson, R. (1982). Effects of cooperative and individualistic instruction on the achievement of handicapped, regular, and gifted students. *Journal of Social Psychology, 116*, 277-283.

Smith, K., Johnson, D. W., & Johnson, R. T. (1984). Effects of controversy on learning in cooperative groups. *Journal of Social Psychology, 122*, 199-209.

Smith, K. A., Petersen, R. P., Johnson, D.W., & Johnson, R.T. (1986). The effects of controversy and concurrence seeking on effective decision making. *Journal of Social Psychology. 126*(2), 237-248.

Smith, M.E., Hinckley, C.C., & Volk, G.L. (1991). Cooperative learning in the undergraduate laboratory. *Journal of chemical Education, 68*(5), 413-415.

Smith, M. J. (1984). A comparison of cooperative and individualistic learning in associate degree nursing students_ (Doctoral Dissertation, University of Minnesota). *Dissertation Abstracts International, 45*/09, 2739-A.

Stevahn, L., Johnson, D.W., Johnson, R.T., Green, K., & Laginski, A.M. (1997). Effects on high school students of conflict resolution training integrated into English literature. *The Journal of Social Psychology, 137*(3), 302-315.

Stevahn, L., Johnson, D.W., Johnson, R.T., Laginsky, A.M., & O'Coin, I. (1996). Effects on high school students of integrating conflict resolution skills and peer mediation training into an academic unit. *Mediation Quarterly, 14*(1), 21-36.

Stevahn, L., Johnson, D.W., Johnson, R., Oberle, K., & Wahl, L. (2000). Effects of conflict resolution training integrated

into a Kindergarten curriculum. *Child Development*, in press.

Stevahn, L., Johnson, D.W., Johnson, R., & Real, D. (1996). The impact of a cooperative or individualistic context on the effectiveness of conflict resolution training. *American Educational Research Journal, 33*(3), 801-823.

Stevahn, L., Johnson, D.W., Johnson, R.T., & Schultz,R. (1997). *Effects of conflict resolution training integrated into a high school social studies curriculum*. Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly, 22*(4), 433-454.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly, 22*(4), 433-454.

Stevens, R.J., & Slavin, R.E. (1995). Effects of a cooperative learning approach in reading and writing on academically handicapped students. *The Elementary School Journal, 95*(3), 241-262.

Stevens, R.J., & Slavin, R.E. (1995). The cooperative elementary school: Effects of student's achievement, attitudes, and social relations. *American Educational Research Journal, 32*(2), 321-351.

Stevens, R.J., Slavin, R., & Farnish, A.M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*(1), 8-16.

Supinnasri, S. (1997). Cooperative learning in upper secondary mathematics. In Aranador, L.C., Valencia, I.N. & Vui, T. (Eds.) *Proceedings: 1997 International Conference on Cooperative Learning and Constructivism* (pp. 18-1 - 18-14). Penang, Malaysia: SEAMEO Regional Centre for Education in Science and Mathematics.

Tanamai, A. (1990). A comparison of computer-assisted cooperative learning with independent learning. (Doctoral dissertation, University of Kansas, 1988), *Dissertation Abstracts International, 50*/11, 3471A.

Tjosvold, D. (1982a). Effects of cooperative and competitive interdependence and task complexity on subordinates' productivity, perception of leader, and group development. *Canadian Journal of Behavioural Science, 14*(1), 24-34.

Tjosvold, D. (1982b). Effects of approach to controversy on superiors' incorporation of subordinates' information in decision making. *Journal of Applied Psychology, 67*(2), 189-193.

Tjosvold, D. (1985b). Dynamics within participation: An experimental investigation. *Group and Organization Studies, 10*(3), 260-277.

Tjosvold, D. (1988). Effects of shared responsibility and goal interdependence on controversy and decision making between departments. *Journal of Social Psychology, 128*(1), 7-18.

Tjosvold, D., & Deemer, D. (1980). Effects of controversy within a cooperative or competitive context on organizational decision making. *Journal of Applied Psychology, 65*(5), 590-595.

Tjosvold, D., & Deemer, D. (1981). Effects of a control or collaborative orientation on participation in organizational decision making. *Canadian Journal of Behavioral Sciences, 13*(1), 33-43.

Tjosvold, D. & Field, R. (1983). Effects of social context on consensus and majority vote decision making. *Academy of Management Journal, 26*(3), 500-506.

Tjosvold, D., & Field, R. (1984). Effect of concurrence, controversy, and consensus on group decision making. *Journal of Social Psychology, 125*(3), 355-363.

Tjosvold, D., & Johnson, D. W. (1978). Controversy within a cooperative or competitive context and cognitive perspective-taking. *Contemporary Educational Psychology, 3*, 376-386.

Torrance, E. P. (1970). Influence of dyadic interaction on creative functioning. *Psychological Reports, 26*, 391-394.

Valentino, V.R. (1988). A study of the achievement, anxiety, and attitude toward mathematics in college algebra students using small-group interaction methods. (Doctoral dissertation, West Virginia University, 1988), *Dissertation Abstracts International, A-50* (02).

Vasquez, B., Johnson, D. W., & Johnson, R. (1993). Impact of cooperative learning on the performance and retention of U.S. Navy Air Traffic Controller Trainees. *Journal of Social Psychology, 133*, 769-783.

Webb, M.D. (1992). The effects of the jigsaw cooperative learning technique on racial attitudes and academic achievement (Doctoral dissertation, California State University, Fresno, 1988). *Masters Abstracts International, 32*/01, 0354.

Yager, S., Johnson, D. W,, & Johnson, R. T. (1985). Oral discussion, group-to-individual transfer, and achievement in cooperative learning groups. *Journal of Educational Psychology, 77*(1), 60-66.

Yager, S., Johnson, R. T. & Johnson, D. W., & Snider, B. (1986). The impact of group processing on achievement in cooperative learning groups. *Journal of Social Psychology, 126*(3), 389-397.

Yueh, J. & Alessi, S. M. (1988). The effect of reward structure and group ability composition on cooperative computer-assisted instruction. *Journal of Computer-Based Instruction, 15*(1), 18-22.

Ziegler, S. (1981). The effectiveness of cooperative learning teams for increasing cross-ethnic friendships: Additional evidence. *Human Organization, 40*(3), 264-269.

Teaching Students To Be Peacemakers: A Meta-Analysis

Roger T. Johnson
and
David W. Johnson
University of Minnesota

Concern about violence in schools has resulted in numerous conflict resolution and peer mediation programs being implemented on very little evidence that they are effective. The exception is the ***Teaching Students to be Peacemakers Program****. Between 1988 and 2000 we conducted sixteen studies on the effectiveness of conflict resolution training in eight different schools in two different countries. Students involved were from kindergarten through ninth grades. The studies were conducted in rural, suburban, and urban settings. The findings indicate that students learn the conflict resolution procedures taught, retain their knowledge throughout the school year, apply the conflict resolution procedures to actual conflicts, transfer the procedures to nonclassroom and nonschool settings, use the procedures similarly in family and school settings, and, when given the option, engage in problem-solving rather than win-lose negotiations. The studies demonstrate that conflict resolution procedures can be taught in a way that increases academic achievement.*

Why Teach Students To Be Peacemakers

A sizable proportion of students in many countries report feeling unsafe, fearful, or that their school has a violence problem (Kaufman, et al., 1999). Between nine and 16 percent of U.S. high-school students fear going to school because of violence. Providing students with an orderly environment in which to learn and even guaranteeing students' safety is becoming more and more difficult in many schools in the United States (Posner, 1994). An increasing number of public and private school teachers and administrators face situations involving serious conflicts among students and between students and faculty (Stop the Violence, 1994). In the United States, as well as other countries, the general public perceives school safety as the top problem facing U.S., schools (e.g., Elam & Rose, 1995; Elam, Rose, & Gallup, 1994, 1996; Everett & Price, 1995; Furlong & Morrison, 1994; Morrison, Furlong, & Morrison, 1997; Rose, Gallup, & Elam, 1997; Rose & Gallup, 1998). According to international surveys, a sizable percentage of elementary and secondary students rate their own schools as unsafe places (e.g., Benbenishty, Zeira, & Astor, 1998, 1999, 2000; Smith, 1999; Kaufman et al., 1999).

With the concern about violence in schools has come a proliferation of conflict resolution and peer mediation training programs. These programs have diverse origins, including researchers in the field of conflict resolution, advocates of nonviolence, anti-war activists, and members of the legal profession. Very few of these programs are based on theory and research in the field of conflict resolution. Almost none of the programs have been evaluated and very little research has been conducted on the effectiveness of teaching conflict resolution and peer mediation skills (Johnson & Johnson, 1996). Only one of the programs, **Teaching Students To Be Peacemakers** (Johnson & Johnson, 1995a), has generated a series of research studies. Between 1988 and 2000 sixteen studies on the effectiveness of the Peacemaker training were conducted in eight different schools in two different countries. The purpose of this article is to conduct a meta-analysis summarizing the results of these sixteen studies.

Teaching Students To Be Peacemakers Program

We began formulating the **Teaching Students to be Peacemakers** Program in the 1960s. It originated from:

1. Our research on integrative negotiations (Johnson, 1967), perspective taking in conflict situations (Johnson, 1967, 1971a), conflict resolution in the school (Johnson, 1970, 1971b; Johnson, Johnson, & Johnson, 1976), communication in conflict situations (Johnson, 1974), and constructive conflict (Deutsch, 1973; Johnson, 1970; Johnson & Johnson, 1979).

2. Social interdependence theory (Deutsch, 1949,

1962; Lewin, 1951; Johnson, 1970; Johnson & Johnson, 1989; Watson & Johnson, 1972).

3. Our training of elementary, junior-high, high school and college students and adults in how to manage conflicts constructively (Johnson, 1970, 1972/2000, 1978/1991, 1983; Johnson & F. Johnson, 1975/2000). We have implemented the peacemaker program in schools throughout North America, Central and South America, Europe, the Middle East, Asia, and Africa.

Students are taught to be peacemakers in five steps. **First,** they learn what is and is not a conflict and that conflicts potentially have many positive outcomes when they are managed constructively. **Second,** they are taught how to negotiate integrative agreements to conflicts of interests by:

1. **Describing what you want.** "I want to use the book now." This includes using good communication skills and defining the conflict as a small and specific mutual problem.

2. **Describing how you feel.** "I'm frustrated." Disputants must understand how they feel and communicate it accurately and unambiguously.

3. **Describing the reasons for your wants and feelings.** "You have been using the book for the past hour. If I don't get to use the book soon my report will not be done on time. It's frustrating to have to wait so long." This includes expressing cooperative intentions, listening carefully, separating interests from positions, and differentiating before trying to integrate the two sets of interests.

4. **Taking the other's perspective** and summarizing your understanding of what the other person wants, how the other person feels, and the reasons underlying both. "My understanding of you is..." This includes understanding the perspective of the opposing disputant and being able to see the problem from both perspectives simultaneously.

5. **Inventing three optional plans to resolve the conflict that maximize joint benefits.** "Plan A is..., Plan B is..., Plan C is..." This includes inventing creative optional agreements that maximize the benefits for all disputants and solve the problem.

6. **Choosing the wisest course of action to implement and formalizing the agreement with a hand shake.** "Let's agree on Plan B!" The agreement must maximize joint benefits and strengthen disputants' ability to work together cooperatively and resolve conflicts constructively in the future. It specifies how each disputant should act in the future and how the agreement will be reviewed and renegotiated if it does not work.

Third, they are taught how to mediate their classmates' conflicts by:

1. **Ending hostilities and cooling down disputants.** Usually disputants ask the mediator for help. In some cases the mediator may see individuals engaged in a dispute and ask if he or she can be of service.

2. **Ensuring disputants are committed to the mediation process.** To ensure that disputants are committed to the mediation process and are ready to negotiate in good faith, the mediator asks students if they want to solve the problem and does not proceed until both answer "yes." The mediator then explains that mediation is voluntary, he or she is neutral and will not take sides, each person will have a chance to speak without being interrupted, and each person must follow the a set of six rules [(1) agree to solve the problem, (2) no name calling, (3) do not interrupt, (4) be as honest as you can, (5) if you agree to a solution, you must abide by it (you must do what you have agreed to do) and (6) anything said in mediation is confidential].

3. **Helping disputants successfully negotiate with each other.** The mediator carefully takes the disputants through the integrative negotiation procedure.

4. **Formalizing the agreement.** The mediator formalizes the agreement by completing a **Mediation Report Form** and having disputants sign it and shake hands as a commitment to implement the agreement and abide by its conditions. The mediator becomes the keeper of the contract and checks back with the disputants a day or so later to see if the agreement is working.

Fourth, once students have completed the initial training, the teachers implement the Peacemaker Program. Each day the teacher selects two class members to serve as official mediators. Any conflicts students cannot resolve themselves are referred to the class mediators. The mediators work in pairs. They wear official T-shirts, patrol the playground and lunchroom, and are available to mediate

any conflicts that occur in the classroom or school. The role of class mediator is rotated throughout the class or school so that all students serve as class mediator an equal amount of time. If peer mediation fails, the teacher mediates the conflict. If teacher mediation fails, the teacher arbitrates by deciding who is right and who is wrong. If that fails, the principal mediates the conflict. If that fails, the principal arbitrates.

Fifth, students receive further training throughout the academic year to enhance and refine their negotiation and mediation skills.

The initial Peacemaker training typically lasts for ten to twenty hours and is usually spread out over several weeks and given to all students in a class or school. The Peacemaker Program is based on the assumptions that (a) all students should know how to negotiate integrative agreements to their conflicts and how to mediate schoolmates' conflicts, (b) all students should have the skills to use the negotiation and mediation procedures effectively, (c) the norms, values, and culture of the school should promote and support the use of the negotiation and mediation procedures, (d) peer mediators should be available to support and enhance students' efforts to negotiate, and (e) the responsibility for peer mediation should be rotated throughout the entire student body so that every student gains experience as a mediator.

Despite the implementation of a wide variety of conflict resolution and peer mediation programs, what is unknown is whether the training programs result in students actually learning negotiation and mediation procedures, applying the procedures in actual conflict situations, using the procedures in nonclassroom and nonschool settings, and retaining the procedures over time. It is also unknown whether the training can be integrated into academic subject areas in ways that increase achievement. We have conducted a systematic program of research to assess the effectiveness of the **Teaching Students To Be Peacemakers Program** and validate the theory on which it is based. The studies have focused on the following concerns:

1. Are schools justified in being concerned about the frequency and destructiveness of conflicts among students?

2. When given training, do students successfully learn the conflict resolution procedures?

3. Do students maintain their knowledge over time?

4. Do students apply the conflict resolution procedures to actual conflicts?

5. Do students transfer the conflict resolution procedures to nonclassroom and nonschool settings?

6. Do students use the conflict procedures in the family setting?

7. Do students achieve at a higher level when learning the conflict procedures in combination with academic learning?

8. Do students use a problem-solving approach to negotiations when placed in a situation where they can use either a win-lose or problem solving approach.

9. Do faculty, administrators, and parents perceive the conflict resolution in positive ways?

10. Will the integration of peacemaker training into academic units enhance or interfere with academic learning or the learning of the negotiation and mediation procedures?

11. Will the peacemaker training affect the frequency with which students use the integrative negotiation procedure when they are given a choice between seeking distributive or integrative outcomes?

12. Can the peacemaker training be initiated with very young children such as kindergartners?

Method Of Our Mediation Studies

Participants

Participants came from eight different schools in two different countries and were from kindergarten through the ninth grade (see Table 1).

Independent Variable

In the studies included in the meta-analysis, the independent variable was the presence versus the absence of conflict resolution training. In the **experimental condition,** students participated in the Peacemaker training. In the **control condition,** students did not receive the training. The training programs lasted from nine to fifteen hours in length. Eleven of the studies involved control groups. In seven of the studies, classrooms and/or controls were selected randomly from the school; in four studies students were assigned randomly to conditions. In nine of the studies teachers were rotated across conditions.

Dependent Measures

A variety of dependent variables were employed in the studies. Not every dependent variable was assessed in every study and different versions of a dependent version were used in different studies. One dependent variable was *academic achievement and retention*. It was measured by

Table 1

Summary of Conflict Resolution and Peer Mediation Studies

Study	Year	Setting	School	Grades	Training Length	Sample Size	Control	Randomization	Teacher Rotation
Johnson, Johnson, Dudley	88 - 89	Suburban	Cornelia Elementary, Edina, MN	1 - 6	15 Hours 30 Days	138 (WC)	Yes	Control Classes Randomly Selected	Yes
Johnson, Johnson, Dudley, Acikgoz	90 - 91	Suburban	Highlands Elementary, Edina, MN	1 - 6	15 Hours 30 Days	92 (WC)	No	Classes Randomly Selected	Yes
Johnson, Johnson, Dudley, Magnuson	91 - 92	Suburban	Creek Valley Elementary, Edina, MN	2 - 5	9 Hours 15 Days	227 (WC)	Yes	Classes Randomly Selected	No
Johnson, Johnson, Dudley, Ward, Magnuson	91 - 92	Suburban	Creek Valley Elementary, Edina, MN	2 - 5	9 Hours 15 Days	227 (WC)	Yes	Classes Randomly Selected	No
Johnson, Johnson, Mitchell, Cotton, Harris, Louison	91 - 92	Urban	Miller Park Elementary, Omaha, NE	3 - 4	10 Hours 1.5 Days	47 (Cadre)	No	None	NA
Johnson, Johnson Cotton, Harris, Louison	92 - 93	Urban	Miller Park Elementary, Omaha, NE	3 - 4	10 Hours 1.5 Days	39 (Cadre)	No	None	NA
Johnson & Johnson	93-94	Urban	Miller Park Elementary Omaha, NE	3-4	10 Hours 1.5 Days	34 (Cadre)	No	None	NA

Table 1 (continued)

Summary of Conflict Resolution and Peer Mediation Studies

Study	Year	Setting	School	Grades	Training Length	Sample Size	Control	Randomization	Teacher Rotation
Johnson, Johnson Dudley, Mitchell, Fredrickson	93 - 94	Suburban	Valley View Middle School Edina, MN	6 - 9	14 Hours 12 Weeks	235 (WC)	Yes	Control Classes Randomly Selected	Yes
Dudley, Johnson, Johnson	93 - 94	Suburban	Valley View Middle School Edina, MN	6 - 9	14 Hours 12 Weeks	235 (WC)	Yes	Control Classes Randomly Selected	Yes
Stevahn, Johnson, Johnson, Green, Laginski	93 - 94	Rural, Suburban, Canada	Durham High School, Ontario	9	9.5 Hours 2 Weeks	40 (WC)	Yes	Students Assigned to Conditions	Yes
Stevahn, Johnson, Johnson, Real	94 - 95	Rural, Suburban, Canada	Durham Region K - 8, Ontario	7 - 8	23 Hours 4 Weeks	111 (WC)	Yes	Students Assigned to Conditions	Yes
Stevahn, Johnson, Johnson, Laginski, O'Coin	94 - 95	Rural, Suburban Canada	Durham High School Ontario	9	10 Hours 4 Weeks	42 (WC)	Yes	Students Assigned to Conditions	Yes
Stevahn, Johnson, Johnson, Schultz	95-96	Suburban Diablo HS, Concord, CA	Mount	9	17.5 Hours 3 Weeks	92	Yes	Classes Assigned to Conditions	No (Same)

Table 1 (continued)

Summary of Conflict Resolution and Peer Mediation Studies

Study	Year	Setting	School	Grades	Training Length	Sample Size	Control	Randomization	Teacher Rotation
Stevahn, Johnson, Johnson, Oberle, Wahl	96-97	Suburban	Highlands Elementary, Edina, MN	Kinder-garten	9 Hours 4 Weeks	80	Yes	Students Assigned to Conditions	Yes
Stevahn, Munger, Kealey	97-98	Suburban	Dorval, Quebec	K-6	6 Months	302	No	None	No
Stevahn, Oberle Johnson, Johnson	99-00	Suburban	Highlands Elementary, Edina, MN	Kinder-garten			Yes	Students Assigned to Conditions	Yes

paper and pencil tests and interviews in which participants responded to questions about the material they had studied. Retention measures were administered anywhere from three weeks to one year after the study had ended. Another dependent variable was *knowledge of the negotiation and mediation procedures* taught. The *How I Manage Conflicts Measure* was aimed at assessing students' total recall of the steps in the negotiation procedure. Students were asked to write out step by step how they would resolve a conflict. Responses were scored for the presence of the six steps of the negotiation procedure. Retention was measured anywhere from three weeks to one year after the training had ended.

Another dependent variable was the *application of the negotiation procedure to actual conflicts*. The use of the negotiation and mediation procedures in actual conflicts were measured. Students received one point for each step of the procedures they used correctly.

The *strategy* used to manage conflicts was measured by two scales. The *Conflict Scenario Written Measure* was given to all participating students usually the day before the training began and the day after training ended. A brief scenario was given that ended in an unresolved conflict. Each student was asked to indicate what he or she would do if actually in the situation. Three different forms of this measure were used—a paper and pencil measure in which students read the scenario and responded in essay form, an interview in which students were read the scenario and described what he or she would do, and a video-taped enactment in which students were asked to role play with another student what they would do to resolve the conflict. Responses were categorized in two ways—a category system derived from a series of content analyses of students' responses and the categories specified by Conflict Strategies Constructiveness Theory (Johnson & F. Johnson, 2003). In the content analysis, the strategies were placed on a continuum from 0 to 12 from most destructive (physical and verbal aggression and avoidance) to most constructive (invoking norms for appropriate behavior, proposing alternatives, and negotiating). The continuum was built by consensus among two professors and two graduate students in social psychology. Responses were classified according to the strategy used to deal with the conflict. The strategies were arranged in a hierarchy ranging from most destructive (physical aggression) to most constructive (full negotiations) and assigned points. The strategies were (with number of points given in parentheses): physical aggression (1), verbal threats (2), unsatisfactory withdrawal (3), telling the teacher (4), commanding or requesting the other to give in (5), satisfactory withdrawal (6), invoking norms the other should conform to (7), proposing alternatives for the other to do (8), expressing intent to negotiate ("I would negotiate") (9), negotiating for mutual agreement with some steps present ("I would negotiate an agreement we both would like") (10), negotiating for mutual agreement with most steps present ("I would try to understand his/her point of view and negotiate an agreement we both would like) (11), and full integrative negotiation (all steps in the procedure present) (12). The data were also classified according to the Two-Concerns Conflict Strategies Theory (Johnson, 1972/2000; Johnson & F. Johnson, 1975/2000). This theory assumes participants in a conflict have two concerns: achieving their goal and maintaining a good relationship with the other person. When those two dimensions are combined, the five strategies of withdrawing, forcing, smoothing, compromising, and negotiating result. The strategies were placed on a continuum from 0 to 5.

Another dependent variable was the degree to which students engaged *in distributive versus integrative negotiations*. It was measured by the Integrative Bargaining Measure given to all participating students before the study began and after the study ended. The measure was adapted from ones developed by Johnson (1967) and Pruitt (1981) and comprised two bargaining exercises. In the first, students were randomly assigned to pairs consisting of a buyer and a seller negotiating over the exchange of three commodities (e.g., pencils, pens, and markers). In the second, students were randomly assigned to pairs consisting of co-workers with equal status in a manufacturing company, (each representing a different department) negotiating over additional vacation hours for the completion of three company projects. In both exercises, students within each pair could negotiate in ways that would maximize their own outcomes (distributive negotiation) or maximize joint outcomes (integrative negotiation). The degree to which the agreement maximized joint benefit (determined by adding together the outcome of both negotiators) was used as a measure of integrative negotiations.

Another dependent variable was attitudes toward conflict. It was measured in two ways. The Conflict Word Association Measure is a paper and pencil measure that asks students to write words that came to mind when thinking of conflict. Words were counted and categorized as (a) positive (feelings, actions, or outcomes associated with the constructive resolution of conflict), (b) negative (feelings, actions, or outcomes associated with the destructive resolution of conflict), or (c) neutral (neither positive nor negative nor a definition). The second measure of attitudes consisted of two forms included in required student notebooks. Students were required to keep an academic notebook in which they completed assignment sheets. Two forms, the Teaching Reflection Form and the Weekly Reflection Form, were unobtrusive measures of positive interest in the conflict resolution training.

Another dependent measure was the *resolutions of the conflicts*. A content analysis on student responses resulted in the following categories: no agreement, third-party imposed agreement, avoidance of each other, other wins, you win, forgiving and apologizing, new solution agreeable to both developed (such as a compromise or a decision determined by chance), and fully negotiated mutual. The solutions were placed on a continuum from 0 to 8 from most destructive (no solution, authority imposed solution, winner take all) to most constructive (proposing new solution,

reaching an integrative agreement).

Another dependent measure was a *mediator report form* on which peer mediators recorded the following information: Date, the names of the mediators, the names of the students involved in the conflict, who referred the students for mediation, the type of conflict, the strategies used during the conflict, whether the conflict was resolved or not, and if so the agreed on solution. The information contained in the conflict reporting forms can be analyzed into type of conflict, strategies students used during the conflict, and the agreed upon solutions. Two classification systems were used to determine the *types of conflicts* occurring in the school. Deutsch (1973) defined five types of conflicts: control of resources (such as books, computers, athletic equipment, television sets), preferences (what game to play, what activity to do first), values (what "should be"), beliefs (what "is"), or the nature of the relationship between the individuals involved (who is dominant, what kind of friendship to have). In our research studies we have conducted a series of content analyses that resulted in categorizing conflicts into seven types: aggression/fights, insults/put-downs/rumors, playground conflicts, turn-taking conflicts, possession/access conflicts, preferences/values/beliefs (what I prefer, what I believe is fact, what I think "should be"), and conflicts over academic work.

All student responses were coded independently by two different coders who were advanced doctoral students in social psychology. A 97 percent agreement level was found using the ratio of agreements to coded occurrences.

Effect Size

Whenever possible, the findings of the peacemaker studies were converted to effect sizes (Cohen, 1987; Cooper, 1989; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Hunter, Schmidt, & Jackson, 1982). The *effect size d* was the difference between treatment divided by the pooled standard deviation of the two groups (Cohen, 1989).

Results Of Our Mediation Studies

Before Training

Before training, most students were found to be involved in conflicts daily. These conflicts were rarely managed constructively. Untrained students indicated they would resolve the conflicts by either forcing the other person to submit (repeating their request, using threats, aggression, commands for the other to give in, and other competitive strategies to "win") or withdrawing unsatisfied from the situation and the relationship. In the urban elementary school, physical force and verbal intimidation were the two most frequent strategies reported. Very few students indicated they would give up their goals to maintain a high-quality relationship with the other person. No untrained student used problem-solving negotiations as a means to resolve the conflicts. Students often would refer the conflict to the teacher. One of the teachers, in her log, stated, "Before training, students viewed conflict as fights that always resulted in a winner and a loser. To avoid such an unpleasant situation, they usually placed the responsibility for resolving conflicts on me, the teacher." These findings indicated that students were generally not being taught negotiation procedures and skills in the home or community at large and, therefore, that they needed to be trained in how to manage conflicts constructively.

Types And Frequencies Of Conflicts

Data were collected on the types of conflicts student were involved in and how frequently each type occurred. In the urban elementary school the vast majority of conflicts referred to mediation involved physical and verbal violence. In the suburban schools, the majority of conflicts were over the possession and access to resources, preferences about what to do, playground issues, and turn-taking. A small percentage of the conflicts involved physical and verbal aggression. More different types of conflicts were reported at school than at home. The conflicts at home tended to be over preferences, possessions, and access; few conflicts were reported over beliefs and relationships or involved physical fights and verbal insults. Very few conflicts occurred over academic work in either setting. Value conflicts were almost never reported.

Learning of Negotiation and Mediation Procedures

Following training, students were given a test requiring them to write from memory the steps of negotiation and the procedures for mediation. Across our studies, over 90 percent of the students accurately recalled 100 percent of the negotiation steps and the mediation procedure. Up to a year after the training had ended, on average over 75 percent of students were still able to write out all the negotiation and mediation steps. These results indicate that the training was quite effective in teaching students the negotiation and mediation procedures. The average effect size for the studies was 2.25 for the immediate post-test and 3.34 for the retention measures (see Table 2).

Application of Negotiation and Mediation Procedures to Conflicts

Learning the negotiation and mediation procedures does not necessarily mean that students will use them in actual conflict situations. We used both paper-and-pencil and observation measures to determine whether students could use the procedures in actual conflicts. Students completed Conflict Report Forms on any conflict they were involved in or were observed in actual conflicts with their classmates. The results (see Table 2) indicate that students were quite good at applying the negotiation and mediation procedures. Immediately after training, students applied the procedures almost perfectly (effect size = 2.16) and were still quite good

Table 2

Mean Weighted Effect Sizes for Peacemaker Studies

Dependent Variables	*M*	*SD*	Number of Effects
Learned Procedure	2.25	1.98	13
Learned Procedure - Retention	3.34	4.16	9
Applied Procedure	2.16	1.31	4
Application - Retention	0.46	0.16	3
Strategy Constructiveness	1.60	1.70	21
Strategy Constructiveness Retention	1.10	0.53	10
Two-Concerns-Retention	0.45	0.46	5
Integrative Agreement Measure	0.98	0.36	5
Quality of Solutions	0.73	0	1
Positive Attitude	1.07	0.25	5
Negative Attitude	-0.61	0.37	2
Subject Matter Learning	0.88	0.09	5
Subject Matter Retention	0.70	0.31	4

months after the training was over (effect size = 0.46).

Strategies Used To Manage Conflicts

Five measures were used to measure the strategies students used in managing their conflicts:

1. Conflict Report Forms on which students periodically recorded the conflicts they were involved in.

2. Written responses to conflict scenarios: Students were given descriptions of conflict situations (such as a conflict over access to a computer or a personal insult through name calling) and asked to write out how they would resolve the conflict.

3. Oral responses to conflict scenarios given in an interview: Students were individually interviewed. A description of a conflict was read to them, and they were asked to explain how they would resolve it.

4. Role playing responses to conflict scenarios that were videotaped: Students were randomly assigned to pairs, each assigned a role in a common conflict, and asked to role play how they would resolve it. They were videotaped doing so.

5. Actual conflicts created with a classmate: Students (a) ranked several alternative ways of completing an assignment, (b) were paired with another student who ranked the alternatives differently, (c) resolved the conflict by deciding which alternative to adopt, and (d) wrote out the actions they took (step by step) to resolve the conflict. each pair member described what was done to reach the agreement on a form.

The results from the different measures were consistent. The diversity of these measures adds validity and generalizability to the results. Responses were categorized in two ways. A content analysis was conducted that resulted in a 12-point continuum from destructive (physical and verbal aggression and avoidance) to constructive (invoking norms for appropriate behavior, proposing alternatives, and using the problem-solving negotiation procedure) actions. The data were also classified according to the Two-Concerns Conflict Strategies Theory (Johnson & F. Johnson, 2000) which proposes that **disputants have two concerns: achieving their goals and maintaining a good relationship with the other person.** When these two dimensions are combined, the five strategies of withdrawing, forcing smoothing, compromising, and negotiating to solve the problem result. Overall, from Table 2 it may be seen that for the Strategy Constructiveness Scale, the average effect size was 1.60 on the post test and 1.10 for the retention tests. For the Two-Concerns Scale, the post-test effect-size was 1.10

and the retention effect size was 0.45. Trained students tended to use the integrative negotiation and mediation procedures in resolving the conflicts. There were no significant differences between males and females in the strategies used to manage conflicts. Although the training took place in school, and focused on school conflicts, there were no significant differences between the strategies used in school and in the home. Students used the strategies learned in school just as frequently in the home as they did in the school.

Resolutions

On the Conflict Report Form students were asked to report the nature of the resolution of the conflict. Very few of the conflicts were reported to be arbitrated by adults or resolved through forgiving in either the control or experimental groups. The number of integrative solutions which resulted in both sides achieving their goals was much higher in conflicts among trained (rather than untrained) students. Untrained students left many conflicts unresolved. There was no significant difference between the solutions arrived at for conflicts in school or at home. Only one study had the necessary analyzes to determine an effect size (ES = 0.73). This indicates that trained students tend to find more constructive resolutions than do untrained students.

Transfer Of Training: Spontaneous Use Of Negotiation And Mediation Procedures In Nonclassroom And Nonschool Settings

An important issue in conflict training is whether the procedures and skills learned will transfer to nonclassroom and nonschool settings. Three types of measures were used to determine the extent of transfer:

1. The spontaneous use of the negotiation and mediation procedures in settings other than the classroom.

2. Written descriptions of conflicts students were involved in outside of the classroom.

3. Systematic observation of students in nonclassroom settings.

The integrative negotiation and mediation procedures students learned in the classroom did generalize to nonclassroom and nonschool situations. The use of the procedures was reported on the playground, in the lunchroom, in the hallways, on school buses, and in the home. Students spontaneously wrote stories about using the negotiation and mediation procedures, students spontaneously presented skits in the school variety show involving the negotiation and mediation procedures, and parents reported that students used the negotiation and mediation procedures and skills with their brothers and sisters, their neighborhood friends, and even their pets.

In a number of studies, students regularly filled out Conflict Report Forms detailing the conflicts they were involved in and how they were resolved. Over half of these conflicts occurred in the home. An example from a fourth-grade student is, "My conflict was when me and my brother were fighting about who would get the dry baseball and who would get the wet one. So me and my brother did conflict resolution and it ended up that I first got the wet ball and my brother got the dry ball, then I got the dry ball and my brother got the wet ball."

Finally, students were directly observed on the playground and other nonclassroom settings in the school (Johnson, Johnson, Dudley, Acikgoz, 1994; Stevahn, Johnson, Johnson, Oberle, & Wahl, 2000). The conflicts were classified as either low-investment or high-investment. **Low-investment conflicts** were usually light-hearted and lasted thirty seconds or one minute. An example of a low-investment conflict is a girl who wanted to give a picture to somebody and asked, "Who wants this?" More than one student wanted it and this started a conflict. When one of the students got the picture, the others stopped and began to laugh, one of them said, "Do you see what you gave up?" The entire conflict lasted about one minute and did not consume any of the classes emotional or academic energy. Formal negotiation and mediation procedures were not used, even by trained students. **High-investment conflicts** affected the students emotionally by detracting from the students' ability to work academically, or interact with classmates in a positive manner. Often times these conflicts would last for days or longer. Once trained, students involved in high-investment conflicts did enter into integrative negotiations and seek out mediation. A group of sixth grade girls, for example, engaged in a prolonged conflict over who were "best friends" and who were no longer "best friends." In a fifth-grade class Bill and his friends believed Doug and his friends were passing notes saying nasty things about Bill. Doug denied it. But Bill responded, "since you are doing it to us we are going to do it to you." So Bill and his friends started passing nasty notes about Doug. The resulting conflict took several days to resolve and required mediation. Another example occurred when two girls became very angry at each other because both wanted the same part in a play. Mediation was finally required. Trained students were observed to be very high on effort to resolve the conflict, attentive listening to each other, emotional seriousness about the conflict and its resolution, commitment to the negotiation and mediation procedures, and expression of personal respect for the other individuals involved.

Distributive ("Win-Lose") Versus Integrative (Problem-Solving) Negotiations

A number of studies examined the impact of the Peacemaker training on the students' approach to negotiating. Students were placed in a negotiation simulation involving the buying and selling of commodities in which

they could adopt a "win-lose" (maximize own outcomes) or an integrative (maximize joint outcomes) negotiation approach. Students who had received the Peacemaker training used the integrative approach significantly more frequently than did the untrained students. No untrained student adopted an integrative approach to negotiations while almost all students in the experimental condition did so. From Table 2 it may be seen that the average effect size was 0.98. Trained students were more likely to engage in integrative negotiations than were untrained students.

Academic Achievement

The Peacemaker Training has been integrated into both English literature and history academic units to determine its impact on academic achievement. The basic design for these studies was to randomly assign students to classes in which the Peacemaker training was integrated into the academic unit studied or to classes in which the academic unit was studied without any conflict training. Students in the experimental classes both studied the academic material and learned the negotiation and mediation procedures. Students in the control classes spent all their time studying the novel. Students who received the Peacemaker training as part of the academic unit tended to score significantly higher on achievement and retention tests than did students who studied the academic unit only. Students not only learned the factual information contained in the academic unit better, they were better able to interpret the information in insightful ways. The higher achievement is all the more notable as students in the control classes spent all their time studying the academic material, while students in the experimental classes had to learn both the novel and the negotiation and mediation procedures in the same amount of time. From Table 2 it may be seen that the average post-test effect size was 0.88 and the average retention effect size was 0.70.

Linking conflict resolution training with academic learning is important, as the history of innovations in schools indicates that new programs are not widely adopted and maintained over a number of years unless they increase students' academic achievement. The integration of conflict resolution training into academic units, furthermore, provides an arena in which frequent and continued practice of the conflict resolution procedures can take place.

Attitudes Toward Conflict

Attitudes toward conflict was measured by a word association task. One of the goals of conflict resolution and peer mediation training is to create positive attitudes toward conflict. Before training, overwhelmingly the students held negative attitudes toward conflict seeing almost no potential positive outcomes. While still perceiving conflict more negatively than positively, the attitudes of trained students became markedly more positive and less negative while the attitudes of untrained students stayed essentially the same (highly negative). The average effect size was positive attitudes toward conflict was 1.07 and for negative attitudes was -0.61.

Interviews Of Teachers, Principals, And Parents

A number of participating teachers and principals were interviewed. All endorsed the program. The teachers reported that the training resulted in conflicts among students becoming less severe and destructive and, therefore, the classroom climate became more positive and the teachers (and principals) spent much less time resolving conflicts among students. Conflicts referred to them were reduced by 80% and the number of conflicts referred to the principal was reduced to zero. The teachers and principals interviewed said that without qualification they would become involved in the training in the future. One teacher commented, "The negotiation and mediation skills we are teaching our students will have a definite positive impact on the way our students interact with each other...these skills go beyond the scope of the classroom, and contribute to the betterment of our community, and our world." Parents reported that students mediated conflicts at home and in other contexts from the school. A third grade girl, for example, attempted to mediate a conflict between her older sister and the older sister's boyfriend while the three of them were at a drive-in movie theater. In one family the parents were arguing over whether to refinish the dining room table after her younger brother accidentally spilled nail polish on it. In the midst of their heated argument the fourth grader interrupted and mediated the conflict. She took her parents through all the steps of resolving conflicts. Many parents whose children were not part of the project requested that their children receive the training next year. A number of parents of trained children requested that they themselves receive the training so that they could use the negotiation and mediation procedures to manage conflict management in the family.

Discussion

Concern about the way in which children, adolescents, and young adults manage their conflicts has led to numerous conflict resolution and peer mediation training programs. Few of these programs have been evaluated (Johnson & Johnson, 1996). Only one, **Teaching Students To Be Peacemakers** (Johnson & Johnson, 1995a), has generated a series of research studies. This article contains a meta-analysis summarizing the results of these sixteen studies.

The results indicate that students are involved in conflicts with each other daily and either refer the conflict to the teacher for arbitration or use destructive strategies that tended to escalate the conflict. The vast majority of students did not know how to negotiate resolutions to conflicts. In suburban schools, the conflicts dealt with control of resources, preferences, and verbal and physical aggression and playground issues, while in inner-city schools studied

the conflicts dealt with violence and aggression. In light of these findings, educators are justified in being concerned about the occurrence of conflicts in schools and the way students resolve their conflicts.

A major barrier to ensuring that conflicts are managed constructively is students' lack of knowledge of and skill in using effective conflict resolution procedures. Students do have procedures for managing conflicts, but often the procedures are not constructive and different students use different strategies. Especially when students are from different cultural, ethnic, social class, and language backgrounds the different strategies used to manage conflicts may create some chaos. Life in schools gets easier when all students (and staff members) use the same set of negotiation and mediation procedures in managing conflicts.

Given the extent of the problems, an important issue is whether conflict resolution training will change the way students behave in conflict situations. The results of our research indicate that students do in fact learn how to negotiate and mediate to a total recall level. The conflict resolution and peer mediation training, however, can only be considered successful if students retain what they have learned over a long period of time. Throughout the academic year, months after students received their training, almost all trained students maintained their knowledge about negotiation and mediation. One possible explanation for the high rate of retention is that students were using the procedures to manage their day-to-day conflicts with classmates and peers.

Knowing what the procedures are, however, does not mean that students will be able to apply them in real-life conflicts. The results of the scenario measures (written, interview, and videotaped simulation) indicate that trained students were able to apply the negotiation and mediation procedures to actual conflicts. Observation results and interviews with teachers, principals, and parents indicate that students did use the negotiation and mediation procedures in playground, neighborhood, and family settings. Even though the types of conflicts students face differ in school and home situations, the results of our studies indicate that the negotiation and mediation procedures were used in both settings.

Not only can students be taught how to negotiate and mediate, the training can be integrated into instructional units in ways that promote academic achievement. When the Peacemaker training was integrated into an English literature class, academic achievement increased even though time was taken away from studying the novel to learn the negotiation procedure. Since much of literature, history, and science deals with conflicts, the possibilities of integrating conflict resolution training and subject matter learning seem quite possible and promising.

Just because students knows the procedures for integrative negotiations does not mean they will also do so. They can still go for a "win" when faced with a conflict with peers. When placed in a conflict situation in which either a distributive ("win-lose") or an integrative ("problem-solving") negotiation strategy could be used, all untrained students used the "win-lose" strategy, students who had the Peacemaker training primarily used the integrative strategy. This finding not only confirms the effectiveness of the training, it also provides a bridge to the overall research efforts on conflict in the social sciences.

In our studies, the adults in the school perceived the peer mediation program to be constructive and helpful. Teachers, principals, and parents stated the conflict resolution program reduced the incidence of destructively managed conflicts and resulted in a more positive classroom climate. The implication is that when students regulate their own behavior, the need for teachers and administrators to monitor and control student actions declines. The fewer the discipline problems teachers have to manage, the more teachers can devote their energies to teaching.

Finally, the success in training children as young as kindergartners in the Peacemaker procedures fills an important gap in the research on conflict resolution training in the schools.

Learning how to manage conflicts constructively may have consequences for cognitive, social, moral, and psychological development (Johnson & Johnson, 1989, 1995a, 1995b). Teaching students how to (a) negotiate integrative agreements to resolve conflicts and (b) mediate schoolmates' conflicts may, for example, increase students' ability to regulate their own behavior and give them a developmental advantage over peers who have not been so trained. Self-regulation is the ability to act in socially approved ways in the absence of external monitoring by others. It is a central and significant hallmark of cognitive, social, and moral development. The opposite of self-regulation is adult monitoring and control. When adults in children's lives act as referees and judges, adults place children in a dependent position and deprive them of opportunities to learn valuable self-regulation and social skills. The more students master and use the integrative negotiation and mediation procedures, the more they independently regulate their own behavior and the less monitoring and control required by adults. In addition, children and adolescents who are able to negotiate and mediate may have a developmental advantage over children who do not know how to do so. When faced with conflicts in school, work, home, or leisure settings, they may be able to be more successful in negotiating agreements that maximize mutual benefit and increase the quality of the relationship.

While any conclusions drawn from our studies have to be tentative, the implications are interesting. Children and adolescents have frequent conflicts daily in school. Without training, they tend to manage their conflicts in destructive ways. When given training, they learn how to engage in integrative negotiations and how to mediate their schoolmates' conflicts and they maintain their ability to do so months after the training has ended. They apply the learned procedures to actual conflicts in classroom, school, and family settings. Learning the negotiation and mediation procedures can be integrated with academic learning in a

way that enhances subject matter understanding. Trained students tend to stick with an integrative approach to negotiations, even when they are tempted to go for a "win." The adults in the school, furthermore, develop positive attitudes toward and perceptions of the conflict resolution training and its impact on students. These implications set the stage for teaching all students in every school how to manage conflicts in constructive ways. The systematic results from our studies add important empirical confirmation to the anecdotal testimonies to the effectiveness of peer mediation in schools and provide validation of the assumptions underlying the total student body approach to peer mediation.

The conclusions and implications of our studies are both strengthened and limited by the nature of the research methodology. The strengths of the studies are in the diversity of the schools and students studied (suburban and urban), the variety of age groups included in the studies (first through ninth), the fact that two different approaches to peer mediation were studied (total-student-body and cadre), the frequency with which control groups were included, the randomly sampling of students and classrooms, the rotation of teachers across conditions, the inclusion of schools from two different countries (United States and Canada), the long duration of the studies (up to eight months), the careful development and validation of the dependent variables, the grounding of the Peacemaker Program in conflict resolution theory (constructive conflict [Deutsch, 1973], perspective taking [Johnson, 1971], communication in conflict situations [Johnson, 1974], and integrative negotiations [Johnson, 1967; Pruitt, 1981]), and the careful controlled way in which the studies were carried out. In addition, these studies were carried out in the "real life" setting of schools. While there is a great deal of research in the field of conflict resolution, almost all of it consists of small, tightly-controlled, experimental studies of short duration conducted in the laboratory. Relatively few studies that examine the effectiveness of the recommended procedures in "real-life" settings such as schools over a relatively long period of time.

Despite these strengths, however, the results of our studies are limited by the specific operationalizations of the independent and dependent variables used, the ability of the teachers to implement the training programs, the ages of the students involved, the ethnicity of the students (caucasion and black only), the geographic locations of the school districts (midwest and North American), and many other factors. Considerable more research is needed on teaching students how to negotiate and mediate effectively and on the operation of peer mediation programs in the schools. Future investigations need to continue to identify and examine the variables that mediate the effectiveness of conflict resolution training. The conditions need to be established under which academic learning and conflict resolution training may be integrated to enhance each other's effectiveness. Another productive avenue of work would be to compare the effectiveness of cadre and total-student-body approaches to peer mediation to delineate (a) their comparative strengths and weaknesses and (b) the conditions under which each will have optimal effects. Finally, it would be useful to examine the interaction between trained and untrained students when they are in conflict with each other.

Every student needs to learn how to manage conflicts constructively. Without training, many students may never learn how to do so. Teaching every student how to negotiate and mediate will ensure that future generations are prepared to manage conflicts constructively in career, family, community, national, and international settings. There is no reason to expect, however, that the process will be easy or quick. It took over 30 years to reduce smoking in America. It took over 20 years to reduce drunk driving. It may take even longer to ensure that children and adolescents can manage conflicts constructively. The more years students spend learning and practicing the negotiation and mediation procedures, the more likely they will be to actually use the procedures skillfully both in the classroom and beyond the school door.

References

Benbenishty, R., Astor, R., & Zeira, A. (1998, November). *A national study of school violence in Israel: Fall 1998.* Jerusalem, Israel: Israeli Ministry of Education.

Benbenishty, R., Zeira, A , & Astor, R. (1999, November). *A national study of school violence in Israel—Wave II: Fall 1999.* Jerusalem, Israel: Israeli Ministry of Education.

Benbenishty, R., Zeira, A , & Astor, R. (2000, January). *Principals and teachers reports of school violence in Israel—Wave II: Fall 1999.* Jerusalem, Israel: Israeli Ministry of Education.

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations, 2,* 129-152.

Deutsch, M. (1962). Cooperation and trust: Some theoretical notes. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 275-319). Lincoln, NE: University of Nebraska Press.

Deutsch, M. (1973). *The resolution of conflict.* New Haven, CT: Yale University Press.

Elam, S., & Rose, L. (1995). The 27th annual Phi Delta Kappa/ Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 77,* 41-56.

Elam, S., Rose, L., & Gallup, A. (1994). The 26th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 76,* 41-56.

Elam, S., Rose, L., & Gallup, A. (1996). The 26th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 78,* 41-56.

Everett, S., & Price, J. (1995). Students' perceptions of violence in the public schools: The MetLife Survey. *Journal of Adolescent Health, 17*(6), 345-352.

Furlong, M., & Morrison, G. (1994). Introduction to mini-

series: School violence and safety in perspective. *School Psychology Review, 23*, 139-150.

Johnson, D. W. (1967). Use of role reversal in intergroup competition. *Journal of Personality and Social Psychology, 7*, 135-141.

Johnson, D. W. (1970). *Social psychology of education.* Edina, MN: Interaction Book.

Johnson, D. W. (1971a). Role reversal: A summary and review of the research. *International Journal of Group Tensions, 1*, 318-334.

Johnson, D. W. (1971b). Students against the school establishment: Crisis intervention in school conflicts and organizational change. *Journal of School Psychology, 9*, 84-92.

Johnson, D. W. (1974). Communication and the inducement of cooperative behavior in conflicts: A critical review. *Speech Monographs, 41*, 64-78.

Johnson, D. W. (1983). *Resolving marital conflicts constructively.* Edina, MN: Interaction Book Company.

Johnson, D. W. (1991). *Human relations and your career* (2nd Ed.). Englewood Cliffs, NJ: Prentice Hall. First Edition, 1978.

Johnson, D. W. (2003). *Reaching out: Interpersonal effectiveness and self-actualization* (8th ed.). Boston: Allyn & Bacon. First Edition, 1972.

Johnson D. W., & Johnson, F. (2003). *Joining together: Group theory and group skills* (7th ed.). Boston: Allyn & Bacon. First Edition, 1975.

Johnson, D. W., & Johnson, R. (1979). Conflict in the classroom: Controversy and learning. Review of *Educational Research, 49*, 51-61.

Johnson, D. W., & Johnson, R. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R.(1994). *Leading the cooperative school* (2nd Ed.). Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R. (1995a). *Teaching students to be peacemakers.* Edina: Interaction Book Company.

Johnson, D. W., & Johnson, R. (1995b). *Creative controversy: Intellectual challenge in the classroom.* Edina: Interaction Book Company.

Johnson, D. W., & Johnson, R. (1995c). Why violence prevention programs don't work—and what does. *Educational Leadership, 52*(5), 63-68.

Johnson, D. W., & Johnson, R. (1996). Conflict resolution and peer mediation programs in schools: A review of the research. *Review Of Educational Research, 66*(4), 459-506.

Johnson, D. W., & Johnson, R. (1999). Implementing cooperative learning: Training sessions, transfer to the classroom, and maintaining long-term use. In N. Davidson, C. Brody, & C. Cooper (Eds) (1995). *Staff development for cooperative learning: Issues and approaches.* New York: Teachers College Press.

Johnson, D. W., Johnson, R., Dudley, B., & Burnett, R. (1992). Teaching students to be peer mediators. *Educational Leadership*, September.

Johnson, D. W., Johnson, R., & Holubec, E. (1998). *Cooperation in the classroom* (6th ed.). Edina, MN: Interaction Book Company.

Johnson, D. W., Johnson, R., & Johnson, F. (1976). Promoting constructive conflict in the classroom. Notre Dame *Journal of Education, 7*, 163-168.

Kaufman, P., Chen, X., Choy, S., Chandler, K., Chapman, C., Rand, M., & Ringel, C. (1999). *Indicators of school crime and safety, 1999.* Washington, D. C.: U.S. Departments of Education and Justice. NCES 999-251/ NCJ-172215.

Lewin, K. (1951). *Field theory in social science.* New York: Harper.

Morrison, G., Furlong, M., & Morrison, R. (1997). The safe school: Moving beyond crime prevention to school empowerment. In A. Goldstein & J. Conoley (Eds.), *School violence intervention:* A practical handbook (pp. 236-264). New York: Guilford Press.

Posner, M. (1994). Research raises troubling questions about violence prevention programs. *The Harvard Education Letter, 10*(3), 1-4.

Pruitt, D. (1981). *Negotiation behavior.* New York: Academic Press.

Rose, L., & Gallup, A. (1998). The 30th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 80*, 41-56.

Rose, L., Gallup, A., & Elam, S. (1997). The 29th annual Phi Delta Kappa/Gallup of the public's attitudes toward the public schools. *Phi Delta Kappan, 79*, 41-56.

Singley, M., & Anderson, J. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.

Smith, P., Morita, YU., Junger-Tas, J., Olweus, D., Catalano, R., & Slee, P. (1999). *The nature of school bullying: A cross-national perspective.* New York: Routledge.

Stop the Violence (1994, January). *Scholastic Update,* 2-6.

Watson, G., & Johnson, D. W. (1972). *Social psychology: Issues and insights* (2nd ed.). Philadelphia: Lippincott.

Webster, D. (1993). The unconvincing case for school-based conflict resolution programs for adolescents. *Health Affairs, 12*(4), 126-140.

Roger T. Johnson is a Professor of Education at the University of Minnesota. He specializes in science education.

David W. Johnson is a Professor of Educational Psychology at the University of Minnesota. He specializes in cooperation and competition, conflict resolution, and organizational change.

They are co-directors of the Cooperative Learning Center.

Appendix A

Studies Included in Meta-Analysis

1. Dudley, B. S., Johnson, D. W., & Johnson, R. (1996). Conflict-resolution training and middle-school students' integrative negotiation behavior. *Journal of Applied Social Psychology, 26*, 2038-2052.
2. Johnson, D. W., & Johnson, R. (2001). Peer mediation in an inner city elementary school. *Urban Education, 36*(2), 165-178.
3. Johnson, D. W., Johnson, R., & Dudley, B. (1992). Effects of peer mediation training on elementary school students. *Mediation Quarterly, 10*, 89-99.
4. Johnson, D. W., Johnson, R., Dudley, B., & Acikgoz, K. (1994). Peer mediation: Effects of conflict resolution training on elementary school students. *Journal of Social Psychology, 134*, 803-817.
5. Johnson, D. W., Johnson, R., Dudley, B., Mitchell, J., & Fredrickson, J. (1997). The impact of conflict resolution training on middle school students. *Journal of Social Psychology, 137*(1), 11-22.
6. Johnson, D. W., Johnson, R., Cotten, B., Harris, D., & Louison, S. (1995). Using conflict managers to mediate conflicts in an elementary school. *Mediation Quarterly, 12*(4), 379-390.
7. Johnson, D. W., Johnson, R., Dudley, & Magnuson, D. (1995). Training of elementary school students to manage conflict. *Journal of Social Psychology, 135*(6), 673-686.
8. Johnson, D. W., Johnson, R., Dudley, B., Ward, M., & Magnuson, D. (1995). Impact of peer mediation training on the management of school and home conflicts. *American Educational Research Journal, 32*(4), 829-844.
9. Johnson, D. W., Johnson, R., Mitchell, J., Cotton, B., Harris, D., & Louison, S. (1996). Conflict managers in an elementary school. *Journal of Educational Research, 89*(5), 280-285.
10. Stevahn, L., Johnson, D. W., Johnson, R. T., Green, K., & Laginski, A. M. (1997). Effects on high school students of conflict resolution training integrated into English literature. *Journal of Social Psychology, 137*(3), 302-315.
11. Stevahn, L., Johnson, D. W., Johnson, R. T., Laginski, A. M., & O'Coin, I. (1996). Effects on high school students of integrating conflict resolution and peer mediation training into an academic unit. *Mediation Quarterly, 14*(1), 21-36.
12. Stevahn, L., Johnson, D. W., Johnson, R., Oberle, K., & Wahl, L. (2000). Effects of conflict resolution training integrated into a kindergarten curriculum. *Child Development, 71*(3), 772-784.
13. Stevahn, L., Johnson, D. W., Johnson, R. T., & Real, D. (1996). The impact of a cooperative or individualistic context on the effectiveness of conflict resolution training. *American Educational Research Journal, 33*, 801-823.
14. Stevahn, L., Johnson, D. W., Johnson, R., & Schultz, R. (2002). Effects of conflict resolution training integrated into a high school social studies curriculum. *Journal of Social Psychology, 142* (3), 305-331.
15. Stevahn, L., Munger, L., & Kealey, K. (submitted for publication). First-year effects of teaching all students conflict resolution in a French Immersion elementary school.
16. Stevahn, L., Oberle, K., Johnson, D. W., Johnson, R., (2001). *Effects of role reversal training and use of integrative negotiation on conflict resolution in kindergarten*. Paper presented at Annual Meeting of the American Educational Research Association, Seattle, April.

Ethnicity, 'Role Models' and Teaching

Bruce Carrington
University of Newcastle, UK

During the past five years, a range of measures has been introduced by the Teacher Training Agency (TTA) in England and Wales to broaden access to pre-service teacher education and, in particular, to increase the number of ethnic minority teachers. The paper begins by examining the assumptions underpinning this policy that stresses the importance of such teachers as 'role models' for minority children. Drawing upon the findings of a recently completed study of ethnic minority student teachers and newly qualified teachers, it moves on to discuss a range of issues relating to their professional socialization. Detailed consideration is given to participants' images of teaching, motives for entering the profession and experiences while training (especially during school placement). The paper concludes by exploring the policy implications of the research.

Introduction

For more than two decades, concern has been voiced in the UK about the under representation of ethnic minority groups in teaching and their apparent lack of opportunities for career advancement (Mahony & Hextall, 2000; Department for Education & Employment, 1998; Osler, 1997). Similar disquiet has also been expressed in the United States, where negative public images of teaching – as a poorly paid, stressful occupation with relatively low status - continue to serve as a barrier to minority recruitment. As Gordon (2000) has argued, whereas teaching was once seen as an avenue of upward social mobility by "working class students of color", attitudes towards the occupation have shifted perceptibly during the past four decades. She notes: "As education has moved from a privilege to a right, more and more students have come to see teachers in the role of laborers rather than professionals" (Gordon, 2000, p.5).

In 1997, the Teacher Training Agency (TTA) in England and Wales responded to demands for a more ethnically-diverse teaching force by introducing a range of measures to broaden access to pre-service teacher education: for example, ethnic monitoring and target setting (at both an institutional and national level), mentoring schemes and so-called "taster" (i.e. introductory) courses. Addressing a conference jointly convened by the TTA and the Commission for Racial Equality (CRE) in November 1997, the former head of the CRE, Sir Herman (now Lord) Ouseley, defended this policy by saying that "underachievement could only be tackled if there were more ethnic minority teachers as role models" (Ghouri, 1997). Moreover, as the following government press notice illustrates, similar rhetoric has subsequently been employed to justify interventions to boost minority teacher recruitment: "The TTA has set targets to increase the proportion of teaching recruits from minority ethnic groups from six to nine per cent by 2006 so that minority pupils can have more positive role models" (Department for Education & Employment, 2000).

The drive to increase the number of ethnic minority teachers, as the foregoing remarks indicate, may in part reflect official disquiet about the obstacles that continue to hinder the progress of many African Caribbean and Bangladeshi children in UK schools (Gillborn & Mirza, 2000). An underlying assumption of such discourse is that the targeted recruitment of minority teachers is needed to bring about a change in the educational *status quo*. On the one hand, it would make available advocates for marginalized sectors of the school population, raising overall levels of academic achievement by challenging low expectations and cultural stereotypes. On the other, the provision of additional positive 'role models' in the classroom would serve to swell the ranks of significant others in schools with whom alienated minority youth might identify.

In the United States, Harold Levy, the Chancellor of the New York City Board of Education, has employed almost identical rhetoric to legitimate the Board's present measures to address the dual problem of teacher shortages and failing schools (Goffe, 2001). Prompted by continuing concern about the academic performance of African-American students, the current drive to recruit 8,000 teachers from the Caribbean reflects Levy's belief that such teachers would "bring a no-nonsense approach – strong on discipline and the three Rs – and provide a model of black achievement". Leaving this rhetoric aside, there can be little doubt that such initiatives are needed to counter the insidious effects of cultural stereotyping on minority group students, particularly African-Americans (Weiner, 2000; Quiocho & Rios, 2000; Steele & Aronson, 1995). However, as Weiner has shown, in her timely appraisal of developments in

the field of urban education, the teaching profession remains a quintessentially white domain even though white student enrollments have steadily declined during the past two decades and "hypersegregation" has become the defining characteristic of many American cities (with African-American and Hispanic students more likely than other groups to attend ethnically homogeneous schools). She states: "Urban teachers in the 90s were less likely to be working with white students and more likely to be teaching immigrants, and when teaching African-Americans and Hispanic students might well be working in schools that were not racially diverse" (pp. 387-388).

Statement of the Problem

The limitations of the 'role model' argument

It should be noted that parallel policy initiatives to address the so-called 'feminization' of teaching in the UK and elsewhere have been predicated upon similar assumptions about the salience of male teachers - as 'role models' - in the schooling of boys (Lahelma, 2000; Pepperell & Smedley, 1998). However, despite widespread acceptance of the view that the 'gender gap' in academic achievement and male disaffection from school stems from the dearth of male teachers, a recent Finnish study by Lahelma (2000) has suggested that the potential influence of a teacher's gender on educational outcomes should not be exaggerated. Drawing upon interviews with ninety 13- and 14-year-olds, she found that although the students often commented on the lack of male teachers in their schools, the issue of gender did not figure prominently in their observations about the quality of teaching that they had experienced. For the large part, the teacher's gender was perceived to be irrelevant. Her findings indicate that students tend to value teachers who – regardless of gender - are able to maintain discipline in the classroom in a friendly, sensitive and impartial manner.

In the United States, Latham (1999) has also drawn attention to the problematic nature of the 'role model' argument. Citing the work of Ehrenberg, Goldhaber & Brewer, (1995), he points to evidence from the National Education Longitudinal Study of 1988 which shows that while matching teachers and students by gender and ethnicity has little effect on educational achievement, it can have an impact on students' attitudes to school. Their attitudes, he suggests, "would appear to be influenced by whether they had at least some teachers of similar race, gender, and ethnicity". Likewise, Quiocho & Rios (2000), in a synthesis of predominantly North American literature on minority group teachers and schooling, note that Ehrenberg *et al*'s work revealed that "the teacher can be a salient factor in the subjective judgments of minority group students (more highly rated by minority group teachers) even if there was no difference in academic achievement for the students" (p. 510). Their wide-ranging review of research also serves to highlight the potential of such teachers to challenge the *status quo* in schools, empathize with minority students and, as a consequence, make a "real difference" in their lives:

> Many minority group teachers, in comparison to their European-American counterparts, are more likely to bring a critical, social justice orientation and consciousness that stems from their real, lived experiences with inequality. In a related way, they are often more willing to work actively to dismantle personal and institutional bias that they find in schools as well as to move toward culturally responsive school-based reform more than are European-American teachers (Quiocho & Rios, 2000, p. 522).

While ethnicity or gender may have some bearing upon the selection of 'role models', the link between such choices and personal identity are at best tenuous. As Allen (2000) has argued, the concept of a 'role model' - with ethnic- or gender-based preferences - is an ambiguous one used variously to describe the following: "an ethical template for the exercise of adult responsibilities"; "a symbol of special achievement"; and "a nurturer providing special educational services" (i.e. mentor). According to Allen, while all teachers provide students with "an ethical template", only a few come to be recognized as "symbols of special achievement". Matching students and teachers by ethnicity, she argues, can give no guarantee that such recognition will occur: "Being a minority group member is neither necessary nor sufficient for being a role model to minority students".

Pole's (1999) qualitative study of minority teachers' lives and careers in the UK lends empirical support to Allen's claims. When invited to account for their decisions to take up a career in teaching, his respondents invariably referred to the enduring influence of teacher 'role models'. However, Pole is concerned to emphasize that these significant others were not always of the same ethnicity: "Role models were not race-specific in all cases. What seemed most important was exposure to a person of influence who showed an interest in the individual and made them realize that teaching was a possibility for them" (Pole, 1999, p. 318). Pole's findings also serve to illustrate the problematic nature of the 'role model' argument in other ways. For example, his interviews revealed "a reluctance on the part of some black pupils to accept black people in the role of the teacher' (1999, p. 322). One black teacher recounted how – "in the early days of her appointment" - she received "little respect" from black children who dismissed her as a "coconut": that is, black on the outside but white within. [The identification of the role of the teacher with "whiteness" is, no doubt, symptomatic of the widespread disaffection among this sector of the English school population (Majors, 2001; Sewell, 1997).] In a similar vein, Pole draws attention to the plight of a Muslim teacher whose authority had been "openly challenged" by a group of Sikh boys.

Ethnicity and teaching in England and Wales

Various small-scale studies have indicated that British African-Caribbeans and Asians tend to enter teaching for altruistic reasons: they wish to "make a difference" to the lives

of children from ethnic minority groups and often see themselves as potential 'role models' for such children (Pole, 1999; Garewal, 1999; Osler, 1997; Robbins, 1995). This particular social construction, however, is not accepted by some minority entrants to teaching who resent being cast as advocates for Black and Asian children, or seen as experts on their cultures (Wolverhampton Race Equality Council Consortium, 1999). There is also evidence which suggests that it is not uncommon for minority student teachers to encounter racism while training, particularly during school placement, where they are more likely to experience difficulties with discipline than their white counterparts (Osler, 1997; Siraj-Blatchford, 1991). They not only have to deal with racist comments from white children, but, as Pole's observations testify, may also find that children from the same or similar backgrounds disparage them. The possibility of undertaking a placement in a school where levels of racism are high is an additional source of stress for minority student teachers. Moreover, there are indications that such students may also experience a degree of social isolation during placement, especially when working in wholly or predominantly white schools. Many feel more comfortable working in multi-ethnic settings, where they can relate more readily to staff, parents and pupils (Osler, 1997; Jones, Maguire & Watson, 1997; Wolverhampton Race Equality Council Consortium, 1999). Finally, despite their concerns about confronting racism and social isolation in schools, ethnic minority students tend to have little or no say in decisions about the allocation of their placements (Osler, 1997; Robbins, 1995; Siraj-Blatchford, 1991).

Methodology

The study, upon which the remainder of this paper draws, set out to extend this body of research (Carrington, Bonnett, Demaine, Hall, Nayak, Skelton, Smith, Short & Tomlin, 2001). This independent inquiry, which was supported by a TTA grant, began in the summer of 1998 and had five main strands. The first comprised a national survey of ethnic minority students enrolled on one-year, full-time Postgraduate Certificate in Education (PGCE) courses. [The PGCE provides pre-service education and training for the bulk of graduates embarking upon a teaching career in England and Wales.] Initially, 776 students were invited to complete a written questionnaire at the beginning of their courses. Of these, 289 (73 males and 216 females) volunteered to take part in the survey, giving a response rate of 37 per cent. Because of the time scale of the project, no effort was made to follow up non-respondents.

The breakdown of this convenience sample, by official ethnicity and gender, was as follows: 'Black Caribbean' (5 males, 24 females); 'Black African' (12 males, 16 females); 'Black Other' (2 males, 11 females), 'Asian Indian' (19 males, 45 females); 'Asian Pakistani' (4 males, 20 females); 'Asian Bangladeshi' (5 males, 11 females); 'Asian Chinese' (3 males, 4 females); 'Asian Other' (6 males, 19 females); and 'Other' (17 males, 66 females). Around half (38) of those checking the latter category described themselves as having mixed parentage or heritage: for example, "Mixed; Scottish/Mauritian"; "Anglo-Indian, mixed race"; "Eurasian-Portuguese/English/Dutch/Chinese"; "Irish-Egyptian"; "Human being, White/Burmese". The remainder of the group came from a wide range of backgrounds. [NB From the outset of the study, the research team took cognizance of the limitations of ethnic monitoring and the adequacy of official categorizations of ethnicity (Bonnett & Carrington, 2000; Runnymede Trust, 2000). Such categories are utilized in this account for descriptive purposes only.]

The PGCE questionnaire, which comprised both closed and open-ended items, was designed to elicit data on the following: respondents' motives for entering teaching; the factors influencing their choice of course and institution; their images of the profession and perceptions of the barriers to ethnic minority recruitment; and their views on current policy initiatives in this sphere. Respondents were also invited to provide additional written comments on any of the issues raised. They were assured that their replies would be used anonymously. Finally, those willing to take part in follow-up interviews were asked to supply the necessary contact details.

The second strand in the research, undertaken during the summer and autumn of 1998, considered the responses of senior PGCE staff (course directors, admissions tutors and other key personnel) in sixteen institutions to the policy in question. In the third strand of the investigation, semi-structured interviews were conducted in eleven of these institutions with a cross-section of volunteers from the initial student survey (11 males, 38 females). These took place during the spring and summer of 1999. As well as giving participants an opportunity to elaborate upon their written responses to the questionnaire, the interviews were used to explore various issues relating to the quality of support provided during the PGCE and their experiences during school placement. Once again, the students were given formal assurances that the interviews would be conducted under conditions of strict confidentiality and that appropriate steps would be taken to maintain both individual and institutional anonymity.

The fourth and fifth strands in the investigation, undertaken between January and June 2000, set out to examine the experiences of newly qualified minority teachers, from the 1998/9 PGCE cohort, in their first posts. The fourth strand drew upon data from a written questionnaire completed by a convenience sample of 149 respondents (33 males and 116 females). The questionnaire consisted mainly of closed items dealing with these issues: respondents' reflections on the quality of the support provided by university and school staff during the PGCE; their experiences during school placement, while looking for a teaching job and, subsequently, in their first posts; and their perceptions of contemporary policy measures to address the under representation of ethnic minority groups in teaching. Finally, respondents were invited to provide details of any incidents of racial harassment that they had encountered during the PGCE or induction year. The fifth strand, a series of follow-up telephone interviews (9 males, 35 females), enabled the research team to glean extra information about the teachers' experiences, both prior to and during the induction year.

Findings and Discussion

In the discussion that follows, particular consideration is given to the findings that relate, directly or indirectly, to the 'role model' argument. I begin by examining the students' and newly qualified teachers' reasons for embarking upon a teaching career and their images of the profession, showing that only a small proportion was motivated by the belief that ethnic minority children would benefit from their presence as 'role models' in the classroom. Participants' perceptions of current policies to make teaching a more ethnically inclusive profession are then explored. Following a brief discussion of the constraints upon minority recruitment, more detailed consideration is given to their experiences in schools, both as PGCE students during placement and as newly qualified teachers in their first posts. Interview data are used to show that matching children and teachers by ethnicity can, in certain circumstances, add to the strains and uncertainties of classroom life; having the same or a similar background to the children is not a sufficient condition for winning their respect or gaining recognition as a 'role model'. An attempt is also made to show that although participants expressed anxieties about meeting social isolation and racial harassment in predominantly white schools, the difficulties actually encountered by those who worked in such settings were little different from those teaching in multi-ethnic schools. The paper concludes by exploring some of the policy implications of the findings.

Images of teaching and motives for entering the profession

A consistent feature throughout the research was the importance placed by participants on the intrinsic satisfactions of a teaching career. When the student teachers were invited at the beginning of the PGCE to describe their images of teaching, a sense of achievement or job satisfaction was overwhelmingly the most important positive factor in their responses. On the negative side, the initial questionnaire survey revealed that concerns about pay, conditions and status were paramount. This also formed a recurring theme throughout the research. Given their often contradictory images of the profession, the PGCE students' replies to the open-ended question, "What attracts you to teaching?", were of particular interest. Around two thirds (66%) indicated that they found the prospect of teaching intrinsically satisfying, while nearly a fifth (18%) simply stated that they wanted to work with children or young people. Of the remainder, a few noted that the length of the school holidays was a 'pull' factor (5%), while others said that teaching was a 'steady job' (4%). Only three per cent indicated that they saw themselves as potential 'role models' for ethnic minority children when explaining their motives for entering the profession.

The follow-up interviews, undertaken towards the end of the PGCE, also revealed that the prospect of becoming an ethnic minority 'role model' had not figured prominently in the participants' decisions to take up a teaching career. Although their motives for entering teaching were often complex and varied, an overall pattern was discernible and confirmed the findings of the initial questionnaire survey. Despite concerns about teachers' pay, conditions of work and levels of morale within the profession, the majority had a positive view of teaching and underlined the importance of intrinsic considerations when addressing the question, "What made you want to become a teacher?". The following remarks, made by Ayesha, a student of Pakistani origins, are illustrative: "I'm not very materialistic. I just want a job, career or profession that's really rewarding, but challenging at the same time. I like the idea of continually learning". [NB Pseudonyms are used throughout the paper.]

A few interviewees (6) referred directly to the perceived importance of minority 'role models' when accounting for their decisions to enter the profession. As the ensuing comments indicate, these students either saw themselves as advocates for British African Caribbean and Asian children, or justified their stance on the grounds that an enhanced presence of minority teachers in schools was necessary to raise expectations and counter stereotyping:

> I know what it's like for black kids in the education system. I wanted to be part of it to help them [...] and also to be a role model for black kids in schools (Madeleine, 'Other').

> You need to see more [minority teachers] who (sic) the children can turn to as role models. I think that's important. They need to see that their own people, from their own background are there (Edith, 'Black African').

> When I went through school, I didn't see any black role models [...]: my mum was a nurse, my dad was a bus driver and that was it! [...] When I was at school, I was told, "Yes you can be a hairdresser". I just think we're failing black children at the moment (Liz, 'Other').

Teacher recruitment and affirmative action

In general terms, the participants were supportive of current policy initiatives to make teaching a more ethnically inclusive occupation. For example, almost every respondent (94%) to the postal survey of newly qualified teachers agreed that it was important for all children to be taught by teachers from a range of ethnic backgrounds. Arguably, in adopting this position, they appeared to take issue with the somewhat circumscribed construction of ethnic minority teachers as 'role models' for ethnic minority children only. As one respondent of Indian heritage wrote on his questionnaire: "I see myself as a role model for any ethnic group". However, although the vast majority of those surveyed acknowledged that minority teachers have a crucial part to play in the education of all children (irrespective of ethnicity), many were themselves committed to teaching in multi-ethnic settings. Thus, when asked about the factors influencing their choice of teaching post, more than two thirds (69%) agreed or strongly agreed with the proposition, "I was keen to work in an ethnically mixed school". This

preference, which was also evident during the PGCE, may have reflected, in part, enduring concerns about the possibility of experiencing social isolation or racism in mainly white schools (see below).

Leaving aside the issue of matching teachers and children by ethnicity, reservations were encountered throughout the research about any measure that could be construed as conferring a comparative advantage on minority recruits. For example, Julie, a student of mixed heritage, voiced the following concerns during interview about the practicalities of affirmative action. When asked, "What advice would you offer the Government about its current policy to boost ethnic minority recruitment to teaching?", she replied: "The Government has to be careful not to show positive bias towards ethnic minorities [...] because whites think, 'Oh, you're taking something away from us.'" In a similar vein, one of the newly qualified teachers wrote on her questionnaire: "I feel that if black people wish to become teachers, positive discrimination should not take place, as they can then be accused of entry through the back door. They need to be in a position to stand their ground, [that is] with equal qualifications to other students".

Despite their reservations about affirmative action, the participants accepted the need for measures to make teaching a more ethnically diverse occupation. When asked to comment on policies directed towards this goal, both student and newly qualified teachers alike identified a number of barriers to ethnic minority recruitment and retention. Not surprisingly, the so-called "image problem" facing the teaching profession in the UK (Mahony & Hextall, 2000; Department for Education & Employment, 1998) figured prominently in their replies. Consider, for example, the following observations about teachers' status, pay and conditions of work made by respondents to the PGCE survey:

> As regards attitudes to teaching, I believe it is respected by non-professionals but derided by other professionals as being a "soft option" for people unable to handle the pressures of commerce. (Male, 'Asian Indian')

> I have found that in my ethnic community teaching is not considered a respectable job. It is stressful: it pays little (which is the main reason it is not liked); and much of the work is brought home. (Female, 'Asian Indian')

> It is my belief that teaching is often seen as a second option after seeking other career paths. Pay is often linked to the prestige of the job. Teachers' pay is often seen as very low. (Female, 'Black Other')

In addition to these factors – which, arguably, influence all potential recruits to teaching, irrespective of ethnicity - respondents to the PGCE survey also wrote about the specific barriers confronting those from ethnic minority groups. Echoing the findings of Osler (1997), their comments suggested that many were deterred from even contemplating a teaching career by the prospect of facing social isolation or racism in schools:

> Judging from own experience and that of some close ethnic minority teachers [...] white staff impose an inferiority complex. They [minority teachers] have to be constantly on guard, work twice as hard, need to prove that they are as good as anyone else. Promotion and acknowledgement of contribution is withheld. Especially Asians do not feel comfortable in staff rooms etc.. But in schools with mainly ethnic pupils and staff, the atmosphere is more relaxed and comfortable. (Female, 'Pakistani').

> The lack of an adequate level of discipline amongst students is the greatest barrier ethnic minorities have towards teaching. Many students are unruly and take advantage of racism in the society to make the life of most teachers from the ethnic minorities very stressful (Male, 'Black African').

> There is a lot of racism in East London. To stand up and teach in these areas can be difficult. I have been working as a play teacher for many years [and] I am able to work with disruptive children (Female, 'Asian Indian').

Teaching in mainly white schools

The follow-up interviews revealed that the PGCE students had been understandably cautious about undertaking placements in mainly white schools where they anticipated problems because of their ethnicity. A number recounted their feelings of trepidation about teaching in such settings when responding to the question, "What have been your experiences during school placement?". For example, Jill, a student teacher of mixed heritage, made this observation: "The first time I went to that school, I saw all these white faces. And it really scared me because it was the first time I was ever in a situation like that." In a similar vein, Nuzrat ('Asian Pakistani') described how vulnerable she had felt during her initial encounters with students at a mainly white boys' grammar school: "It took a lot of time for them to get used to me. You know when people stare - and they just stare at you! [They're] not being rude: they're just getting used to your face and they're getting used to you". Despite their concerns about teaching in mainly white schools, several interviewees indicated that their anxieties had proved to be ill founded and that their experiences in such settings had been positive:

> I was more hesitant about that school and wondered how the students would react to having an Asian teacher. But they were wonderful! The support I received [...] was phenomenal. Not even once during the whole placement did I feel that there was any race issue from the head of department or from the other teachers (Ayesha, 'Asian Pakistani').

However, this was not the case with other interviewees. Four made allegations of racism against teaching staff. In addition, four students also confided that they had been subjected to racist abuse from white children or their parents.

Of these, Liz ('Other'), recalled the following incident that had occurred in a predominantly white primary school:

> [The children would ask]: "Why have you got a brown face?"; "Why have you got brown hands?". They just weren't used to black people at all. They thought I was an alien, or something. And they'd go home and tell their parents. A lot of the parents [...] didn't really want me there and they'd tell the children [who] would come back and say, "My mum says you're this and my dad says you're that - and you're not allowed to touch me!"

As the interviews with PGCE students revealed, the prospect of experiencing social isolation and racism during school placement were common causes of anxiety and stress. However, in the main, the students' actual experiences in schools were far better than they had anticipated and the number of reported incidents of racial harassment was relatively low. Aware of the possibility of underreporting, the research team decided to investigate the matter further in the survey of newly qualified teachers.

The respondents were invited to engage with a number of propositions relating to their experiences of racial harassment and social isolation during their pre-service education and training and induction (see Table 1). They were also asked to provide written details of any incidents of racism that had been directed towards them.

When invited to reflect on their experiences during the PGCE, two fifths (41%) of the newly qualified teachers agreed or strongly agreed that they had been anxious about the reception they would get in the staff room. More than a third (37%) indicated that they experienced problems during training that were different from those faced by majority students. Although a quarter (27%) reported that they had encountered racism in their placement schools, relatively few (13%) believed that minority student teachers have a harder time with discipline than their majority counterparts. Turning to the induction year, only a small minority (18%) complained that they had experienced some form of racial harassment in their first posts and, in almost every case (15%), children were identified as the perpetrators. Twenty-six percent of those surveyed provided details of the alleged incidents. As following comments illustrate, accusations of racism were occasionally directed against school staff and parents:

> [During] school placement I was referred to on a number of occasions as "The Blackie Teacher" and "The Paki Woman" by parents of children who were attending the school (Female, 'Asian Indian')

> The first time I met my class, one child, during the carpet session, decided to stand up repeating the word "Paki" (Female, 'Asian Other').

> I worked in a predominantly working class school. I have been mistaken (by other teachers) for a classroom assistant. Furthermore, students have been surprised to see a black teacher. I find [that] I have to tell them I have been to university for four years and [that I] am qualified. Initially, I had lots of racist comments, like students singing Reggae when I walked past, or shouting "Bob Marley". Once a student called me a "nigger". Being a woman and black confuses the children, who have negative attitudes and stereotypical views of both these groups. I find I have to work harder for respect, unlike [...] a white male (Female, 'Other').

Matching by ethnicity

All inexperienced teachers, regardless of ethnicity, are likely to face difficulties in managing classes. Furthermore, as the above vignettes suggest, minority teachers may have to deal with the added stress of handling incidents of overt racism when faced with direct challenges to their authority in the classroom. As well as finding that they may have to 'work harder' than their majority colleagues to secure the respect of disaffected white youth, they may also discover that their authority is called into question by African-Caribbean or Asian students, who view teaching as a quintessentially white domain and dismiss minority members of the profession as 'selling out' (Pole, 1999). Consider, for example, the observations made by the following PGCE student during interview as he described reactions of black children towards him during the early stages of a school placement:

> [At first] they're trying to weigh you up: "Are you one of us, or are you not?" "Are you a coconut, basically?" [...] They try you out to find out what you're like initially, [during] the first week or so. Until the word gets around, "We've got a new black guy in the school", everyone is looking out for you. [...] They eye you up, to try and find out how you'll respond. And [when] the word gets round that, "He's not a bad bloke; he's one of us", then you're OK! I dread to think what it would be like in some schools if you were "not one of them" (Kevin, 'Black Other')

Another PGCE student, Deepak ('Asian Indian') had also experienced similar antipathy during one of his placements from African Caribbean school students. He told the researcher that although racism was "relatively rare" among staff at the school, black students often accused teachers of discriminatory behaviour and "used racism as an excuse for disruptive behavior". Deepak went on to state that although similar allegations had been leveled against him, he found them less intimidating than his white colleagues:

> If I'm called racist by a kid it doesn't really bother me. It's very hard, given the fact that I come from an ethnic minority background. So, I don't feel

Table 1

Newly Qualified Teachers' Experiences of Racial Harassment and Social Isolation During Their Pre-service Education and Induction

		N	Percent
Pre-service Education and Training			
I was anxious about the reception I would get in the staff room.	Strongly agree	18	12.2
	Agree	42	28.4
	Neither	28	18.9
	Disagree	36	24.3
	Strongly Disagree	24	16.2
As an ethnic minority student, I encountered problems during my training that were different from those faced by majority students.	Strongly agree	25	16.9
	Agree	29	19.6
	Neither	24	16.2
	Disagree	37	25.0
	Strongly Disagree	33	22.3
I encountered racism during school placement.	Strongly agree	14	9.4
	Agree	26	17.4
	Neither	26	17.4
	Disagree	41	27.5
	Strongly Disagree	42	28.2
Ethnic minority students generally have a harder time with discipline.	Strongly agree	6	4.0
	Agree	14	9.4
	Neither	38	25.5
	Disagree	54	36.2
	Strongly Disagree	37	24.8
Induction Year			
I have experienced racial harassment from students.	Strongly agree	6	4.1
	Agree	16	11.0
	Neither	24	16.4
	Disagree	35	24.0
	Strongly Disagree	65	44.5
I have experienced racial harassment from staff.	Strongly agree	2	1.4
	Agree	2	1.4
	Neither	20	13.7
	Disagree	38	26.0
	Strongly Disagree	84	57.5

> threatened by it at all. But I cannot think of a single English teacher at the school that does not feel threatened. [...] When I told my teacher [mentor], I'd been called "racist", she asked: "Why didn't you fill in an equal opportunities form? This is really serious!" To her it was a big deal and you had to get the paper work written up!

In view of these additional pressures, it came as no surprise that some of our interviewees expressed unease about being construed as 'role models' for minority children. Of these, Fiazza, a student of Pakistani descent, described a placement in a multi-ethnic school where she encountered a degree of hostility from Asian girls. She felt that this group tended to accord her white colleagues (particularly men) more respect than they did female Asian staff. She claimed: "I don't think they can take anybody seriously, especially if you are female, unless you're white and wearing a suit." She went on to explain that she felt far more comfortable during a subsequent placement in a mainly white school, where she was not obliged to adopt the mantle of a 'role model' for Asian children:

> I think that not only can you be a role model because of the color of your skin but also [because of] your personality. And if you don't suit your own group, then why be pushed into that? I felt a little bit pushed. [...] I find it really disappointing to say I fitted into a white school, but not an Asian school. But the fact of the matter is, that that's how it happened!

Another Asian student had encountered similar difficulties while working in a predominantly Asian school. She believed that that her authority as a teacher had been undermined as a result of having the same (or a similar) background to the children.

> I wanted to be a role model [but] that's not the way it turned out to be! They [Asian students] didn't see me like that; they didn't see me as a role model. They saw me [...] like an older sister or auntie. So instead of giving me the respect - which I thought I'd get - it didn't turn out that way. They were treating me like, "Oh, she's just one of us". The disrespect was quite unbelievable to be honest! (Yasmin, 'Asian Bangladeshi')

In common with Fiazza, Yasmin said that she felt more at ease working in a mainly white school, where the children had been far more accommodating towards her: "They accepted me wholeheartedly and quite warmly as well. Like one of the boys asked me, 'Why can't I see your hair miss?' I like the openness; I like the honesty of it! So I'd tell him, 'You know [it's] because of my faith; it's what I believe'."

Conclusions

In this paper, an attempt has been made to show that recent policy measures to address the under representation of ethnic minority groups in teaching have been predicated upon largely unexamined assumptions about the salience of such teachers as 'role models' for minority children. In official rhetoric, the matching of teachers and children by ethnicity has been justified on the grounds that it is both a precursor to improved academic performance and improved attitudes towards education. Despite the emphasis in official discourse on the need for positive 'role models' to inspire minority youth, relatively few participants in our study referred directly to this issue when accounting for their decisions to teach. The primary motivation for most recruits was the intrinsic satisfactions of a teaching career. Although some expressed reservations about affirmative action, an overwhelming majority were supportive of current moves to make teaching a more inclusive occupation and believed that minority teachers have a vital part to play in the education all children, whatever their ethnicity. Notwithstanding this, our participants often voiced concerns about working in mainly white schools, fearing that they could be 'cold-shouldered' by staff or find themselves exposed to racism. These anxieties were common causes of stress throughout the PGCE. Although some reported that they had experienced racial harassment in schools - either during the PGCE or in their first posts - their actual experiences were often better than they had anticipated. The findings suggest that the problem lies more in the expectation than in the reality. While many of the participants expressed a preference for teaching in multi-ethnic schools, such settings presented them with additional challenges as minority children openly questioned their authority in the classroom. Matching teachers and children by ethnicity is not a sufficient condition for winning their respect or gaining recognition as a 'role model'.

The research offers a number of pointers to policy. First, it underlines the need for a change in the rhetoric employed to justify measures to increase ethnic minority recruitment to teaching. The restricted view of minority teachers as 'role models' for minority children needs to be replaced with a more inclusive form of representation. Gordon (2000) puts it so aptly when she states: "Positive role models and the removal of stereotypes are as important for White children as for children of color. Teachers of color are pivotal in this awesome task of breaking down centuries of misinformation and fear that have kept us separated and ignorant" (p.2).

Secondly, the research also points to a need for change in pre-service teacher education. Among other things, it suggests that steps should be taken to engage with minority students' anxieties about working in mainly white schools. Thus, in the run up to the initial school placement, seminars could be used to explore a range of issues relating to antiracism, citizenship and justice. In these seminars, all students, irrespective of ethnicity, should be encouraged to reflect on their own assumptions and prejudices and to articulate their concerns.

The experience of both PGCE and school staff should be used as part of a process to build the confidence of all students to deal with new and challenging situations. The impact of this will be felt not only during the PGCE but also throughout the student's subsequent teaching career.

References

Allen, A. (2000). *The role model argument and faculty diversity.* Available: http://onlineethics.org/abstracts/fac-diverse.html

Bonnett, A. & Carrington, B. (2000). Fitting into categories or falling between them? Re-thinking ethnic classification. *British Journal of Sociology of Education, 21*, 485-498.

Carrington, B., Bonnett, A., Demaine, J., Hall, I., Nayak, A., Skelton, C., Smith, F., Short, G. & Tomlin, T. (2001). *Ethnicity and the Professional Socialization of Teachers. Report to the Teacher Training Agency.* London: Teacher Training Agency. Available: http://www.canteach.gov.uk//research/tsr/carringtonrep/newcastle_research.pdf

Department for Education & Employment (1998). *Teachers: Meeting the Challenge of Change.* London: HMSO.

Department for Education & Employment (2000). Press notice. #2000/0061.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Education Longitudinal Study of 1988. *Industrial and Labor Relations Review, 48,* 547-561.

Garewal, G. (1999). Experiences of racism in initial teacher training. *Multicultural Teaching. 17*, 17-23.

Ghouri, N. (1997). Wanted: The best minority teachers. *Times Educational Supplement,* 7 November. Available: http://www.tes.co.uk/.

Gillborn, D. & Mirza, H. (2000). *Educational Inequality. Mapping race, class and gender: a synthesis of evidence.* London: Office for Standards in Education (OfSTED).

Goffe, L. (2001). Caribbean staff can make it in New York. *TimesEducational Supplement.* April 11. Available: http://www.tes.co.uk/

Gordon, J. (2000). *The color of teaching.* London: Routledge Falmer.

Jones, C., Maguire, M. & Watson, B. (1997). The school experience of some minority ethnic students in London schools during initial teacher training. *Journal of Education for Teaching, 23*, 132-144.

Lahelma, E. (2000). Lack of male teachers: A problem for students or teachers? *Pedagogy, Culture and Society, 8*, 173-85.

Latham, A. (1999). 'The Teacher-student mismatch', *Educational Leadership, 56.* Available: http://www.ascd.org/readingroom/edlead/9904/extlatham.html

Mahony, P. & Hextall, I (2000). Reconstructing teaching: Standards, performance and accountability. London: Routledge Falmer.

Majors, R. (Ed.) (2001). Educating our black children: new directions and radical approaches. London: Routledge Falmer.

Osler, A. (1997) The lives and careers of black teachers: Changing identities, changing lives. Buckingham: Open University Press.

Pepperell, S. & Smedley, S. (1998) Calls for more men in primary teaching: Problematizing the issues. *International Journal of Inclusive Education, 2*, 341-57.

Pole, C. (1999) Black teachers giving voice: Choosing and experiencing teaching. *Teacher Development, 3*, 313-328.

Quiocho, A. & Rios, F. (2000). The power of their presence: Minority group teachers and schooling. *Review of Educational Research, 70*, 485-528.

Robbins. M. M. (1995). Black students in teacher education, *Multicultural Teaching, 14*, 15-22.

Runnymede Trust (2000). The future of multi-ethnic Britain: The Parekh report. London: Profile.

Sewell, T. (1997). *Black masculinities and schooling.* Stoke-on-Trent: Trentham.

Steele, C. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797-811.

Siraj-Blatchford, I. (1991). A study of black students' perceptions of racism in initial teacher education. *British Educational Research Journal, 17*, 35-50.

Weiner, L. (2000). Research in the 90s: Implications for Urban Teacher Preparation, *Review of Educational Research, 70*, 369-406.

Wolverhampton Race Equality Council Consortium (WRECC) (1999). *Recruitment and Retention of Teachers from the Ethnic Minority Communities.* Wolverhampton: WRECC.

Bruce Carrington is a Professor of Education at the University of Newcastle, UK.

Footnote

[1]This following schedule was employed: What made you want to become a teacher? Why did you choose this training institution? Has it been the right choice for you? How happy have you been with the quality of support provided on your PGCE course (both in the University and in your placement schools)? Would you recommend this course to others? In your view, how well does the institution meet the needs of ethnic minorities? What specific provisions, if any, is made available for them? What, if anything, could the institution do to make itself more attractive to ethnic minorities? What have been your experience during school placement? How happy are you with 'official' descriptors of ethnicity, such as those used on the Graduate Teacher Training Registry application form? What advice would you offer the Government about its current policy to boost ethnic minority recruitment to teaching?

Author Note

I would like to thank all those who took part in the study and acknowledge the contribution of other members of the research team: Alastair Bonnett, Ian Hall, Anoop Nyak, Chris

Skelton and Fay Smith (University of Newcastle); Jack Demaine (Loughborough University); and Geoffrey Short (University of Hertfordshire). I am also grateful to the Teacher Training Agency for funding the research. Finally, I would like to record my thanks to Donald Biggs and the anonymous reviewers for their constructive feedback on an earlier draft of this article.

For further information contact the author, Bruce Carrington, Department of Education,University of Newcastle, Newcastle uponTyne,NE17RU,UnitedKingdom.Email: Bruce.Carrington@ncl.ac.uk

Developing Authentic Instruction in the Social Studies

Patricia G. Avery
and
Carol Freeman
University of Minnesota

Dana Carmichael-Tanaka
Minneapolis Public Schools

This study describes a professional development program characterized by strong connections between content knowledge, pedagogical content knowledge, and pedagogical knowledge. Forty-five secondary social studies teachers from one urban school district participated in the six-month program, Linking Authentic Instruction to Students' Achievement Using Peer Coaching. Teachers participated in content seminars led by social scientists, pedagogical content seminars led by master teachers and teacher educators, and peer coaching sessions in which colleagues observed one another's teaching.

The project was grounded in the research on authentic student achievement conducted by Fred Newmann and his associates at the University of Wisconsin in the 1990s (Newmann & Associates, 1996). The goal was to improve teachers' ability to engage in authentic instruction and assessment, and to enhance the quality of students' performance. Observations and document analyses were used to rate the "authenticity" of instruction, assessment, and student performance. Results indicated significant improvement in teachers' instruction and students' achievement, but little change in the authenticity of the assessment tasks used by the teachers.

Introduction

The purpose of this study is to examine the effects of an inservice program, *Linking Authentic Instruction to Students' Achievement Using Peer Coaching* (hereinafter referred to as the *Authentic Instruction Project*) on the authenticity of teachers' instruction, assessment tasks, and their students' work. The study extends the research on authentic student achievement conducted by Fred Newmann and his associates at the University of Wisconsin in the mid-1990s (Newmann & Associates, 1996; Newmann, Secada, & Wehlage, 1995), and is grounded in a model of professional development that connects pedagogical knowledge with content knowledge through peer coaching.

Related Literature and TheoreticalFramework

Authentic Intellectual Achievement

Numerous studies indicate that thoughtful class discussions about significant social and political issues, in which students engage in substantive conversation, challenge one another's thinking, and build a collective understanding of problems, alternatives, and consequences, are rare (Kahne, Rodriguez, Smith, & Thiede, 2000; Newmann & Associates, 1996). In the mid-1990s, Fred Newmann and his colleagues at the University of Wisconsin's Center on Organization and Restructuring Schools (CORS) devoted considerable efforts to thinking about the nature of meaningful teaching and learning. They proposed an integrated conception of "authentic intellectual achievement" with three elements: *student construction of knowledge* through *disciplined inquiry* that has *value beyond the classroom*. They reasoned that authenticity could be reflected in three areas: assessment tasks, instruction, and student performance. In their model, each area has different but parallel indicators of authenticity. For example, to gauge the "authenticity" of instruction, the framework suggests asking questions such as the following: To what degree does instruction involve students in manipulating information and ideas to arrive at conclusions that produce

new meaning and understandings for them (higher order thinking)? To what extent does instruction address the central ideas of a social studies discipline or topic (disciplined inquiry)? To what degree are students involved in extended conversational exchanges with the teacher and/or their peers about subject matter (substantive conversation)? To what extent does the lesson provide connections to the larger social context in which students live (value beyond the classroom)? Using similar questions, CORS researchers developed standards and scoring criteria for assessing the authenticity of teacher instruction, assessment tasks and student performance.

In an extensive study of authentic intellectual achievement in 24 restructured elementary, middle, and high schools, CORS researchers found a high correlation between level of authentic pedagogy and student performance, i.e., the higher the quality of instruction and assessment, the higher the quality of student performance. Their observations of classrooms and review of student work, however, suggested that in most classrooms, the authenticity of pedagogy (i.e., instruction and assessment) was fairly low. When teachers *did* implement more authentic pedagogy, all students benefited, regardless of ethnicity, socioeconomic status, gender, or achievement level. Subsequent studies have supported the original findings of the CORS researchers (Avery, 1999; Newmann, Bryk, & Nagaoka, 2001; Palmer & Avery, 1999).

In a recent study of more than 5,000 elementary teachers in Chicago, Smith, Lee and Newmann (2001) found that teachers who had been involved in high quality professional development activities were more likely to use instructional strategies that are associated with authentic pedagogy (i.e., less didactic instruction and more interactive instructional strategies). The Chicago teachers' report of good professional development was more strongly associated with interactive instructional approaches than either a degree beyond a Bachelor of Arts, or more than 40 credit hours of work in the teachers' primary subject matter. Although the difficulty of changing teachers' instructional practices has been widely documented (Firestone, Mayrowetz, & Fairman, 1998), high quality professional development has usually been central to any significant improvements in pedagogy.

Professional Development

Effective professional development generally (a) links content and pedagogy, (b) reflects and models interactive learning, and (c) supports and sustains professional community over a period of time.

Links Content and Pedagogy

Teachers report greater increases in knowledge and skills and more positive benefits for student learning when professional development focuses on subject area content and on students' learning of particular subject matter than when it is focused on teaching techniques, such as simulations or role play (Garet, Porter, Desimore, Birman, & Yoon, 2001). Teachers' beliefs about the nature of subject matter influence the degree to which professional development impacts their instructional practice. It *matters* whether teachers believe their subject area discipline is characterized by established, static knowledge, or by evolving, dynamic knowledge (Loucks-Horsley, Hewson, Love & Stiles, 1998). For example, teachers who believe the study of history to be simply knowledge of the past are less likely to be open to professional development that advocates inquiry-based learning, problem solving, student investigation and discovery, and application of knowledge than are those with a more constructivist orientation.

Reflects and Models Interactive Learning

Effective professional development takes into account how teachers learn. Lieberman (1995) suggests that the ways teachers learn may be more like the ways students learn than we have previously recognized. Other researchers agree that instructional methods for teachers should mirror the methods teachers are encouraged to use with students (Darling-Hammond & Ball, 1998; Darling-Hammond & McLaughlin, 1995), and that instruction of teachers should: (a) build on teachers' current knowledge, skills, and beliefs; (b) allow teachers to construct their own knowledge through interactive learning; (c) provide opportunities for teachers to collaborate with and support one another (e.g., observing one another teach, analyzing students' work together); and (d) structure continuous support to address new issues and problems that inevitably arise from classroom practice.

Supports Professional Community

Teachers need to have opportunities "to discuss, think about, try out, and hone new practices" (Lieberman, 1995, p. 593). This type of inquiry and examination of practice requires a focus on communities of practice rather than on individual teachers (Darling-Hammond & McLaughlin, 1995). Teachers need supportive, collegial communities when inquiring into significant questions about subject matter, as well as into questions about learning and pedagogy (Loucks-Horsley et al., 1998). If teachers are to take risks and experiment with new ways of teaching, they need to feel supported. One approach to enhancing collegiality is through peer coaching programs.

Peer coaching. Over the past 20 years, peer coaching has emerged as an important tool in professional development (Joyce & Showers, 1980; Kohler, McCullough, Crilley, & Shearer, 1997). Peer coaching involves pairs or teams of teachers observing one another's instruction and then providing one another with constructive feedback. Research suggests that peer coaching can be very effective in improving teaching techniques and enhancing collegiality (Kohler et al., 1997).

Method

This project integrates current research on authentic pedagogy with research on effective professional development. The study addresses the following questions: (a) How effective was the professional development model in promoting authentic instruction and assessment? and (b) What were the effects of authentic instruction and assessment on the learning and achievement of students?

The Authentic Instruction Project

The *Authentic Instruction Project* was designed to improve teachers' ability to engage in authentic instruction and assessment. The project involved a strong collaborative effort between a large inner-city school district and a major university. The project made use of the concepts, standards and scales developed by CORS to design a series of professional development seminars for secondary social studies teachers.

Teachers involved in the project participated in four professional development seminars between January 2001 and June 2001. The first seminar was devoted to familiarizing teachers with the authentic pedagogy framework, the second and third seminars focused on developing teachers' content knowledge through model lesson demonstrations, and the final seminar revisited the authentic pedagogy framework and provided time for teachers to engage in in-depth reflection on their work. Content specialists in civics, geography, and U.S. history led the second and third seminars. The teachers critiqued the experts' "model lessons" in terms of the standards for authentic instruction. Throughout the project, pairs of teachers were given opportunities to observe one another's teaching and debrief their instruction according to the criteria for authentic instruction. Training in peer coaching helped teachers develop the skills to give one another constructive feedback.

The professional development model was thus based on *Pedagogical Knowledge* (the authentic pedagogy framework) and *Social Science Content Knowledge* (the model lessons developed by content area experts) and *Peer Observations of Instruction*. Each of these components played a very important role in the teachers' professional development.

Design

To address the research questions, a one-group pre/post design was used to collect data on teachers' instruction, teacher-designed assessment tasks, and student work at the beginning and the end of the project. Data were gathered through student and teacher surveys, classroom observations and document analysis.

Teachers selected a target class on which they would concentrate to improve instruction, assessment tasks, and student work. From these target classes, data were collected on teachers' instruction, teacher-designed assessment tasks, and student work at the start of the project and at the end of the project.

Data Sources

Classroom instruction. Project staff observed the target classes of social studies teachers at the beginning of the project (January-February) and three to four months later near the end of the school year and the project. All observers were trained in the standards for authentic instruction. Observers spent the entire class period scripting what they saw and heard. Later, they rated what they had observed according to the standards for authentic instruction, i.e., the degree to which students are given the opportunity to (a) use higher order thinking, (b) develop deep knowledge, (c) engage in substantive conversation, and (d) make connections to the world outside the classroom. Each standard is defined as a multi-dimensional construct on a five-point scale. The scoring criteria for each standard are defined in considerable detail to assist observers in rating what they observe (see Newmann et al., 1995).

Assessment tasks. Eight standards were used as scoring criteria for the assessment tasks. Six standards had a three-point scale and two standards had a four-point scale. Assessment tasks are considered more authentic when the task required students to (a) organize information, (b) consider alternatives, (c) show understanding of disciplinary content, (d) use disciplinary processes such as inquiry and research, (e) elaborate through written communication, (f) address a problem connected to the world beyond the classroom, and (g) communicate to an audience beyond the classroom. An eighth standard that was not in the original CORS criteria was included: clarity of directions. In our previous work with teachers, we found that although a task was often *intended* to require students to consider alternatives, the directions for the task could be confusing and difficult to follow.

Student performance on assessment tasks. Student work, collected at the beginning and the end of the project, was rated using three standards, each having a four-point scale. Student work is considered more authentic when students (a) demonstrate strong analytical abilities, (b) show an understanding of significant disciplinary concepts, and (c) convey their ideas through extended writing.

Participants

Teachers. All social studies teachers in one school district were invited to join the *Authentic Instruction Project*. Forty-five teachers (approximately 33% of the total number of social studies teachers in the district) participated in some or all of the project activities. Thirty-one teachers were observed at both periods of time. Twenty teachers submitted assessment tasks at the beginning of the project and 17 submitted assessment tasks at the end of the project. Sixteen

Table 1

Demographic Characteristics of Students Participating in the Project

	Middle School (n = 102)	High School (n = 179)	Total (N = 281)
Percent female	55.7	56.0	55.9
Percent white	12.1	42.2	30.9
Eligible free/reduced lunch	86.6	54.3	66.4
Born in the U. S.	74.5	74.1	74.2

Note: Of the 281 students, from 14 to 22 students skipped one or more of these questions

teachers submitted assessment tasks and student work at both time periods.

Male and female teachers were almost equally represented in the group; however, middle school teachers tended to be female, a bit younger with fewer years of teaching in their current school, and fewer years in teaching generally. High school teachers were more likely to be male, white, and a bit older. The average participant was 36.7 years old and had nine years of teaching experience; 65.3% of the participants had an advanced degree. Sixty-three percent of the project teachers were female and about 66.7% were white.

Students. Observers noted a total of 690 students in classrooms during the pre-observations, and 650 during the post-observations. Twenty teachers submitted work from a total of 331 students at the beginning of the project and work from a total of 224 students at the end of the project. There were 98 students for whom there was work from both time periods. This district is urban with a highly mobile population, which may account for the low percentage of students with both pre and post project completion of assessment tasks.

As shown in Table 1, middle school students were less likely to be white, more likely to say they qualified for free or reduced lunch, and more likely to report getting lower grades in school and in social studies than the high school students. Both groups reported getting lower grades in social studies than in other subjects. Approximately one-fourth (26%) of the students were born outside the United States.

Analysis

Reliability. Developers of the scoring criteria found good inter-rater reliability for the ratings on instruction (.77 to .79 for the four standards) (Newmann et al., 1995). They found lower inter-rater reliability for the scoring of student work (range from .49 to .62). The percent of scores that were in exact agreement or within one scale point ranged from 88% to 97% on all standards of instruction and student work. For this study, one research assistant rated all of the student work after being trained by the research coordinator. On a random sample of 10% of the student work, the research coordinator and the research assistant were within one scale point for all standards.

Statistical procedures. All comparisons of instruction, assessment tasks, and student work scores were made using matched pairs t-tests comparing the score from the beginning of the project to the score for the same person at the end of the project. With instruction, all teachers who were observed at the beginning of the project were observed at the end. However, not all of these teachers submitted assessment tasks and student work at both time periods. Particularly with student work, the number of matched pairs (98) was much less than student work at the beginning (331) and at the end (224). For this reason, independent samples t-tests were also run keeping all assessment tasks and student work in the analysis. The results of the independent samples t-test were very similar to the results for matched pairs, and, therefore, the results for the matched pairs tests are included here. The formula used to calculate effect sizes is one suggested for matched samples situations. It is $E=m_1-m_2(\sigma_{x1-x2}(\sigma_{x1-x2}=2(1-p))$ with p being the correlation of the two scores.

Results

Impact on Instruction, Assessment and Student Performance

Instruction. The scoring criteria for instruction consists of four standards: higher order thinking, deep knowledge, substantive conversation, and connections to the world beyond the classroom. In looking at changes between the first observation and the second, we saw significant improvements in the extent to which teachers engaged in authentic instruction (see Table 2). In particular, teachers showed improvement in engaging students in "higher order thinking" and developing "deep knowledge" of subject matter. There were also some increases in "substantive conversation," although this standard began with and continued receiving the lowest scores.

Table 2

Comparison of Authenticity of Instruction, Beginning of Project to End of Project

	Potential Range	Beginning Mean	Ending Mean	Effect Size
Standard 1. Higher Order Thinking	1 - 5	2.39	3.16	.81***
Standard 2. Deep Knowledge	1 - 5	2.42	3.03	.69**
Standard 3. Substantive Conversation	1 - 5	1.90	2.29	.35*
Standard 4. Connections to the World Beyond the Classroom	1 - 5	2.29	2.71	.36*
Total Authentic Instruction Score	4 - 20	9.03	11.19	.61**

*$p = .08$, **$p < .01$, ***$p < .001$ using a matched pairs *t*-test, $n = 31$

Note: Effect size = $E = m_1 - m_2 / s_{x1-x2} (s_{x1-x2} = 2(1-p))$. Researchers differ as to what constitutes a significant effect size. Jacob Cohen (1988) suggests that an effect size of .20 can be considered small; an effect size of .50 medium, and an effect size of .80 large. Rosenthal and Rosnow (1984), however, categorize an effect size between .1 and .3 as small; between .3 and .5 medium; and over .5 large.

Assessment tasks. Teachers submitted copies of assessment tasks (assignments) they had used with their target class at the beginning of the project period and at the end of the project period. As shown in Table 3, the quality of the assessment tasks in terms of authenticity improved on every standard except "consideration of alternatives" and "elaborated written communication." The substantial decrease in the amount of written work required (effect size = -.43) may simply indicate that teachers gave less extensive assignments toward the end of the year. At both the beginning and the end of the project, the greatest weaknesses across the tasks were in not requiring students to consider alternatives and not having students share their work with audiences beyond the classroom. Though only the "clarity" standard showed statistically significant improvement, effect sizes of several other standards range from small to medium: "disciplinary content," "disciplinary process," and "audience beyond the school."

Student performance on assessment tasks. Teachers provided copies of student work resulting from the assessment tasks that they used with their target class before the project began and student work resulting from the assessment task they used toward the end of the project. Student work is considered more authentic when students demonstrate strong analytical abilities, show an understanding of significant disciplinary concepts, and convey their ideas through extended writing. In the *Authentic Instruction Project,* students demonstrated gains on each of these standards from the beginning to the end of the project (see Table 4). The improvement ranges from medium to large as indicated by the effect sizes. Without a comparison group, it is unknown what portion of these gains would have been evident as part of normal student growth over a six-month period.

Discussion

The goal of the *Authentic Instruction Project* was to improve instruction and assessment, and of course, student achievement. This project used the concepts of authentic pedagogy and achievement as the foundation for a professional development model. Ultimately, the focus was on getting students to think more deeply and to become more engaged in substantive conversation about the key concepts included in social studies subject matter, as well as making more connections beyond the classroom.

Using the authentic pedagogy framework developed by Fred Newmann and his colleagues at CORS (Newmann et al., 1995; Newmann & Associates, 1996) both as a tool for professional development and as the measure of authentic pedagogy, we noted statistically significant improvement in the quality of teachers' instruction and students' work.

We believe the inclusion of all three dimensions of the professional development model (content, pedagogy, and peer coaching) was central to the results. The project provided training in pedagogy through content, followed by the on-site application using peer coaching. The peer coaching component seems to have provided teachers with strong support as they worked to improve their instruction. Teachers were not simply encouraged to observe each other. Teachers were given training in coaching so that they would be able to observe and provide feedback to each other using

Table 3

Comparison of Authenticity of Assessment Tasks, Beginning of Project to End of Project

	Potential Range	Beginning Mean	Ending Mean	Effect Size
Standard 1. Organization of Information	1 - 3	2.38	2.44	.11
Standard 2. Consideration of Alternatives	1 - 3	1.94	1.88	-.10
Standard 3. Disciplinary Content	1 - 3	2.25	2.44	.39
Standard 4. Disciplinary Process	1 - 3	2.19	2.44	.55
Standard 5. Elaborated Written Communication	1 - 4	3.31	2.94	-.43
Standard 6. Problem Connected to the World Beyond the Classroom	1 - 3	2.06	2.06	.00
Standard 7. Audience Beyond the School	1 - 4	1.13	1.25	.32
Standard 8. Clarity (added for this project; not part of the framework designed by Newmann et al., 1995)	1 - 3	2.25	2.56	.65*
Total Authentic Task Score	8 - 26	17.50	18.00	.16

*$p < .02$, using a matched pairs t-test, $n = 16$

the criteria for authentic instruction. Participants were also given enough release time so that they had ample opportunity to observe one another.

The content seminars were developed with the criteria for authentic instruction in mind. Initially, this was done to model the framework, but as the project progressed, it became apparent that the criteria for authentic instruction permeated all project activities. It shaped the planning process for each of the four seminars; it was the tool for debriefing the content lessons during seminars two and three; it redefined the peer coaching training; and it was the instrument with which the researchers did pre and post observations. This created a powerful set of cohesive experiences for teachers.

The three key features of this project—content, pedagogy, and peer coaching—are grounded in theory and practice. Together, they form a powerful model for professional development. That this project evidenced statistically and educationally significant changes in teacher instruction and student performance in a short six-month period suggests that the general model of professional

Table 4

Comparison of Authenticity of Student Performance, Beginning of Project to End of Project

	Potential Range	Beginning Mean	Ending Mean	Effect Size
Standard 1. Analysis	1 - 4	1.94	2.55	.83***
Standard 2. Disciplinary Concepts	1 - 4	1.94	2.41	.68***
Standard 3. Elaborated Written Communication	1 - 4	1.92	2.29	.48***
Total Authentic Performance Score	3 - 12	5.80	7.24	.75***

***$p = .000$ using matched pairs t-test, $n = 98$

development, as well as the specific model employed here, deserve further examination.

References

Avery, P. G. (1999). Authentic assessment and instruction. *Social Education, 65*, 368–373.

Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership, 57*(8), 28–33.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Darling-Hammond, L., & Ball, D. L. (1998). *Teaching for high standards: What policymakers need to know and be able to do*. Retrieved December 14, 2001, from The National Commission on Teaching & America's Future and Consortium for Policy Research in Education Web site: http://www.nctaf.org/publications/index.html

Darling-Hammond, L., & McLaughlin, M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(8), 597–604.

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis 20*, 95–113.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership, 37*(5).

Kahne, J., Rodriguez, M., Smith, B., & Thiede, K. (2000). Developing citizens for democracy? Assessing opportunities to learn in Chicago's social studies classrooms. *Theory and Research in Social Education, 28*(3), 311–338.

Kohler, F., McCullough, Crilley, K., & Shearer, D. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90*(4), 240–250.

Lieberman, A. (1995). Practices that support teacher development. *Phi Delta Kappan, 76*(8), 591–596.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

Newmann, F. M., & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass.

Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* Chicago: Consortium on Chicago School Research.

Newmann, F. M., Secada, W. G., & Wehlage, G. G. (1995). *A guide to authentic instruction and assessment: Vision, standards and scoring*. Madison, WI: Wisconsin Center for Education Research.

Palmer, E., & Avery, P. G. (1999). *Professional development for authentic pedagogy in the social studies: An evaluation*. Minneapolis: University of Minnesota Center for Applied Research in Educational Improvement (CAREI).

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw Hill.

Smith, J. B., Lee, V. E., & Newmann, F. M. (2001). *Instruction and achievement in Chicago elementary schools*. Chicago: Consortium on Chicago School Research.

Patricia G. Avery is a Professor in the Department of Curriculum and Instruction at the University of Minnesota. Her research interests include social studies instruction and political socialization.

Carol Freeman is a Research Associate with the Center for Applied Research and Educational Improvement, University of Minnesota. Her research and evaluation focuses on the efficacy of teacher professional development and on school change efforts such as block scheduling.

Dana Carmichael-Tanaka has taught secondary social studies for over 15 years. The past four years, she has been the K-12 Minneapolis Social Studies Curriculum Specialist. She is a doctoral candidate in Educational Policy and Administration at the University of Minnesota.

Author Note

This project was funded by a grant from the Minnesota Department of Children, Families, and Learning.

Perceptions of Care Among Black Female High School Students

Erin McNamara Horvat
and
Kristine S. Lewis

Temple University

Researchers have consistently noted the importance of care in the lives of students. Care has been examined as a relational construct as well as one that has important cultural linkages. However, this research has not examined the role of social context in student perceptions of care. In this paper we demonstrate how students at different schools perceived care in fundamentally different ways. While the findings presented here are suggestive rather than conclusive, we highlight the importance of social context in student's abilities to perceive and receive care in educational settings. We suggest that future research ought to more closely examine the contextual factors that affect student's perceptions of care.

Studies on academic achievement (Levine, 1994; Polite, 1993), college aspirations (Glazer & Wughalter, 1991; Stage & Hossler, 1989) and college decision-making (Sevier, 1993) have all referenced the importance of care in the lives of students. With special emphasis on the one-to-one relationship between students and adults (Lareau, 1997; Stage & Hossler, 1989; Wells & Crain, 1997), researchers have consistently concluded that care plays a major role in shaping educational experiences.

However, little attention has been paid to the contextual dimension of care in educational settings. That is, while researchers have documented the importance of care in the lives of students, the context in which caring relationships are developed and nurtured has gone relatively unexamined. We argue that the context in which the caring relationship exists plays a very important role in students' perceptions of care and their ability to receive care in educational settings.

Varying definitions and conceptualizations of care exist in the literature. According to Mayerhoff (1971), caring is an involved, nurturing exchange between two people for the benefit of the cared-for individual. It is through caring relationships that adults teach students to be responsible for their own lives. Caring, for Mayerhoff, is more than the pairing of an available adult and a child. Rhodes and Davis's work (1996) suggests that voluntary mentoring relationships "may not be as influential or enduring as those that form naturally" (p. 221).

Noddings (1992) concurs with Mayerhoff (1971). Extending his argument, Noddings asserts that rather than viewing the student as a passive recipient of the care flowing from the adult, we recognize the student as an active participant. She places a condition on the relationship between the care-giver and the cared-for:

> In order for the relation to properly be called caring, both parties must contribute to it in characteristic ways. A failure on the part of either the care-er or the cared-for blocks completion of caring, and although there may still be a relation - that is, an encounter or connection which each party feels something towards the other - it is not a *caring* relation. No matter how hard teachers try to care, if the caring is not received by the students, the claim "they don't care" has some validity. (p. 15)

For Noddings (1992), care is a reciprocal construct. If the student does not recognize the behavior of the adult as a demonstration of care, a caring relationship does not exist. It is imperative that the student perceive the care of the adult, for there will be no other way for the caring relationship to affect the student's academic achievement or personal development.

Other scholars incorporate cultural linkages in their conceptualization of care. They affirm that a shared cultural heritage between the cared-for and care-giver assists in the recognition of care. This is especially important in the case of Black students (Dempsey & Noblit, 1993). Dempsey and Noblit outline the importance of Black teachers in the educational experiences of Black students. Employing Irvine's (1990) concept of cultural synchronicity, Dempsey and Noblit discuss how shared cultural heritage facilitates the development of caring relationships between teachers and students. Foster (1997) points to the faith that Black teachers have in Black students' abilities and the culturally-

informed understanding that Black teachers have of Black students' needs. [See also Delpit (1988) and Ladson-Billings (1994)]

We argue that school context is yet another way in which students' perceptions of care are mediated. Virtually unexplored terrain in the literature, research has been so intensely concentrated on the relationship between students and adults, that few researchers have pulled back their focus to look at the school contexts in which these relationships develop and flourish. This paper contributes to our understanding of students' caring relationships in school settings by considering variations in students' perceptions of care as mediated by school context.

The importance of care in the lives of the students in this study emerged during the course of data collection. The original goal of the ethnographic research that generated this data set was to understand how race and class influence college decision-making. In the process of conducting research to address this initial question, the importance of care in students' lives consistently emerged as a central construct. Interviews with fourteen Black, female, college-bound, high school seniors enrolled in three different high schools contained a strong theme of care. Using these data we explore the differences in the students' perceptions of care and illustrate how these variations may be influenced by school context.

Method

The research questions guiding the original qualitative study considered how an intersection of race and social class influence college choice. Data were collected between the months of October 1994 and June 1995 and consisted of fieldnotes, collected documents, and interviews. More than one hundred hours were spent at each of the three high schools. During these visits, relevant supporting materials were collected including college counseling information provided to students and parents, sample college recommendations written for students, general information generated by the college counseling offices, and the high schools' curriculum plans. The first author interviewed fourteen primary participants, each of their best friends and one of their parents. She also interviewed all of the high school counselors and principals at their respective schools. A total of fifty-three interviews, each approximately 90 minutes long, were completed.

The primary study participants were all college bound students with a grade point average between 2.75 and 3.5. They all had been or were currently enrolled in at least one honors level or Advanced Placement level class. The students' SAT I and PSAT scores ranged between 800 and 1270. The college counselor in each school identified potential study participants. The study was explained to each student individually and they were asked if they would like to participate. Each participant read and signed a statement indicating that they understood the study and freely agreed to participate. One student who was asked to participate in the study declined. All other students accepted.

The interviews with the fourteen main study participants were usually conducted in the student's home, though some took place in private offices and classrooms at the school sites. All but one parent was interviewed in their homes. Most best friend interviews were conducted in the students' homes. The fourteen primary student interviews focused around the student's educational history and plans. The goal was to gain an understanding of the general educational history of the student and her family, as well as a more in-depth understanding of the critical incidents in the participants' lives that placed them on the path to college attendance. While the interviews were semi-structured to allow for variation among school contexts, specific questions in the protocol elicited descriptions of caring relationships. These questions included (but were not limited to): when did you first start thinking about college; when do you first remember talking with someone about college; did anyone else influence your thinking about going to college.

Three schools were selected to participate in the study. Each represents a mixture of race and class contexts. All three schools are located in the same metropolitan area. Access to the schools was gained through professional and personal contacts. Enrolling 3500 students, Lincoln, a large, public comprehensive high school, resides in a low-income neighborhood that is predominantly Black and Hispanic. Wilson is also a large, public comprehensive high school. Located in an upper middle class suburban neighborhood, the student body of 2854 pupils is uniquely diverse, representing a large number of American ethnic groups and other nationalities. The student body is also diverse socio-economically; students come from the surrounding suburban neighborhoods and travel from the poor neighborhood in which Lincoln is located. Nestled in an exclusive, upper-class enclave, Hadley is a small, independent girls' college preparatory school. Enrolling 500 students, the young women who attended this school were from middle and upper-middle class families. The school is predominantly white.

Over the course of her research, the first author noticed the recurring theme of care in her discussions with the primary participants. In 1997, a post hoc analysis of the student interviews using Noddings' (1992) reciprocal construct of care was completed. Three themes emerged from the data set: adult involvement in decision-making, pushing, and cultural connection.

Table 1 provides an overview of the main themes of care and exemplar quotes that we found at the three schools in the study. Each participant in the study discussed a minimum of two themes reported in Table one. Generally, the adults in the students' lives were much older than the student. However, there are a few exceptions where the adults were young adults. Often these young adults were college students who were simply farther along in their education. Caring adults offered the support, encouragement, and direction that these students needed as they made their way through high school and on to college.

It is very important to remind the reader that this paper

Table 1

Sample of Illustrative Quotations by School

	Lincoln	Wilson	Hadley
Pushing	A lot of teachers kind of pushed me and made me work hard.	There has to be people in schools that will push people because there are some people that may not have that push in the home, but if they come to school and someone pushes them, then they'll pursue a further education.	{My brother and I} have always thought of college. It's never been an option. As far as our futures ever. When we were little it was always kind of pushed on us that we were gonne have like a career kind of future.
Adult Involvement in Decision-making	I was talking to one of my teachers and he said, "Debra, you got to have backup. If you don't get accepted you got to have some backup."	Well, {my cousin and my mentor}, they are really smart. They're into college or whatever and she's really smart. She's always coached me on what to do and what not to do and what I should do in regards to school.	I think she's a really good college counselor. I mean she's helped me a lot in my decision.
Cultural Connection	*Of a Black counselor:* She's like a friend. She's nice. She helps me.	I'm close to a lot of Black teachers there and they help me also. They're really like on me and stuff like that.	*Of a Director of Minority Affairs at a college she applied to:* And she's like I know it's gonna be hard and all this stuff and it feels good when somebody knows what it's gonna be like you know. You can get a first hand experience so it's not gonna be the same for everyone at least you've got something in common

does not present the findings of interviews explicitly conducted to investigate the influence of school context on students' perceptions of care. Unsolicited, each of the participants in the study discussed the importance of caring relationships in their lives. These women identified and talked about their mothers, teachers, and other adults who cared for them as they responded to questions regarding their academic success.

Findings: Caring in Context

All students in our study discussed care as a relationship with an adult. By far the most prevalent means by which a student felt cared for was when adults became involved in their decision-making. We called this *adult involvement in decision-making*. However, we also found contextual variations in how students perceived care. At Lincoln, the predominantly Black, inner city school, many students felt cared for when adults "pushed" them and at Hadley, the predominantly white, elite, private school, caring was marked by a cultural connection between the care-er and the student receiving the care. Our findings are divided into two sections. First, we review the most dominant theme found in the data – adult involvement in decision-making. In the next section of our findings, we contrast this overall theme of adult involvement in decision-making with contextual variations found at Lincoln and Hadley.

Adult Involvement in Decision-Making

Adults play critically important roles in the college decision-making process for youth. Prior research notes that students with parents who did not attend college are more likely to turn to their employers for help with this decision, whereas students with more highly educated parents are more likely to depend on their parents for advice and direction (Galotti & Mark, 1994; Litten, 1982). Research also reflects the role of teachers in this process. In Suskind's (1998) *A Hope in the Unseen*, Cedric's relationship with Mr. Taylor is indicative of this bond. Adult involvement in decision-making is especially important for young women. Research demonstrates that college-bound girls place greater emphasis on consultation with their parents and peers than young men (Galotti & Mark, 1994; Litten, 1982). Bateman and Kennedy (1999) suggest that mothers play a particularly important role.

Adult involvement in decision-making can be described as an exhibition of care. As college-bound high school students, these Black women encountered a range of choices that could have ramifications on their futures: which high school to attend, which courses to take, which colleges to consider. Students' appreciation for the advice and counsel offered by adults can be characterized as students' recognition of adults' exhibitions of care.

High school selection can be a pivotal moment in a student's life. One participant from Wilson described how her mother decided which high school she would attend. The student recounted: "My mother basically made that decision. She thought that as far as education that I would be better off going to [Wilson]." In considering her daughter's educational options, this student's mother made the decision that her daughter would attend a school that would provide her with the best education (given her options) as opposed to a school that was more conveniently located. Her mother's active involvement demonstrated care to the student; she knew that her mother had her best interest in mind when making this choice.

Students who attended Hadley talked at length about the adults who encouraged them to attend this elite private school. Here a student describes how one teacher suggested she apply to Hadley:

> My seventh grade teacher, when I was applying to schools, told me about Hadley, because her daughter had applied there in the past, but didn't get in. And so since I was one of the top students in the class and I was applying to like [other competitive schools] already, she said, 'Why don't you give Hadley a try?" although she was worried about the financial aid part because she knew it's a lot, it takes a lot of money to apply, but we figured it couldn't hurt to apply and apply for financial aid and see what happens.

Parents and other adults also helped participants with their course selections. When asked how she made choices concerning her courses, one Wilson participant explains her mother's involvement:

> I was going in to school with my mom, you know, thinking that the hardest courses are the ones you want to take, don't take cooking as an elective, take business organization. When I was in junior high, I used to take choir. I used to sing all the time. Then when I went to high school, my mom was like, 'You can't take choir anymore. You're going to college.'

The mother encouraged her daughter to take courses that she recognized as academically rigorous in order to help her daughter prepare for college. She also wanted her daughter to enroll in classes that would look impressive on her transcript. The daughter understood her mother's involvement as care because her mother wanted to prepare her for college.

Similarly, this student from Lincoln described how she consulted both her counselor and her mother regarding course selections. Although she was not interested in math, her mother insisted that she continue to take advanced math courses. When asked who she talks to when deciding what courses to take she responded:

> *Student*:[I] try to talk to my mother and my counselor. My mother – I talk to because, like, her

> thing is about AP classes – I really didn't want to take AP Calculus because I just really didn't want to take anymore math. [But] there was no other option.
> *EH*:If you hadn't takencalculus,you would not have takenmath?
> *Student*: Yeah. I talk to my mother for support. She says, 'Oh you can do this...'

Students from Hadley also called on the expertise from their parents in selecting courses. This student explains why she took calculus:

> My mom really encouraged me to take it [calculus] because she said when she took calculus at Stanford she said she had a really difficult time with it. So with me taking it she kind of influenced my decision and plus to help me with Physics.

Parents and adults were also important in helping the participants to develop a list of potential colleges. A Wilson student explained that she applied to Howard and Spelman because of her parents influence: "[They] wanted me to go to a historical Black college." Many Black parents are aware of the potential benefits that many young Black people receive from attending a historically Black college or university (HBCU). Research documents the significant role that HBCUs often play in students' lives (Fleming, 1985; Allen, 1992). The parents' involvement in the student's decision-making process surrounding college can be described as an act of care.

Some adults discouraged participants from considering HBCUs. One Wilson student narrates:

> I was talking to one of my teachers and I was expressing to him some of the colleges that I wanted to go to and when I brought up Howard, he was like, "I don't think you should go to Howard. I don't think you'd be good there." I said, "Why?" He goes, "Because they like hold their students' hands and show them what to do, I don't think you need that kind of guidance. I think you should just be let to go by yourself and find your own way. That's the way that you would be able to develop. I don't think you'll be good there." The kind of education they provide is ... because you know they always say, "Well, you know, we're all Black and we all help each other out." That sort of brotherhood/sisterhood kind of thing and, he said, "I don't think you need that." So that kind of stuck in my head, too. I don't know where I want to go.

This teacher appears to be sincerely interested in this student's well-being and development. He wants the best experience for her after high school. He shares his reservations concerning HBCUs not to confuse her, but to warn her against what he understands to be an educational experience that may stunt rather than further her growth. Although the student appears confused about her college choices after listening to this teacher, the fact that she took his reservations regarding HBCUs to heart speaks to the student's recognition of his sincerity and good intentions.

Other teachers played critical roles by providing a nurturing environment and preparing students to think of themselves as "college material." They assisted students with the identification of potential colleges to attend. This student from Lincoln talks about a popular and respected teacher with whom she described a close relationship:

> He's just like my role model. It's like so much of a personal relationship too that makes me want to work hard. It makes me believe in myself. He kind of boosts my self-esteem, in a way. It's like he just had his first child and it's like – it's been like a family – like, 'How's your baby?' and stuff like that. It feels a little closer. It's more of a personal connection.

Later in the interview, this student talks about how this same teacher assisted her with her college choice process:

> I ask him everything and he said, 'Oh yes. They're [a certain college] real good.' And I was, 'Oh, they probably won't want me if they're real good.' And he was like, 'Don't say that. Don't make the choice for them.'

At Hadley, parents, teachers and counselors were involved in the college choice process. After explaining that her college counselor helped her generate some of the colleges on her list this student noted the important influence of her father. She says: "Cornell and Syracuse for my dad. He's been talking about it ever since I was three years old. So I finally woke up and I was like, 'Let me at least see about these schools.' So, I visited and read about them and I was like, 'Hey, he wasn't lying. These are really good schools.'"

Even in their decisions to refrain from commenting on the student's choices, adults and parents communicated care to the participants. After being asked if she talked with her mom about the schools she wanted to attend, one student responded:

> [My mother] always asks what college... She didn't say, "I want you to go here, here, here." She said, "Well, I wish you could go here. Why don't you do this?" But, she really don't ... she lets me make my own decisions about college because she don't have to go to them. I have to. She trusts my judgement.

In this instance, the mother is involved in her daughter's college decision-making process. The mother does not tell her daughter where to apply. She has more of a hands-off approach, discussing the matter with her daughter, but

leaving the ultimate decision to her daughter. Her mother expresses an interest in her choices. She also rests assured in the knowledge that her mother trusts her own ability to make major decisions. A definitive demonstration of care is when a parent or another caring adult can let go and allow the student to grow (Noddings, 1984).

Participants in the study described specific acts of care as instances when their parents and other adults assisted them in making decisions that affected their immediate and future educational experiences. When a parent or another adult decided which high school the student would attend, suggested courses to take, or discussed colleges that the student should consider, care was observed.

Variations by School Context in Students' Perceptions of Care

While adult involvement in decision-making was clearly the dominant way in which students related feeling cared for by adults in their lives, we found that students' perceptions of care seemed to vary significantly among the schools. In this section we present data that shows how students from Hadley and Lincoln placed differing degrees of emphasis on specific aspects of care. Our findings suggest that educators need to be attentive to the school context within which care is expressed and received.

Pushing at Lincoln Preparatory High School

Participants from Lincoln more often identified pushing as demonstrating care than participants from Wilson and Hadley. When a participant referred to care as pushing, they were more likely to be a student at Lincoln- a low-income, predominantly Black, inner-city school.

Pushing can be defined as the possession of high expectations of a student, the insistence on high quality performance, and the refusal to accept anything less than the student's personal best. A number of behaviors encompass and exemplify pushing. For one student at Lincoln, being pushed meant not being "babied." She explains:

> [A] lot of my teachers haven't babied me and I have a tendency to want to be babied ... a lot of teachers kind of pushed me and made me work hard. Basically, helping me to find out what I want out of life more so as in educationally, but knowing what I want out of life it's helped me so that I'm prepared and I'm willing to work for what it is.

For this student, being babied is when a teacher does not make her work hard. Teachers who pushed her made her work hard, and consequently, helped her to find what she wanted out of life. The student recognizes and appreciates this. She understands that the hard work she is being made to do now is preparing her to meet the challenges of her future. The refusal to baby is recognized as a demonstration of care.

Most of the participants at Lincoln referred to the strict policy of certain teachers regarding late assignments. Many students mentioned the same science teacher who would not accept late homework and the same English teacher who would not accept late papers. One student describes a science teacher who pushes his students: "He doesn't baby us. When we have something due, it's due PERIOD. And he's sorry if you don't turn it in, that's too bad. The test; if you fail the test it just goes directly in. He'll just average it out. And that's it." This refusal to accept late assignments and poor performance on tests pushed students to be disciplined and to take responsibility for their actions. No longer were the students being babied with lax homework and test policies. This science teacher forced his students to adhere to high standards and expectations.

Firm policies and high expectations pushed students. In turn, this pushing is perceived as care. One student said this of the Advanced Placement teachers:

> Well, they try to prepare you. ... nice teachers, very intelligent teachers that are willing to work with the students. In that way, they prepare you because they're teaching you, like my English teacher. She'll give you an essay. It's due Monday. She don't sit there and keep begging for the essay. She's like, "If you don't turn it in you get an F. That's all it is to it." Some of the teachers prepare you for the fact that you're not going to be treated as babies anymore. They're not going to beg you to do anything, it's just, if you don't turn it in, that's on you.

Advanced Placement teachers- and more specifically in this instance, the English teacher- pushed students to prepare them for college. The English teacher sets deadlines that the students must meet. By enforcing this policy, the teacher fosters students' growth. Students learn to be disciplined and to take responsibility for themselves. Thus, pushing is an act of care.

The students perceived that this adherence to high standards, this "pushing," prepared them for college. The students generally interpreted pushing as a demonstration of care because the teacher believed in their academic potential and cared for their future well-being. However, pushing consists of more than strict adherence to rules and high expectations. A student notes: "Sometimes [the science teacher] will grade on a curve on the test, so that's kind of good because if everybody fails the test, he'll go back and try to ... help us out." In this example, we see the teacher acknowledging that a particular topic was difficult. Instead of penalizing the students and allowing their grades to stand, the teacher modified their grades and revisited the subject matter to "help us out." Here we see that pushing is a partnership between the teacher and the student, where the teacher assisted the student in meeting and exceeding established goals. More than being demanding, the teacher

possessed an understanding of where the students were academically so that he could help them get to where they need to be – both in the mastery of the subject, as well as in their entrance to college. Often referring to his class as an "in-between class," students understood that the teacher believed his class to be a bridge between high school level work and college level work. His commitment to preparing students for college was described as care by several participants.

For participants at Lincoln, pushing is a specific attribute of care. One student said it best. When asked if Lincoln does a good job preparing students for college, this participant responded: "Yeah, I believe, I don't think so much it's Lincoln. I think it's the kids' — the drive that's in the kids along with the help of the teachers, the ones that really care and really want to help you out, that prepares them for college." This participant recognized the student's role in terms of college preparation, and she acknowledged the importance of caring teachers. At Lincoln, caring teachers helped students by pushing them, by not babying them, by requiring them to meet deadlines, by assisting them in the attainment of high expectations, and by shaping students into disciplined and responsible young adults.

Cultural Connection at Hadley

The theme of cultural connection describes participants' expressed belief that Black adults and mentors offer greater care and concern for them. This concern is rooted in their shared cultural bond and the adult's own lived experiences as a person of African descent. Building on Irvine's (1990) concept of cultural synchronicity, this section of the paper describes the participants' caring relationships (or longing for a relationship) with other Black people.

When compared with the other schools in the study, participants from Hadley more often identified cultural connection as demonstrating care than participants from Lincoln and Wilson. It is important to note that at the time of this study there were no Black faculty or staff at Hadley- an elite, predominantly white, girls independent school. Thus, the importance of a cultural connection offering insight to students' perceptions of a caring relationship was most evident due to its absence.

Particularly revealing were participants' responses to the question: Do you think that your relationship with your counselor would be different if he/she was Black? One participant at Hadley suspected she would have a heightened sense of comfort with a Black counselor:

> When I'm around a lot more Black people I feel a good sense of community and that I can talk about anything I want. So if [the counselor] was Black I really would feel a lot more comfortable and I probably, like, be more closer to her than I would be with [the current counselor].

This student believed that she would feel more comfortable with a Black counselor because of past experiences that she had when she was around "a lot more Black people." Based on a previously encountered sense of community, this student contended that she would feel more comfortable and enjoy a closer relationship with a Black counselor. This participant believed that a Black counselor would care for her.

Another participant elaborated: "Well probably [my relationship with a Black counselor would be different] because... I think maybe he might be more inclined to be concerned about how I'm doing as opposed to the other students who aren't Black in the class." Although this student seems uncertain, her message is clear. She felt that a Black counselor would care for her more than her current counselor because he/she would be concerned with how she was doing. This concern for her personal well-being would be perceived as care by this student.

A caring relationship with a person of African descent would mean connecting with an individual who understood the student's experiences as a Black student in a white school. As one participant explained: "A [Black counselor would] be a been-there, done-that person you know so she would really know. This woman [the participant's current white counselor] has no idea what it would be like for me in college." Later in the same interview, the participant discussed a phone conversation with a director of minority affairs at one of the colleges she was considering:

> ... talking to a woman [of color], she said, "You did a good move taking calculus. Because I've seen too many young Black women, young Black people, applying to our school who just cop out towards the end and it's better for [admissions counselors] to see that you are trying." It feels good when somebody knows what it's gonna be like you know. It's not gonna be the same for everyone [but] at least you've got something in common.

For this participant, and other students at Hadley, connecting with a Black adult who not only knew and understood their current experiences but who could also help them to prepare for their experiences as young Black women in college was very important. The students believed that the Black adults who they yearned to connect with could offer affirmation. By sharing cultural backgrounds, cultural heritages, and personal experiences of being Black in a predominantly white academic setting, the students felt that care could be better facilitated with a Black counselor.

Unable to connect with Black adults at Hadley, many of the students connected with each other. One student discussed supporting another Black student after a classroom incident with racist implications:

> Oh well, I don't know like one of my friends ... afterwards was crying you know and like I just, I'm not even like I guess I get emotional about those types of things, but I'm not the type of person to

> like start crying in front of all these people you know. So after everyone left I was like giving her a hug and stuff just cause like I guess I understood how she felt or something.

In the absence of Black adults to comfort and validate their experiences as Black students in a predominantly white school, many of these students turned to one another for support. By forming a connection with each other based on their cultural heritage, these students cared for one another.

While all of the categories were found in all of the schools, the frequency with which the attributes occurred varied. Adult involvement in decision-making was important to all of the participants. However, we uncovered important contextual variations as well. For participants at the low-income, predominantly Black school, pushing appeared to be more instrumental in terms of demonstrating care to students. Cultural connection was found to be the most important attribute of care for students at the predominantly white school. In summary, the major finding of this research is that dependent upon the school context, a different or specific attribute of care is found to be more important to demonstrating care to students.

Discussion

The lived experiences of the students in this study illustrate the significance of care in schools. Specifically, we highlight the ways in which perceptions of care varied by school context. Caring relationships with parents, teachers, and peers were described as pivotal elements of participants' high school experiences and college decision-making processes. Reflections on their high school experiences yielded memories of those individuals, whose demonstrated care nurtured their academic achievements and college aspirations. These students provide evidence in support of Nodding's (1992) call for an ethic of care to imbue educators' approaches to the profession.

This study's findings resonate with other studies on care and education. Several scholars identify help as an important component of care (Bosworth, 1995; Noblit, Rogers, & McCadden, 1995). Bosworth's (1995) study categorizes guidance as a specific form of help. Our finding of adult involvement in decision-making corroborates and extends our understanding of this form of caring for students. The help that parents, teachers and other adults provided students during this crossroads period of college decision-making can be characterized as acts of care.

Black students' recognition of pushing as an act of care is also well documented (Delpit, 1988; Irvine, 1999). By disciplining students and demanding the best, teachers communicate their high expectations, their faith in students to excel, and their care for students (Irvine, 1999). This study substantiates these earlier works and considers how pushing as an act of care contributes to the increased performance and self-esteem of students.

Several scholars note the significance of Black teachers for Black students (Delpit, 1988; Foster, 1997; Wilder, 2000). Wilder's (2000) qualitative study records Black students' insistence on the necessity of Black teachers in the school system. The importance of cultural connections between students and teachers resounds in this study as well. While a shared cultural heritage is not necessary for the development of a caring relationship, such relationships with adults who share one's cultural heritage are important to the academic achievements and aspirations of Black students (Wilder, 2000). Tatum (1997) explains the significance of Black students' relationships with one another and Black adults as a positive coping strategy given students' racial identity development and the stressors associated with predominantly white educational settings. The ability to respond to Black students' experiences of racism and explorations of racial identity in meaningful and supportive ways is yet another critical component of care.

Adolescence is a period of growth and development. Our findings reveal the importance of caring relationships to the academic achievement and college decision-making processes of a group of Black adolescent girls. We surmise that caring relationships are significant in other areas of adolescent development. Presently, most models used to ascertain racial identity development (Cross, 1991; Sellers, Shelton, Rowley, and Chavous, 1998) and ego identity development (Marcia, 1966, 1980) rely on quantitative instruments and designs. To better understand the process of identity development, we suggest that future research employ qualitative methods to examine the role and influence of caring relationships in the identity development of adolescents.

Current debates among educators weigh the adequacy of the ethics of care, critique, and justice to serve the best interests of students (Callan, 1992; Shapiro & Stefkovich, 2001; Starratt, 1994). Callan suggests a balanced blend of the ethics of care and justice. Starratt argues that each ethic compliments the other to provide a comprehensive ethical framework for educators. Shapiro and Stefkovich further develop and expand Starratt's proposed framework. By recognizing the personal ethics of the educator as well as the diversity of students, Shapiro and Stefkovich advance a fluid model that allows for the complexities that shape, inform and influence the moral deliberations of contemporary educators. The research presented here seeks to contribute to this debate by including the voices of students and grounding these ethical considerations in everyday encounters.

In addition to including students in the chorus of voices who advocate care in schools, this study examines the influential role of school context in the development of relationships with students and other ethical considerations. All aspects of care - adult involvement in decision-making, pushing, and cultural connection – are necessary in order to adequately care for students. However, which aspect is most important to the student receiving the care is contextually mediated. This research suggests that the student's perceptions of care will be affected by the social

location in which she lives and learns. Although these findings are not conclusive, one could speculate that school context plays an important role in mediating students' perceptions of care.

References

Allen, W. (1992). The color of success: African American college student outcomes at predominantly White and historically Black public colleges and universities. *Harvard Educational Review, 62*(1), 26 - 44.

Bateman, M. & Kennedy, E. (1999). Expanding the applicant pool: Exploring the college decision-making of students from single-parent families. *College and University, 75*(1), 2 – 10.

Bosworth, K. (1995). Caring for others and being cared for: Students talk about caring school. *Phi Delta Kappan, 76*, 686 – 693.

Callan, E. (1992). Finding a common voice. *Educational Theory*, 42, 429 – 441.

Cross, W. (1978). The Thomas and Cross models of psychological nigrescence: A review. *Journal of Black Psychology, 5*, 13 – 31.

Delpit, L. (1988). The silenced dialogue: power and pedagogy in educating other people's children. *Harvard Educational Review, 58*, 280 - 298.

Dempsey, V. & Noblit, G. (1993). The demise of caring in an African-American community: One consequence of school desegregation. *The Urban Review, 25*(1), 47 - 61.

Fleming, J. (1985). *Blacks in college: A comparative study of students in Black and White Institutions.* San Francisco: Jossey-Bass Publishers.

Foster, M. (1997). *Black teachers on teaching.* New York: The New Press.

Galotti, K. & Mark, M. (1994). How do high school students structure an important life decision? A short-term longitudinal study of the college decision-making process. *Research in Higher Education, 35*, 589 – 607.

Glazer, J. & Wughalter, E. (1991). Mentor in education: Attracting minority students to teaching careers. *Mentoring International, 5*(1-2), 15 – 20.

Horvat, E. (1996). *African American students and college choice decision-making in social context.* Unpublished doctoral dissertation, University of California, Los Angeles.

Irvine, J. (1990). *Black students and school failure: Policies, practices, and prescriptions.* New York: Greenwood Press.

Irvine, J. (1999). The education of children whose nightmares come both day and night. *Journal of Negro Education, 68*, 244 – 253.

Ladson-Billings, G. (1994). *Dreamkeepers.* San Francisco: Jossey-Bass Publishers.

Lareau, A. (1997). *Home Advantage: Social Class and Parental Intervention in Elementary Education.* London: The Falmer Press.

Levine, A. (1994). African Americans and the three-year college degree. *Journal of Blacks in Higher Education, 3*, 82.

Litten, L. (1982). *Ivy bound: High ability students and college choice.* New York: College Board.

Marcia, J. (1966). Development and validation of ego status. *Journal of Personality and Social Psychology. 3*, 551 – 558.

Marcia, J. (1980). Identity in adolescence. In J Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159 – 187). New York: Wiley.

Mayerhoff, M. (1971). *On caring.* New York: Harper and Row Publishers.

Noblit, G., Rogers, D., & McCadden, B. (1995) In the meantime: The possibilities of caring. *Phi Delta Kappan, 76*, 680 – 685.

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education.* Berkeley: University of California Press.

Noddings, N. (1992). *The challenge to care in schools: An alternative approach to education.* New York: Teacher's College Press.

Polite, V. (1993, April). If only we knew then what we know now: The school-to-work and school-to-college transitions of African-American males in suburbia. Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.

Rhodes, J. and Davis, A. (1996). Supportive ties between nonparent adults and urban adolescent girls. In B. Leadbeater & N. Way (Eds.), *Urban Girls: Resisting Stereotypes, Creating Identities* (pp. 213 – 225). New York: New York University Press.

Sellers, M., Shelton, M., Rowley, S., & Chavous, T. (1998). Multidimensional Model of Racial Identity: A reconceptualization of African American racial identity. *Personality and Social Psychology Review, 2*, 18 – 39.

Sevier, R. (1993). Recruiting African American undergraduates: A national survey of the factors that affect institutional choice. *College and University, 68*(1), 48 – 52.

Shapiro, J. & Stefkovich, J. (2001). *Ethical leadership and decision making in education.* Mawah, NJ: Lawrence Earlbaum Associates Publishers.

Stage, F. and Hossler, D. (1989). Differences in family influences on college attendance plans for male and female ninth graders. *Research in Higher Education, 30*, 301-315.

Starratt, R. (1994). *Building an ethical school: A practical response to the moral crisis in schools.* Washington, D.C.: Falmer Press.

Suskind, R. (1998). *A hope in the unseen: An American odyssey from the inner city to the ivy league.* New York: Broadway Books.

Tatum, B. (1997). *"Why are all the Black kids sitting together in the cafeteria?" and other conversations about race.* New York: BasicBooks.

Wells, A. and Crain, R. (1997). *Stepping over the color line:*

African-American students in white suburban schools. New Haven: Yale University Press.

Wilder, M. (2000). Increasing African American teachers' presence in American schools: Voices of students who care. *Urban Education, 35*, 205 – 220.

Erin McNamara Horvat is an Assistant Professor in and Program Coordinator of the Urban Education Program at Temple University.

Kristine Lewis is a doctoral candidate in the Urban Education Program at Temple University.

Author Notes

All names of students and locations are pseudonyms.

Regretfully, a question on caring specifically was not included in the interview protocol.

One student interview transcript was unavailable.

The authors gratefully acknowledge the helpful comments provided on previous versions of this manuscript by Joan Poliner Shapiro and Steven Jay Gross. All errors and omissions, however, remain our own.

Reflective Racial Tolerance and its Development in Children, Adolescents and Young Adults: Age Related Difference and Context Effects

Rivka T. Witenberg
The University of Melbourne, Australia

Racial tolerance has been under-researched and is largely unexplored from a psycho-developmental perspective. The starting point for defining and exploring tolerance is the assumption that tolerance is not simply the absence of prejudice but that it is a separate and complex construct which is worth serious consideration. This paper will discuss definitions around tolerance and report on a study about racial tolerance using a developmental model involving one hundred and eighty four (184) students aged 11-12, 15-16 and 18-22. The students were asked to make judgments about stories relevant to the Australian context about people from Aboriginal, Asian, and English backgrounds. Analysis revealed age-related differences with older students becoming less tolerant, contradicting previous findings about political and belief discrepancy tolerance. Older students tended to accept the right to express openly intolerant beliefs, both for themselves and others on the basis of democratic rights and freedom of speech. However, racial tolerance was found not to be global construct endorsing previous findings about political and belief discrepancy tolerance. To whom and under what circumstances students were asked to extend their tolerance and what behaviours they were asked to tolerate affected their judgments.

Racial tolerance has been under-researched and largely unexplored, particularly from a psycho-developmental perspective. This contrasts starkly with the large body of research into prejudice (Aboud, 1993; Aboud & Doyle, 1996; Doyle & Aboud, 1995; Bigler & Liben, 1993). Much of the research into prejudice assumes that absence of prejudice reflects tolerance, a claim that has been challenged in recent times. Conceptually prejudice and tolerance are different processes that are often investigated as if they were the opposite of each other, confounding much of the research into tolerance (Vogt, 1997; Robinson, Witenberg & Sanson, 2001; Witenberg, 2000a, b). Thus the starting point for the current research is the assumption that tolerance is not simply the absence of prejudice but that it is a separate construct which is worth serious consideration from a psycho-developmental perspective. Given the significance society places on tolerance and acceptance (Thomas, 1998), it is surprising how little we know about racial tolerance in its own right (Vogt, 1997; Witenberg, 2000a, b).

Tolerance is only necessary when difference or diversity is present because it is only when confronting diversity that our acceptance of others is truly tested (Robinson, et al., 2001; Vogt, 1997; Witenberg, 2000a). As an outcome of changing geo-political realities, we live in a complex world that is racially and culturally diverse but not necessarily harmonious. Multinational migrations of either ethnic workers or refugees, whether for economic or political reasons, is becoming more and more prevalent (Iram, 2001) resulting in a world were monocultural societies are virtually extinct today (Walzer, 1997). Globally, cultural and ethnic diversity has increased the need for tolerance of racial differences. Unfortunately, our knowledge and ability to reduce prejudice and discrimination has proved to be a very difficult task (Aboud, 1993; Aboud & Doyle, 1996). Therefore, understanding tolerance and acceptance as applied to racial and cultural characteristics is an important research goal both theoretically and practically. Such insight may help to shape the debate and practice on social policies related to productive diversity in a multicultural environment and provide a basis for the design of community and school-based education programs (Thomas, 1998). Evaluations of education and/or intervention programs aimed at reducing prejudice have yielded mixed results. Promoting tolerance maybe an alternative route to a more harmonious society. However, where tolerance is concerned, even definitional issues are complex, (some would suggest contentious), and need further consideration. Tolerance is an ambiguous and complex concept open to several interpretations ranging from forbearance to full acceptance.

The most commonly accepted definition of tolerance

emphasises forbearance, endurance or "putting up with" without interference. Tolerance defined as endurance involves a recognition that other people have different opinions and practices but it does not entail any form of acceptance or support of the difference. This definition of tolerance implies that a person can be tolerant and prejudiced simultaneously. One can endure and refrain from acting intolerantly but remain biased, disapproving or judgemental. For example, a schoolteacher or a lecturer may hold prejudiced beliefs towards students from a different racial or cultural group but may act in a tolerant way in order to maintain his or her position. This possibility is rarely acknowledged in the literature, which tends "to assume that tolerance and prejudice are mutually exclusive of each other" (Robinson et al., 2001, p 74). Embedded in the idea of "tolerance as endurance" is also the notion that what is endured is less valuable or even socially or morally wrong (Witenberg, 2000a, b).

Secondly, tolerance can be defined as the absence of prejudice of what is disliked, resented and possibly even abhorred towards those "whose practices, race, religion, nationality, etc., that differ from one's own" (Robinson et al., 2001 p. 74). This definition focuses on absence of prejudice and goes beyond simply enduring or refraining from interfering. This sort of conceptualisation of tolerance has been adopted by much of the research into prejudice, particularly with children (Robinson et al., 2001). However, the absence of prejudice both in children and adults is just that. It cannot be assumed that the lack of one does necessarily entail the presence of the other (Witenberg, 2000b).

"A third view of tolerance is that it involves a conscious rejection of prejudiced attitudes, beliefs and responses. That is, one's own negative stereotypes are recognised, judged against experiential knowledge or value systems, and rejected" (Robinson et al., 2001, p. 74). Recognising and rejecting prejudicial views moves a person from simply being "a narrow-minded bigot who shows restraint" to a reflective individual (Burwood & Wyeth, 1998; p. 469). This definition involves an active, conscious, reflective agent and for these reasons has been adopted for the purpose of this research. Finally, perhaps the most ideal way to think of tolerance hinges on full acceptance of others. This involves a conscious rejection of biased beliefs and behaviour and the valuing of others irrespective of their colour or creed. Acceptance of differences and diversity also entails "enthusiastic endorsement of difference" (Walzer, 1997, p.11). However, there is little research that has specifically addressed such a definition of tolerance (Robinson et al., 2001).

Two bodies of research, political and belief discrepancy tolerance, have enhanced our understanding of tolerance more generally (Avery, 1988; Enright, Lapsley Franklin & Streuck, 1984; Owen & Dennis, 1987). Whilst political scientists and sociologists using large scale surveys have examined political tolerance mostly in adolescents and adults over the last thirty years, only a few studies have investigated tolerance from a psychological or developmental perspective. Where tolerance research has been investigated from a psychological perspective a cognitive developmental model has been utilised at times (eg. Enright & Lapsely, 1981; Wainryb, Shaw & Maianu, 1998). Their findings highlight age-related differences and the contextual nature of tolerance.

In a series of studies using dilemmas, Enright and Lapsely (1981) asked children from grade 1 through to college students to judge people with dissenting beliefs on a range of social, moral and political issues. They found that tolerance involved an age-related progression from less to more tolerance towards those with dissenting beliefs. However, a more recent study by Sigelman and Toebben (1992) who examined both belief-discrepancy and political tolerance with grades 2, 5 and 8 found that age related differences were mediated by context.

Wainryb et al. (1998) also found that tolerance to dissenting beliefs followed an age-related progression which was mediated by contextual factors. They examined what behaviours and under what circumstances students from grade 1 through to college undergraduates would extend their tolerance. Irrespective of age, holding dissenting beliefs was the most tolerated behaviour, while engaging in acts based on these beliefs was least tolerated. Clearly, what behaviours participants were asked to make judgements about influenced their level of tolerance.

Several studies examining political tolerance have also shown that tolerance and intolerance coexist in individuals of all ages including children and adolescents. McColsky and Brill (1983), found that a very high percentage of their participants defended "freedom of speech" as a democratic principle, but such rights would only be accorded to "acceptable" individuals or groups. It appears that people are selective about whom and what they will tolerate and under what circumstances they are prepared to be tolerant. Research has also found that support for political tolerance increased significantly with age (Sigelman & Toebben, 1992). Whilst few age related differences were observed in the primary school years, important developmental changes occur during adolescence and early adulthood which is a vital period in the emergence of political tolerance (Avery, 1989; Zellman & Sears, 1971).

Whilst age-related differences were evident, the progression differed for the two domains of belief discrepancy and political tolerance. The findings relating to belief discrepancy suggest a gradual progression from childhood to adolescence ranging from less to more tolerance towards those with dissenting beliefs. In political tolerance the salient time is adolescence and early adulthood. However, whilst age related differences were evident in both domains, when context was considered, no single global construct of tolerance emerged. What we do not know is whether racial tolerance follows a similar pattern of age related differences and whether it too is multi-dimensional and context dependent. Thus, the main purpose of the investigation was to examine the possibility of a unique

pattern of age related differences in racial tolerance. The study focused on judgements children (11-12 years old), young adolescents (15-16 years old) and young adults (18-22 years old) made about three different contexts (context of tolerance) and about a broad set of behaviours involving beliefs, speech and acts (dimension of tolerance).

Method

Participants

There were 184 participants in the study, 58 were between 11-12 years (in either grades 6 or 7), 55 were between ages 15-16 years (in grade 10) and 71 were between ages18-22 years. The two younger groups attended schools either in the Melbourne metropolitan area or a rural township. None of the teaching institutions were situated in areas of extreme socioeconomic deprivation. About 44% of the participants had parents born in Australia, 20% originated from Northern Europe or North America, 16% from Southern and Eastern Europe, 2% from the Middle East, and 18% originated from either South East Asia or the Indian subcontinent. All the participants were fluent in English.

Procedures and Measures

The stimuli to assess tolerance consisted of three short dilemma-like stories. Each dealt with an event depicting a form of tolerance relevant to the Australian context and based on real life incidents. These were gathered from reports in newspapers, individual experiences and from official sources such as the Human Rights and Equal Opportunity Commission's (1991) study into racial prejudice in the Australian context. The order of the presentation of the stories was counterbalanced. They were presented with competing considerations (see below for more detailed account of the stories) and students were assured that their specific answers could not be identified. Students were debriefed and informed that the researchers did not endorse the views presented in the stories.

Story 1 concerned a person from an Asian background whom a shopkeeper wants to serve last because Asians do not belong in Australia. Story 2 involved an employer who will not hire a young English person because English people are lazy. The final story was about a person from an Aboriginal background whom a teacher at school wants to give lower marks to because the person was black. A pilot study found that the use of the terms Asian, English and Aboriginal posed no conceptual or ethical problems to the participants. After responding to the stories, participants were asked to indicate whether they had any problems understanding the three terms or found them objectionable.

To explore dimensions of racial tolerance, each story was presented three times and each presentation dealt with a different behavioural dimension; holding prejudicial beliefs, expressing these beliefs, and acting on these beliefs. In the first instance the protagonist holds prejudicial beliefs, in the second the protagonist tells others about these beliefs, and in the final presentation the protagonist acts on these beliefs (Wainryb et al., 1998).

The issues in the stories could be viewed from either a tolerant or intolerant perspective. Participants were asked to make a judgement in writing about the events in the stories. They were then asked to explain their judgement and to express their own ideas about the event. Finally, participants were asked to justify their judgements.

Assessment and reliability

Responses were assessed using a four-point rating scale. In the first instance a response was categorised as either tolerant or intolerant. Level of tolerance for both stories and dimensions depends on the number of responses that were coded as tolerant or intolerant. (Scoring directions are available from the author.)

Responses were coded as representing "reflective tolerance" when students' rejected prejudicial beliefs, speech and acts and supported tolerance by appealing to one or more of the following values; fairness/equality, empathy or reasonableness (Witenberg & McDowall, 2001). Fundamental beliefs about equality and fairness were reflected in such responses as "It is wrong", "It's not fair" "We are all equal" and "we are all alike". Statements such as "I would explain that we are all equal and it shouldn't matter where you come from or what you are. We should all be treated fairly and equally." Perspective taking was reflected in responses such as "I would ask you, how would you feel if you were an Asian person and would always be served last." Reasoned thinking was reflected in such responses as "Giving Aboriginal children lower marks is stupid. How could you possibly think that way?" Inter-rater reliability for the classification of these beliefs was established through agreement with a second rater who coded 25% of the responses from the total set. There was an 80% inter-coder agreement between the two raters.

Results

The effect of cultural diversity on Tolerance

Due to the cultural diversity of the Australian society, it was deemed important to assess whether cultural origin as defined by the parents' place of birth was a significant variable in determining tolerance judgements. No significant relationships between the cultural origin of parents and students' tolerance judgements were found.

The effects of age, gender and context on tolerance

In order to examine the relationship between age, gender and context on tolerance judgements both multivariate and univariate analyses were computed. To assess the effect of context on tolerance level, a within subjects analysis of

Table 1

Mean Scores for Level of Tolerance for Context by Age and Gender

				Gender			
				Female		Male	
Context	Age	M_a	*SD*	M_b	*SD*	M_c	*SD*
Asian	11-12	3.79	0.61	3.76_d	0.55	3.83_e	0.71
	15-16	3.55	0.83	3.74_f	0.51	3.25_g	1.15
	18-22	3.41	0.61	3.42_h	0.59	3.38_i	0.64
	Total	3.56	0.70	3.63	0.58	3.47	0.86
English	11-12	3.71	0.54	3.64_d	0.59	3.83_e	0.38
	15-16	3.41	0.80	3.64_f	0.55	3.05_g	1.00
	18-22	3.24	0.75	3.31_h	0.81	3.13_i	0.61
Aboriginal	11-12	3.92	0.30	3.94_d	0.23	3.89_e	0.32
	15-16	3.80	0.40	3.77_f	0.42	3.85_g	0.37
	18-22	3.38	0.55	3.38_h	0.58	3.36_g	0.49
	Total	3.67	0.49	3.67	0.51	3.68	0.47

[a]N=[b]n=107, [c]n=62, [d]n=34, [e]n=18, [f]n=31, [g]n=20, [h]n=42, [i]n=24

variance was conducted with the three stories entered as repeated measures within subjects variable while age and gender were entered as between-subjects variables. Table 1 shows mean tolerance level and standard deviation for stories by age and gender.

The analysis revealed a significant interaction effect for context x age x gender, $F(4,326) = 3.30, p < .01$, partial $\eta^2 = .040$, an interaction effect for context x age, $F(4,326) = 2.52, p < .05$, partial $\eta^2 = .030$, and a significant effect for context, $F(2,326) = 12.70, p < .001$, partial $\eta^2 = .072$. With respect to the main effect, students were most tolerant when the subject of the story was a person from an Aboriginal background in the school situation and least tolerant when the person was from an English background in the work situation. However, the interaction effect showed that context related responses were mediated by both age and gender. This was in the direction of females between 15 and 16 years of age expressing more tolerance when the subject of the story was an English person in the work situation or an Asian person in the shop situation than did males of that same age. Between-subjects analysis indicated a significant effect of age on level of tolerance, $F(2,163) = 13.62, p < .001$, $\eta^2 = .14$. Contrary to expectation, a decrease in tolerance level for all three stories with age was evident as indicated in Figure 1. Tukey (HSD) post hoc test confirmed a significant difference between the three age groups.

The effects of age and gender on individual stories

Since the mutivariate analysis was significant, the effect of each individual variable was explored. The univariate analysis revealed a significant interaction of age x gender on the Asian story, $F(2,175) = 3.69$, $p = .05$, $\eta^2 = .030$, and significant differences between the three age groups on the

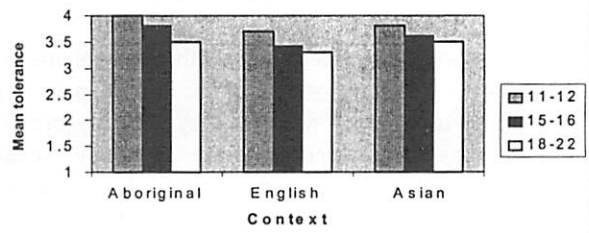

Figure 1: Mean level of tolerance for context by age

same story, $F(2,175) = 5.26, p < .01$, $\eta^2 = .057$. The youngest males were more tolerant than the females but the reverse was true for the two older groups where females tended to be more tolerant. The least tolerant were 15-16 years old males. Pairwise comparisons show significant differences between the three age groups.

The univariate analysis revealed a significant interaction of age x gender on the English story, $F(2,173) = 3.63, p < .05$, $\eta^2 = .040$, and significant differences between the three age groups for the same story, $F(2,173) = 7.96, p = .001$, $\eta^2 = .084$. Examination of the means again shows a general decrease in tolerance with age. However, 11-12 years old males were by far the most tolerant while the least tolerant were 15-16 years old males.

The univariate analysis revealed a significant difference between the three age groups on the Aboriginal story, $F(2,179) = 18.70, p = .001$, $\eta^2 = .173$, with a decrease in tolerance with age. Pairwise comparisons show a significant difference between the 11-12 years olds and 18-22 years old and between 15-16 years old and the 18-22 years old but no difference was found between 11-12 years old and 15-16 years old subjects. The univariate analysis confirmed the decrease of expression of tolerance with age.

The effects of age, gender on dimensions of tolerance

The effect of the three dimensions of tolerance was also examined using a within subjects analysis of variance with the three dimensions of belief, speech and act as repeated measures. Age and gender were entered as between-subjects variables. Table 2 presents mean tolerance level and standard deviations for dimension by age and gender.

Using the Greenhouse-Geisser ε adjustment, the analysis revealed a significant effect of dimensions, $F(1.43,232.83) = 61.07, p < .001$, partial $\eta^2 = .27$. Overall, 3.5% of the students accepted acting with intolerance and only 3.7% accepted holding prejudicial beliefs. In contrast, 30.1% of students (mostly 18-22 years old) believed it was acceptable to express discriminatory beliefs openly and freely. Mean level of tolerance for dimensions by age are shown in

Table 2

Mean Scores for Level of Tolerance for Dimensions by Age and Gender

				Gender			
				Female		Male	
Dimension	Age	M_a	*SD*	M_b	*SD*	M_c	*SD*
Belief	11-12	3.84	0.41	3.82_d	0.46	3.89_e	0.32
	15-16	3.67	0.59	3.81_f	0.47	3.45_g	0.69
	18-22	3.76	0.58	3.74_h	0.63	3.79_i	0.51
	Total	3.76	0.54	3.79	0.53	3.71	0.55
Speech	11-12	3.67	0.73	3.62_d	0.82	3.78_e	0.55
	15-16	3.33	1.01	3.39_f	0.99	3.25_g	1.07
	18-22	2.35	1.23	2.50_h	1.23	2.08_i	1.21
	Total	3.05	1.18	3.11	1.15	2.95	1.23
Act	11-12	3.90	0.41	3.91_d	0.38	3.88_e	0.47
	15-16	3.78	0.58	3.97_f	0.18	3.50_g	0.83
	18-22	3.92	0.27	3.88_h	0.33	4.00_i	0.00
	Total	3.88	0.43	3.92	0.31	3.81	0.57

[a]N=169, [b]n=107, [d]n=34, [e]n=18, [f]n=31, [g]n=20, [h]n=42, [i]n=24

Figure 2.

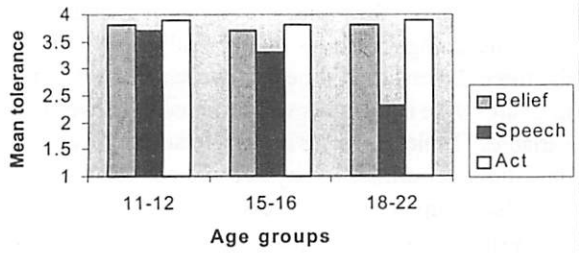

Figure 2: Mean level of tolerance for dimension by age

In addition, a significant interaction effect was found with dimensions x age, $F(2.86,232.83) = 14.70, p < .001$, partial $\eta^2 = .23$. Overall, students were least accepting of intolerant acts (followed by intolerant beliefs) and most accepting of intolerant speech, but this pattern was mediated by age. Most notable is the decrease of level of tolerance for speech. Between subjects analysis also showed a significant effect of age $F(2,163) = 13.93, p < .001$, $\eta^2 = .15$. There was a tendency for the youngest students to reject holding prejudicial beliefs more often than the other two age groups, thus making them the most tolerant group. The reverse was true for acting discriminatory where 18-22 years olds were the most tolerant. The least tolerant group was the 15-16 years old group. The most tolerant for speech were the 11-12 years old group. Least tolerant were the 18-22 year old students who endorsed expressing intolerance openly.

Discussion

The findings of this study offer some insight into the development of racial tolerance, its context dependent nature and to whom and under what circumstances young Australians are willing to extend their tolerance. Students expressed a very high level of tolerance. However, there were notable age related shifts for context and dimensions. There was a tendency for the 11-12 years old to strongly reject intolerance in all contexts and behavior and support tolerance. On the other hand, when intolerance was supported it was more prevalent in the responses of 15-16 years old students and peaked in the responses of 18-24 years old. The youngest students tended to more often make blanket judgements supporting tolerance by appealing to fairness, empathy or reasonableness. On the basis of individual and democratic rights, the older students endorsed intolerance more often than the other two age groups. This shift was also evident for each of the individual stories. Whilst age related shifts emerged, important contextual influences were also clearly evident supporting the findings of the research about political and belief-discrepancy tolerance. Contrary to expectations, greater tolerance was extended when the subject of the story was a person from an Aboriginal background in the school situation, while the subject of the English story was accorded the lowest level of tolerance in the work context.

Such an unexpected result, that young people are most tolerant to an indigenous person in the school situation and least tolerant to a person from an English background in the work situation is puzzling and raises the question about the relationship between person-situation interaction. Despite agreement that tolerance is not a global construct, there has been no extensive investigation into the manner by which tolerance judgements are affected by contextual factors. We do not yet know whether individuals react to the person, the situation or the interaction of a person in a situation when making tolerant or intolerant judgments. Currently a study is under way which is designed to separate person-situation relationships (Witenberg & Guadagno, 2001). This research seeks to investigate which component of the stimulus material influences individuals' judgements and may help to unpack the theoretical underpinning of whether the person or the situation variables influence responses to racial tolerance. Preliminary findings are showing that the participants in all three groups were more tolerant to a person irrespective of whether they are Aboriginal, Asian or English in the school situation and there is also a tendency to be least tolerant in the work situation.

However, the process is more complex as the participants in all three groups tended to be more tolerant in all three situations (school, work and shop) towards the person from the Aboriginal background than they were towards the person from either an English or Asian background.

The finding that tolerance was more often extended towards the Aboriginal person in the story contradicts previous findings in Australia where research has found that 5 to 12 years olds expressed less positive attitudes towards indigenous people in comparison to people from Asian or English backgrounds (Black-Gutman and Hickson, 1996). The current findings could be explained by the contemporary focus on reconciliation which has dominated the social and cultural debates in contemporary Australian society. This vigorous debate coupled with media and education campaigns may have affected the way young people view indigenous Australians today. Similarly, a lower level of tolerance towards people from Asian background could reflect the negative images expressed by vocal anti-immigration, anti-Asian rightwing parties.

What behaviours the students were asked to tolerate affected their responses. In all cases, students rejected holding prejudicial beliefs and acting on them considerably more often that speaking about them. The major constraint to tolerance was not prejudice towards others but beliefs in freedom of speech as a democratic right. Many of the students in both age groups assumed that it was acceptable to openly express prejudicial beliefs "because we had free

speech in this country" or "everybody has the right to have their own opinion and talk about them." These beliefs may again reflect recent, vigorous political debate in Australia about "political correctness" and civic education.

The topic of racial tolerance is one of considerable significance both from a theoretical and a practical perspective. The value of the current research is its focus on the positive aspects of social perceptions and behaviours in contrast to the large body of research into prejudice. Undoubtedly the ultimate aim of both research into prejudice and racial tolerance is to eliminate discrimination. However, in real terms the two perspectives may involve separate paths which cannot be explored as if there was only one. Of course we may follow both paths concurrently.

References

Aboud, F. E. (1993). A fifth grade program to reduce prejudice. In K. McLeod(Ed.),*Multicultural education (The State of the Arts National Study Report no. 1).* Toronto, University of Toronto, Faculty of Education.

Aboud, F. E. & Doyle A. B. (1996). Does talk of race foster prejudice or tolerance inchildren? *Canadian Journal of Behavioural Science, 28*, 161-170.

Avery, P. G. (1988). Political tolerance among adolescents. *Theory and Researchin Social Education, 16*, 183-201.

Avery, P. G. (1989). Adolescent political tolerance: Findings from the research and implication for educators. *High School Journal, 72*, 168-174.

Bigler, R. & Liben, L (1993). A cognitive-developmental approach to racialstereotyping and reconstructive memory in Euro-American children. *Child Development, 64*, 1507-1519.

Black-Gutman, D. & Hickson, F. (1996). The relationship between racial attitudes andsocial-cognitive development in children: An Australian study. *Developmental Psychology, 32*, 448-456.

Burwood, L. & Wyeth, R. (1998). Should schools promote toleration? *Journal of Moral Education, 27*, 465-473.

Doyle, A. B. & Aboud, F. E. (1995). A longitudinal study of white children'sracial prejudice as a social cognitive development. *Merrill-Palmer Quarterly, 41*, 209-228.

Enright, R. D. & Lapsley, D. K. (1981). Judging others who hold opposite beliefs:The development of belief-discrepancy reasoning. *Child Development, 52*, 1053-1063.

Enright, R. D., Lapsley, D. K., Franklin, C. C. & Streuck, K. (1984). Longitudinal and cross-cultural validation of belief-discrepancy reasoning construct. *Developmental Psychology, 20*, 143-149.

Human Rights and Equal Opportunity Commission (1991). *Racist violence: Report of the national inquiry into racist violence in Australia.* Canberra: Australian Government Publishing Service.

Iram, Y. (2001). Education for democracy in pluralistic societies: The case of Israel. In L. Limage (Ed.), *Democratizing education and educating democratic citizens: International and historical perspectives.* NY: Routledge Falmer.

McColsky, H. & Brill, A. (1983). *Dimensions of tolerance: What Americans believe about civil liberties.* New York: Russell Sage Foundation.

Owen D. & Dennis, J (1987). Preadult development of political tolerance. *Political Psychology, 8*, 547-561.

Robinson, J., Witenberg, T. R. & Sanson, A. (2001). The socialisation of tolerance. In M. Augoustinos & K. Reynolds (Eds.), *The psychology of prejudice and racism* (pp 73-88.)Thousand oaks, CA: Sage.

Sigelman, C. K. & Toebben J. L. (1992). Tolerant reactions to advocates of disagreeable ideas in childhood and adolescence. *Merrill-Palmer Quarterly, 38*, 542-557.

Thomas, T. (1998). The great wall of racial divide. *Australian Quarterly Journal of Contemporary Analysis, 5*, 38-41.

Vogt, W. P. (1997). *Tolerance and education: learning to live with diversity and difference.* Thousand oaks, CA: Sage Publications.

Wainryb, C., Shaw, L. A. & Maianu, C. (1998). Tolerance and intolerance: Children's and adolescent's judgments of dissenting beliefs, speech, persons, and conduct. *Child Development, 69*, 1541-1555.

Walzer, M. (1997). *On toleration.* New Haven, MA: Yale University Press.

Witenberg, R. T. (2000a). *Racial tolerance and acceptance as the basis for harmonious living in a civil society.* Paper presented at AME, Glasgow, Scotland, 7th – 11th July 2000.

Witenberg, T. R. (2000b). Do unto others: Towards understanding racial tolerance and acceptance. *Journal of College and Character* [On-Line], *1*. Available www.collegevalues.org.

Witenberg, T. R. & McDowall, J. (2001). *In favour of tolerance: How young adolescents justify their stance about racial tolerance.* Paper presented at The 36th Annual Conference, The Australian Psychological Society, 21st – 24th September, Adelaide, South Australia.

Witenberg, T. R. & Guadagno, B. (2001). *The context sensitive nature of racial tolerance.* Manuscript in preparation.

Zellman, G. L. & Sears, D. O. (1971). Childhood origins of tolerance for dissent. *Journal of Social Issues, 27*, 109-135.

Dr Rivka Witenberg is a lecturer of Developmental and Cognitive Psychology at the University of Melbourne, Department of Psychology. Her main areas of interests are the development of moral and prosocial competencies as well as the influence of knowledge on thinking, problem solving and decision making. Other areas of interests are teaching and supervision.

Author Note

Requests for reprints should be addressed to Dr Rivka T. Witenberg, The University of Melbourne, School of Behavioural Science, Department of Psychology, The University of Melbourne, Melbourne, Australia 3061
r.witenberg@psych.unimelb.edu.au

I would like to acknowledge the helpful comments from Dr M. Ainley and Dr. C. Milnitsky-Sapiro and thank Rochelle Cairns and my students who helped to collect and code the data. The collection and the analysis of the data were supported by a research grant from the Department Research Support Scheme, Department of Psychology, the University of Melbourne, Parkville, Australia.

Promoting the Development of Talent in Technical Areas: Obstacles to Females Pursuing Technical Careers in Europe, Scandinavia Countries, Asia, and the United States.

Gender Inequity Among Academic Olympians Across the Globe: Theoretical Paradigms

Introduction

This article introduces the theme for this issue. The articles that follow examine how parents and teachers nurture or hinder the development of talent. Figure 1 illustrates our conceptualization of the interacting forces that contribute to the development of talent. Certainly, parents occupy a central position, but the schools and teachers share in the process. Our framework also adds gender as an overlapping sphere of influence. We see overlapping spheres of influence among all of these sources.

Nations are especially concerned about the development of technical talent because it is fundamental to both economic and military concerns. In the United State such talent is carefully monitored by the National Science Foundation because the nation needs a steady supply of scientists and mathematicians in the Science and Engineering (S & E) pipeline for continued economic growth.

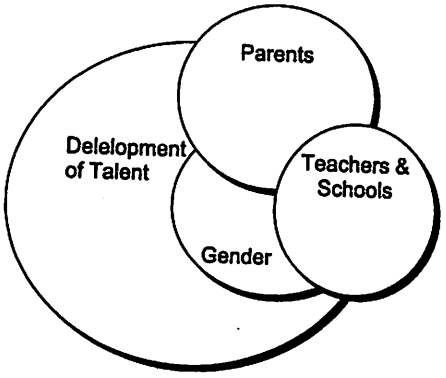

Figure 1. **Talent Development**

Our focus for these articles is on the development of talent in the technical areas (mathematics, chemistry and physics). To secure samples of individuals that possess these talents, we conducted retrospective studies of academic Olympians. Many people around the world are familiar with the winter or summer Olympic competitions, but there are also academic Olympic competitions in mathematics, chemistry, and physics for exceptional high school students. Each participating nation administers a series of demanding national tests to identify the best 20 students in these technical areas. Once these selections have been made, the Olympians are sent to a summer training camp that prepares them for the international competitions. At the completion of these training sessions, the top four or six students are selected and then sent to the International Olympics where they compete with the best students from each participating nation. Currently, there are more than 75 nations that participate in these Academic Olympics.

The Russians were the first to realize the potential of academic competitions and initiated the academic Olympics. The first program involved mathematics and was started in Leningrad in 1934 (Kukushkin, 1996). This mathematics Olympic competition was extended to city programs in Moscow and Kiev in 1935 and eventually spread to the entire USSR and beyond. The Soviets used these competitions to funnel talent into areas where they were needed. A student scoring exceptionally high on one of the academic Olympic exams was given automatic admission to the best universities. This admission placed the exceptional student in the Soviet S & E pipeline. In socialized countries national testing programs were conducted that assured the identification of a steady stream of gifted individuals. Once identified, these talented students could be funneled into areas where development was needed.

Our studies of academic Olympians have been underway since 1994 in China, Taiwan, Korea, Finland, Germany, Romania, and the United States. In each of these countries the different Olympiad cohorts in mathematics, chemistry, and physics have been tracked down for our studies. In some countries interviews are used to secure information; in other cases, mail surveys are used. Our studies accumulated a substantial amount of quantitative and qualitative data. These parallel studies collected data in a number of areas: socio-economic information; psychological dimensions of family; school influences (negative & positive); school grades; scores on standardized tests; high school graduation ranks; extracurricular activities (including musical interests & abilities); awards; college and university degrees; computer literacy; career histories; academic productivity (including publications, patents, software products); and evaluation of Olympiad programs.
The objectives of these studies are:

1. To determine if these academic Olympians fulfill their high potential;
2. To isolate the factors that help or hinder the development of extraordinary talent. Members of our research team have published the results of a number of studies dealing with these academic Olympians (see Campbell, 1996a, 1996b; Campbell & Wu, 1996; Shoho, 1996; Subotnik, Miserandino, & Olszewski-Kubilius, 1996; Wu, 1996; Wu & Chen, 2001; Zha, Liu, & Tao, 1996).

Methodology for the Olympiad Studies

Since the different national Olympiad research teams used the same methods and instruments, a brief description of them is provided. To secure the samples we followed four steps: 1. Secure the names and addresses of Olympians from the organizations that are responsible for running these competitions. 2. Verify the validity of these addresses. (Mobility in some countries makes finding adult Olympians very difficult.) 3. Mail packets of surveys and instruments to the Olympians where we have valid addresses. 4. Refinement of the database: Many packets were never delivered because Olympians have moved without leaving any forwarding address. (One national team used telephone interviews to secure survey information.) After utilizing these steps, stable sets of addresses were established, in effect, finalizing the size of each sample.

To obtain responses from the Olympians we used repeated mailings, telephone calls, and e-mail messages. With all of these contacts we uncovered no bias for the nonrespondents. Some of them replied after the 2nd or 3rd mailing. Some replied to a follow-up mailing that was done two years later. Some are overseas and cannot respond. The only complaint we got from them was that the packets are too long and require too much time to complete. Still, most of them responded eventually (95% of the U.S. Math Olympians responded).

Our research teams were especially concerned with validity issues. Our strategy was to ask for much of the same information from the Olympians and their parents as a means of checking the accuracy of the information. However, a few Olympians did not give us the addresses of their parents, and in a small number of cases we had addresses of parents who felt that their Olympian was too busy to reply. Consequently, we have a few cases where we have data from only Olympians or only from parents, but the overlap strategy in these cases supplied us with complete data. Overall, for over 90% of the Olympians we have data from both sources and after careful analysis concluded that both sets of information are virtually identical.

For the two instruments that were used in these studies (Inventory of Parental Influence, IPI; Self-concepts attitudes attribution Scales, SaaS), extensive factor analyses were done to establish validity. Further factor analyses were done with items within the surveys to isolate valid scales. In all cases reliabilities were derived for the scales used with these studies. For example, for the IPI, the alpha reliabilities for the following factors are: pressure α =.83; support α =.63; help α =.92; books and intellectual stimulation α =.80; tight supervision α =.87. For the SaaS attribution scales, the reliabilities ranged from α =.64 to α =.68. As an example of scales constructed from items within the surveys, we derived two hindrance scales: shortcomings of the schools α =.90; negative affect in the schools α =.76.

One common finding emerges in all the cohorts in the different countries — the scarcity of female mathematics, chemistry, and physics Olympians. The gender gaps are summarized in Table 1. Why are there so few female Olympians? These gender gaps are surprising because there are so many different cultures involved (including samples in Asia, Europe, Scandinavia, and the United States). All of the articles contained in this theme issue of *JRE* address this question. For the most part, the different national teams conducted in-depth qualitative studies to answer this question. Some articles also contain other studies that were related to this fundamental question.

Overview of Paradigms that Offer Explanations for Gender Gaps

Some researchers attribute the gender differences to overall male/female differences (Benbow & Lubinski, 1997; Benbow & Stanley, 1983a, 1983b; Page, 1976). Feingold (1992a, 1992b) attributes the gaps to more variability among males. Other researchers (Campbell & Beaudry, 1998; Eccels, 1982; Eccles, Kaczala, & Meece,1982; Chipman & Thomas, 1987; Chipman &

Table 1

Gender Gaps for Academic Olympians

Country	Ratio (males:females)
Finland	
math	10:1
USA	
math	44:1
physics	4:1
chemistry	6:1
Germany	
math	35:1
physics	95:0
chemistry	10:1
Taiwan	
math	17:1
physics/chemistry	11:1
China	
math	17:1
Korea	
math	15:1

Wilson, 1985; Fennema, 1983, Linn, 1986) believe that differential socialization factors are responsible. Plomin (1997) argues for analyzing the interactions between nature (genes) (44%) and nurture (46%) when analyzing such complex questions. The following brief summaries represent a primer of the paradigms that have been presented to explain gender gaps.

Explanations Depending Largely on the Biological Differences Between Males and Females (Deficit Theory)(Including Genetic & Hormonal Differences)

Deficit Theory. In 1971 the Study of Mathematically Precocious Youth (SMPY) uncovered large gender gaps (Benbow & Stanley 1980, 1982, 1983a; 1983b, 1983c; Lubinski, Benbow, & Morelock, 2000). The following gender gaps were found: SAT-M 500+2:1 (2 boys: 1 girl); SAT-M 600+ 4:1(4 boys: 1 girl); SAT-M 700+ 13:1 (13 boys: 1 girl). Page (1976) believes that these data "brilliantly satisfies [sic] all requirements" for a sex-linked recessive model of inheritance where genes on the X and Y chromosomes were responsible for some of the important traits involved with exceptional math achievement. This deficit would explain why males did better in this one academic area (p304).

Explanations Involving some Balance Between Physical Gender Differences and Socialization Forces

Combination of Nature & Nurture. According to this paradigm (Plomin, 1990, 1997), genes account for no more than 50% of variance for most traits. For example, for IQ, 44% of the variance is due to genes (nature), 23% can be attributed to common family factors (environment — nurture), and 23% can be traced to environmental factors outside the family (nurture). Home environment is important early in life but fades later; some genes are turned on later in life. For gender differences there is a complexity of factors.

Males Exhibit Greater Variability. Feingold (1992a, 1992b) and Hedges and Nowell, (1995) attribute gender gaps to the

fact that males exhibit more variability than females; more males are found at the extreme ends of the normal curve on many traits. For example, there are more feeble-minded males at one extreme and more genius males at the other extreme. This phenomenon occurs for several psychological and cognitive traits.

Explanations Depending Largely on Socialization Factors

Interaction of Numerous Factors Effecting Choice. The Eccles paradigm (Eccels, 1982a, 1982b, 1983, 1984, Eccles, Adler & Meece,1984; Eccles, Barber, Updegraff & O'Brien, 1995; Eccles & Jacobs, 1986) is the most comprehensive theoretical model yet to be developed. It represents a synthesis of other promising theories and frameworks. Much of this model is based on the perceptions of students as they make important decisions. Decisions are the product of complex interactions among 11 sets of factors. Within each set many variables operate. Each specific decision is not dependent on any one variable but is the result of numerous interactions, and gender stereotypes and gender typing occurs over time.

Females Camouflage Talent. Kerr (2000) believes that during preschool and primary school years gifted females are encouraged to develop their talents. However, during early adolescence and adulthood many gifted females learn to camouflage their talents in an effort to gain acceptance by other females and by males for dating and marriage. Furthermore, the families in many cultures encourage gifted girls to comply with traditions, and by doing so limit their career development.

Balance between Development & Nurturing Job Environment. According to this theoretical prospective (Theory of Work Adjustment —TWA) (Lubinski, Benbow & Morelock, 2000), there is a correspondence between each individual's abilities and the abilities inherent in the job (satisfactoriness). The correspondence between each individual's interests and preferences must be matched with the degree to which the job permits the nurturing of these personal qualities (satisfaction). Many gifted females have more balanced goals than gifted males — wanting to divide their time more equitably among marriage, family, and their career. This orientation undercuts their job satisfaction and limits their career growth.

Micro-Inequities/Macro-Inequities. Campbell and Beaudry (1998) break down gender differences into micro-inequities and macro-inequities. This approach emphasizes that there are hundreds or even thousands of socio-psychological variables where gender inequities occur. Many different social agents reinforce these inequities over time. Such socialization occurs time and time again at the micro level where it is subtle, easily missed, or overlooked. A good many of these micro-inequities arise when parents emphasize "masculine" behavior for their sons and "feminine" behavior for their daughters. They want their sons to exhibit masculine ways of acting and their daughters to be "lady-like." This emphasis is responsible for starting boys and girls down very divergent roads. Eventually these subtle micro-inequities accumulate over time to produce observable gender gaps and gender stereotypes. This conceptualization views socialization as a strong outgoing ocean tide that is very difficult to overcome. People get caught up in such a tide and "learn to go with the flow."

The articles that follow summarize qualitative data from Finland, the United States, Germany, and Korea. These articles use the research questions listed below to address the gender inequities that were uncovered among the Olympians.

1. Does an aversion from competition hamper talented females in pursuing technical careers?
2. Are parents somewhat responsible for this lack of encouragement?
3. Do parents discourage the development of their daughter's extraordinary talent in the technical areas?
4. Does family socialization hamper a female's development?
5. Are teachers responsible for not recognizing talented females?

The final article in this theme issue deals with how well the findings from these international Olympiad studies support or refute the eight theoretical paradigms that have been derived to explain the existence of the gender differences in the technical areas. How well do any of these theoretical frameworks fit the Olympiad data?

James Reed Campbell
Guest Editor
St. John's University

References

Benbow, C., & Stanley, J. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 1262-1264.
Benbow, C., & Stanley, J. (1983a). Differential course-taking hypothesis revisited. *American Educational Research Journal,*

20(4), 469-473.

Benbow, C., & Stanley, J. (1983b). Sex differences in mathematical reasoning ability: More facts. *Science, 222*(4627), 1029-1031.

Benbow, C., & Stanley, J. (Eds.). (1983). *Academic precocity: Aspects of its development.* Baltimore, MD: The Johns Hopkins University Press.

Benbow, C., & Stanley, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal respective. *American Educational Research Journal, 19*(4), 598-622.

Campbell, J. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *Journal of Educational Research, 25*(6), 497-521.

Campbell, J., & Wu, W. (1996). Development of exceptional academic talent: International research studies. *Journal of Educational Research, 25*(6), 479-483.

Chipman, S., & Thomas, V. (1987). The participation of women and minorities in mathematical, scientific, and technical fields. *Review of Educational Research in Education, 14*, 387-430.

Eccles. (1983). Expectancies, values and academic behavior. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W.H. Freeman.

Eccles, J. (1982). *Sex differences in achievement patterns.* Paper presented at the Annual Meeting of the American Educational Research Association.

Eccles, J. (1984). *Do students turn off to math in junior high?* Paper presented at the Annual Meeting of the American Educational Research Association.

Eccles, J., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology, 46*(1), 26-43.

Eccles, J., Barber, B., Updergraff, K., & O'Brien, K. (1995). *An expectancy-value model of achievement choices: The role of ability self-concepts, perceived task utility and interest in predicting activity choice and course enrollment.* Paper presented at the Annual Meeting of the American Educational Research Association.

Feingold, A. (1992). The greater male variability controversy. *Review of Educational Research, 62*(1), 89-90.

Plomin, R. (1997). Genetics and intelligence. In N. Colangelo & G. Davis (Eds.), *Handbook of gifted education (2nd ed.)* (pp. 67-74). Boston, MA: Allyn and Bacon.

Wu, W., & Chen, J. (2001). A follow-up of Taiwan Physics and Chemistry Olympians: The role of environmental influences in talent development. *Gifted and Talented International, 17*(1), 16-26.

Zha, Z., Liu, P., & Tao, X. (1996). Nurturing factors that promote Mathematics achievement in Mainland China. *Journal of Educational Research, 25*(6), 535-543.

James Reed Campbell is a Professor at St. John's University, Jamaica, New York (campbelj@stjohns.edu). He has authored over 200 publications and research papers including two popular books for parents, Raising Your Child to Be Gifted: Successful Parents Speak (Brookine Books) and Secrets of Productive Parents: Recipes That Work (Psychological Corp.)(in Mandarin).

Developing Females' Talent: Case Studies of Finnish Olympians

Kirsi Tirri
University of Helsinki, Finland

This qualitative study utilized in-depth interviews with 13 male and 13 female Finnish Olympians (matched by age and profession) to isolate the critical events that occurred in their lives that had an impact on the development of their talent. This sample ranged from pairs of college students to professional scientists in their mid careers and was drawn from a larger sample of math, physics, and chemistry Olympians (N=166). The common factor that contributed to the talent development of both females and males includes a strong inner drive in learning. The results reveal that the males make their professional choices earlier than the females. Furthermore, the males reported having more encouragement from their families and friends to choose science as their career. The results can be used to inform teachers and parents of the possible gender biases at homes and schools. This article suggests some concrete ways to encourage females to enter and remain in professional careers in science.

The Role of Female Scientists in Society

Due to the history of women's rights, Finland offers a culturally interesting context in which to explore the status and the role of female scientists. Women in Finland have a long tradition in the labor force, and the percentage of working women (approximately 80%) is among the highest in the world. Finnish women enjoy an advanced status internationally, measured by indicators such as general suffrage (granted in 1906) and the composition of the parliament (38.5% female). Additionally, in the year 2000, the first female president of Finland was elected. Women's rights and equality issues are under constant public debate. The advanced nature of sensitivities in this area can be seen, for example, in the fact that a father can take a paternity leave when a mother resumes her career. In academia, this kind of role reversal is becoming more common (Wager, 1994).

Given the advanced status of women in the Finnish society, we can assume that talented women in Finland are in control of their own lives, which contributes greatly to their talent and career development (Reis, 1998). Women in Finland gained equal rights to pursue academic studies in 1901, and in 1927 the first Finnish woman received a professorship. Today Finland tops the OECD (Organization for Economic Co-operation and Development) statistics in the percentage of women earning academic degrees. However, women and men continue to study different subjects, and there has been little change in this respect over the past couple of decades. Examples of women's choices are veterinary medicine, health care, pharmacology, and educational sciences, in which more than 80% of the students in 1995 were women. The technical subjects are, by contrast, clearly male dominated, with more than 80% of the students being men (Women in Academia, 1999). The proportion of women earning doctorates in Finland was 40.1% in 1996. However, the gender bias can be seen in the disciplines chosen by women. In 1995 the proportion of women obtaining doctorates was highest in medicine and the nursing sciences (53%) and lowest in technology (11%). The same gender bias concerning the disciplines has been reported for American women as well. American studies reveal that scientifically competent women are especially interested in biology, chemistry, medicine, and health rather than physics, computer science, and engineering. Helping others and doing something worthwhile for society serve as powerful motivators to attract gifted women to science in general and to the biosciences and health in particular (Rosser, 2000).

When gifted women enter academia they soon realize that the research faculty in Finnish universities is still clearly differentiated according to education and sex. The more teaching an appointment involves, the more likely the post is to be occupied by a woman, whereas the highest university appointments are clearly male dominated. Today in Finland the percentage of women professors is still very low (17%). The underrepresentation of women in university research communities, as well as male-led hierarchies, is a global phenomena. One special feature in all of the Nordic countries regarding the female role is that more academic women tend to be married and have children than, for example, in the United States (Women in Academia, 1999).

Evidently, many problems related to the position of

female scientists still exist. Various obstacles gifted females in academia face are related to cultural attitudes and women's roles. Kerr (1994) and Reis (1998) identified external barriers to gifted women, including the attitudes of parents and school personnel, environmental opinions, and possible discrimination or harassment at school or at work. The possible internal barriers for gifted females include self-doubt, self-criticism, and too low expectations. According to Siegle and Reis (1998), gifted girls tend to underestimate their abilities, especially in mathematics, social studies, and science. Research results also show a positive correlation between perceived ability and achievement (Multon, Brown, & Lent, 1991); therefore, women's self-perception of their own abilities becomes a key factor in keeping gifted girls involved in higher-level mathematics and science courses.

Many talented young women lower their aspirations and choose less demanding careers that they believe will enable them to marry. Women gifted in math and science continue not to pursue careers in these fields (Rosser, 2000). Instead they may select a more traditional career such as teaching or nursing. Those women who choose a career in science rarely produce as much creative work in the form of research publications as do their male colleagues (Ajzenberg-Selove, 1994; Piirto, 1991).

Case studies of Finnish Olympians add to the research tradition exploring critical events in the lives of successful people. In the field of gifted education, for example, researchers Kerr (1994) and Kitano (1997) utilized critical events as part of their data analysis. Kerr (1994) compared critical events to crucial moments. She found that major occupational or professional changes were usually in response to a family situation. Kitano (1997) examined women's interpretations of events in their lives, but a systematic analysis of those events is missing. Reis (1998) used similar sections of choices and compromises in the lives of gifted females in her book *Work Left Undone*. Tirri and Koro-Ljungberg (2002) studied the critical incidents in the lives of gifted female Finnish scientists. Their study revealed gender-related choices and compromises that female scientists in different fields and at various ages needed to make in order to balance their lives.

In this article we explore the role of critical events in the lives of gifted Finnish Olympians. The purpose of the study is to identify those events that helped them to succeed in their professional and personal lives. Our goal is to compare the critical events of females and males together in order to identify some gender-related patterns in their life histories.

Data and the Methods

Description of the data

The data of this study are based on interviews with Finnish Mathematical Olympians. The interviews with Olympians lasted one to two hours each, and all of the material from the interviews was transcribed. Additionally, curriculum vitae were collected and field notes taken to support the interpretation of their life stories. The whole sample of Olympians included 166 mathematics Olympians who participated in 1965-1999. Only sixteen of them were females. The quantitative description of the study can be found in Tirri (2001). The data for this article include in-depth interviews of the 26 Olympians (13 males, 13 females) who agreed to be interviewed. Every female was chosen to have a male Olympian who represented the same age group and professional orientation. This selection makes it possible to find similarities and differences in the critical events they experienced in their personal and professional lives.

Our qualitative sample includes 13 male and 13 female Olympians of different ages and fields of science. All the males chose mathematics, physics, or computer science as their major. The females' fields include mathematics, physics and medicine. Two of the females have Ph.D. degrees, and four of them are working on their doctorates. Six of the males have Ph.D. degrees and, all the rest are working on their degrees. Seven of the males are single, one is divorced with children, and five are married or engaged in permanent relationships with children. Ten of the female Olympians are in their twenties or thirties and single. The majority of them are students or graduate students in mathematics or physics. The oldest three female Olympians are in their forties or fifties, married, and two of them (Hanna and Kaisa) are well-known researchers.

According to the curriculum vitae, the male Olympians in academia publish more than their female colleagues. Patric and Matti have almost sixty publications in refereed journals and conference proceedings. The oldest female Olympian Kaisa, who is a well-known researcher in her field, has approximately thirty publications. Additionally, she is the only female with patents. Two of the males (Patric and Heikki) have patents in their list of academic productivity. Two of the male Olympians (Patric and Timo) are currently professors in their fields; none of the females are professors in their fields. The most successful female Olympians carry the title researcher in their curriculum vitae.

Most of the Olympians come from families of two or more children. The majority of participants are firstborn children in their families. The socioeconomic status and the education of their parents varied. On the whole, the Olympians had positive school experiences, and the greater part of them had been excellent students. Furthermore, they all emphasized the role of hard work in achieving good grades in school.

Methods

The researcher interviewed the Olympians by herself. In the personal interviews the professional and personal lives of Olympians were discussed. The themes discussed in the interviews included childhood experiences, school experiences, the choice of career, job, spouse, life-style, friends, and hobbies. A special emphasis was placed on the critical events identified by Olympians. Critical events in

this study were defined as experiences that helped the Olympians to identify their talent or to succeed in their career. In previous research critical events have been used as a tool of qualitative data analysis. John Flanagan (1953) was among the first researchers to describe the critical incident technique. He stated how the critical incident technique became essential when collecting certain important facts related to defined situations. Richardson (1995) postulated that people make sense of their lives in terms of special events and critical events. She went even further and claimed that connection between critical events constituted the meaning of narrative, whereas Denzin (1989) explained how lives are constructed and turned around by significant events.

In each interview the Olympians identified most of the critical events by themselves. In a few interviews the Olympians did not directly mention some of the events as critical in their lives. However, these events were identified as critical by the researcher during the data analysis. These kinds of events include early experiences in reading and mathematics.

Data Analysis

A content analysis was used to analyze the themes of critical events in the lives of the Olympians. In each interview all the events were counted and labeled (see Table 1). The female data were analyzed first, and the coding categories were established based on the critical events in their lives. Secondly, the male data were analyzed using the same coding categories. Most of the categories from the male subjects were the same. However, one new subcategory had to be added. After establishing final coding categories, the reliability of coding categories was tested. The interrater reliability was .90 based on the independent scoring of 26 interviews by two raters and an index calculated by the formula (number of rater agreement)/(number of life stories).

Critical events in the lives of Finnish Olympians

Events in childhood

According to our analysis, the male Olympians identified more important events in their childhood that influenced their talent development (see Table 1). All the males mentioned early reading experiences as critical events that helped them in actualizing their talents. Elina was one of the females who identified early reading experience as influential for her talent development. She learned to read at the age of four. The following quote from her interview demonstrates a typical reading experience by Finnish Olympians:

> I had a big brother and I followed his interests. I remember I read all his books. I can remember an event in the kitchen. I was playing with a milk carton when I noticed the word "milk" was written on it. I identified the letters and matched them together to form the word. I remember, how I felt the letters and sounds in my mouth, and it was wonderful!

The males reported early experiences related to mathematics as being critical as well. Timo, a professor in mathematics, described his early experience related to mathematics in the following way:

> I can remember an event in my very early childhood. I was only two years old, and I understood the number two. We lived in Helsinki in an apartment building, and I had a girlfriend who lived on the second floor. Once I passed her door, and I thought that now I am two years old and she lives on the second floor. When I turn three she will live on the third floor.

The females reported less early childhood experiences that were related to reading and mathematics. However, most of them were early readers, and some of them had equal experiences to men. The experiences reported by females were more associated with school. Riitta remembered a critical event in her childhood related to mathematics that influenced her to choose mathematics as her field. She described this event in the following way: "I was on the way to school, and I solved a math problem I had found very difficult. I had wondered what 15 +15 was. During that trip I realized how to solve that problem. I was very proud of myself. I can still remember the feeling."

The males reported more events related to science experiments and discussions with their parents than the females did (see Table 1). Many times parents encouraged their sons to explore nature and discussed scientific phenomena with them. The parents of females emphasized music and play more than scientific issues with their daughters. Only two females remembered any events related to science or discussions with their parents as critical for their talent development.

Events in school

All the Olympians enjoyed academic competitions during their school years. A young physicist describes her beliefs and values associated with challenge and competition in the following way:

> I need to know how things work, and I want to be challenged in my life. All the competitions have been very important for me. I want to test my knowledge and show to myself and to others what I can master in my field. I simply want to be the best.

A competitive mind was identified as a typical feature of gifted Finnish female scientists (Tirri & Koro-Ljungber (2002). The males seemed to enjoy the competitions even more than the females. One of the females participated in the

Table 1

Critical Events in the Lives of Finnish Olympians

Critical Events	Males (N=13)	Females (N=13)
Events in childhood	43	19
Reading experiences	13	8
Mathematics experiences	13	7
Science experiments	8	2
Discussions with parents	7	2
Events in school	40	31
Academic competitions	13	13
Teachers' encouragement	7	11
Peer support	10	4
Hobbies	10	3
Events in college	18	12
Studies abroad	8	4
Choosing the right domain	7	4
Mentor's support	3	4
Events in adulthood	12	8
International co-operation	6	4
Mentoring the youth	2	1
Partner choice	4	3

Olympics twice. Two of the males, Heikki and Jari, took part in the Olympics in mathematics and physics three or four times. In the interviews all the Olympians emphasized the importance of the Olympiad experience for their identity and talent development. For example, Heikki mentioned:

> The success in the competitions made me realize that I am really gifted, and the competitions have helped me to learn to trust myself in my academic studies as well. I had the self-confidence that was needed in order to succeed in doctoral studies.

Both males and females reported events related to teacher encouragement as critical in their talent development. However, teachers' feedback was more influential to the females than the males (see Table 1). Hanna described the influence of teachers in her talent development in the following way:

> I had wonderful female teachers in mathematics and chemistry. In physics I had a bad male teacher, but he was very nice. He was the type who wanted to be a researcher but ended up teaching. He was such a lousy teacher, but so nice and encouraging and warm. I studied everything by myself from the books; he couldn't help me academically. However, I liked the teacher, and that made me study even more.

Males reported more peer support during their school years. Evidently, it was easier for the boys than for the girls to find friends who had the same interests in science and math. Peer support also influenced their choices for college. Most of the male Olympians chose the same fields of science as their peers. Furthermore, their teachers and parents encouraged them to choose those fields. The females reported loneliness and lack of social contacts as being the most problematic experiences in school. Most of the females found peers with the same interests only in the upper secondary grades. The females reported fewer hobbies that were helpful for the development of their talent. Typical male hobbies included computers and chess. During their leisure

time the females enjoyed reading alone and other hobbies more than science.

Events in college

College had a great influence on the females' career choices. Those females (Hanna and Elina) that had a good mentor continued their studies at the graduate level. The Finnish Olympians reported less mentoring than their American counterparts (Campbell, 1996). Only four females and three males mentioned their mentors as critical persons in their career development. A typical Finnish Olympian reported independent study and inner drive to be the most important things in the development of their talent (Tirri 2001). Furthermore, they emphasized the importance of choosing the right field and studies abroad as more influential events in their career development (see Table 1).

The majority of them were happy with their career choices. The most successful researchers in both groups were those Olympians who were able to choose the right field; for example, mathematics or computer science, and had become internationally known researchers in that field. A typical answer to the question of why the Olympians chose mathematics and science as their domain is expressed in the quote by Sirpa:

> I have enjoyed my field because it involves studies in mathematics and science. It has been easy for me to advance in my work because I chose an academic area that allows me to enjoy the beauty and logic of math and science.

Twelve Olympians out of 26 mentioned studies abroad as a critical event in their lives (see Table 1). These Olympians include the most well known researchers with the widest publication activity of all the Olympians. According to our study, studies abroad have a crucial impact on the career development of Finnish scientists. Additionally, years spent abroad at various universities brought some of the Olympians to a deeper understanding of the opportunities and challenges in scientific work. As Olympian Hanna emphasized: "A researcher in the natural sciences has to work in different places and learn how to do research at the top."

Events in adulthood

Three of the females and four of the males mentioned the choice of a partner/spouse as an important choice affecting their success (see Table 1). The male professor in mathematics admitted that:

> My wife takes care of our three kids and runs the household so that I can concentrate on my research. I love the esthetics of mathematics, and it takes time to create a masterpiece. I have the circumstances that allow me to work the way I want.

The females emphasized the importance of having cooperative husbands who share responsibilities. Such a relationship is a critical asset in their lives. As Hanna reported:

> We are both researchers and we share the responsibilities at home. Both of us have the opportunity to work and take care of our home. My husband has always been very supportive because he understands the demands of this kind of work.

Four female Olympians and six males utilized cooperative liaisons with foreign colleagues to keep up with worldwide progress in their field, which has been reported as being significant in promoting one's career (Sawyer & Csikszentmihalyi, 1994). The most well-known researcher of all the Finnish Olympians, Patric, has worked with more than twenty colleagues from ten different countries in his international publications. Evidently, international co-operation opens doors to important publications and makes Finnish science known to the international research community. Three of the Olympians have continued their Olympic experience by mentoring future Olympians. These individuals give their leisure time to train and mentor young talented youth in mathematics. Matti acts as a leader of the Finnish math team; Jari and Riikka help him to train future Olympians.

Concluding Discussion

Our study reveals critical events that helped Finnish Olympians in actualizing their mathematical talents. The males had more critical events in their childhood than the females. Male Olympians identified early reading and mathematics experiences as more influential. Furthermore, males had more science experiments and discussions with their parents. According to previous research with gifted children, more boys than girls are encouraged by their parents to pursue a specific career (Reis, 1998). In early childhood, a child's reading, mathematics, and science experiments are greatly dependent on their parents' readiness to provide books and other materials. The case studies of Finnish Olympians reveal that males are provided more early encouragement in math and sciences than females.

During the school years both male and female Olympians enjoyed academic competitions and the Olympiad experience. However, the males took part in more competitions. The males had more peer support at school and hobbies that supported their talent development. However, the female Olympians more often reported teachers' encouragement as an influential factor than their male counterparts.

The most successful Olympians studied abroad during their graduate studies. Both males and females mentioned the choice of the right field as an influential factor for their academic success. The Finnish Olympians were very independent students, and only seven of them reported mentoring as a critical event in their talent development.

Those females who found a good mentor continued their graduate studies and chose an academic career.

International cooperation was identified as the most influential factor for academic success. The most well known researchers in our sample used international co-operation in their work and published in international journals. Both males and females identified a choice of a supportive partner/ spouse as a critical event. The females who referred to their husbands reported having equal responsibilities with child-care and opportunities for both spouses to actualize their talents. The male perspective emphasized the freedom to concentrate on academic work because of a wife who takes care of the household. Three of the Olympians reported positive events in mentoring future Olympians as influential in their talent actualization.

Our study suggests that parents should actively and at an early age encourage their daughters to do science experiments and mathematical problem solving. Furthermore, girls should be guided to pursue intellectual hobbies that support their talent development during their school years. Teachers should understand their important role in encouraging girls to choose mathematics and science for their college majors and careers. Furthermore, girls need help in finding support from their peers at school. Educators and mentors should guide gifted females to choose the right field and encourage them to study abroad. According to our results, the most successful scientists in academia need international cooperation to gain eminence in their fields. Furthermore, the importance of partner choice in the lives of gifted scientists should be addressed and discussed with gifted girls. The life histories and critical events experienced by gifted people can guide our efforts in providing equal opportunities for both females and males in actualizing their talents.

References

Ajzenberg-Selove, F. (1994). *A matter of choices: Memoirs of a female physicist.* Brunswick, NJ: Rutgers University Press.

Campbell, J. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *Journal of Educational Research, 25*(6), 497-521.

Denzin, N. (1989). *Interpretive biography.* Newbury Park, CA: Sage Publications.

Flanagan, J. (1953). *The critical incident method.* Pittsburgh, PA: American Institute for Research and University of Pittsburgh.

Kerr, B. (1994). *Smart girls: A new psychology of girls, women and giftedness. (Rev. ed.).* Scottsdale, AZ: Gifted Psychology Press.

Kitano, M. (1997). Gifted Asian American women. *Journal for the Education of Gifted, 21*(1), 3-37.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30-38.

Piirto, J. (1991). Why are there so few? (Creative women: visual artists, mathematicians, musicians). *Roeper Review, 13*(3), 142-147.

Reis, S. (1998). *Work left undone.* Mansfield Center, CT: Creative Learning Press.

Richardson, L. (1995). Narrative and sociology. In J. Van Maanen (Ed.), *Representation in ethnography* (pp. 198-221). Thousand Oaks, CA: Sage.

Rosser, S. (2000). *Women, science, and society.* NY: Teachers College Press.

Sawyer, R. K., & Csikszentmihalyi, M. (1994). Social influences on the creative process: An autobiographical memory study of creativity through the lifespan. In N. Colangelo & S. Assouline & D. Ambroson (Eds.), *Talent development: Proceedings from the 1993 Henry B. and Jocelyn Wallace National Research Symposium of Talent Development* (Vol. 2, pp. 467-471). Dayton, OH: Ohio Psychology Press.

Siegle, D., & Reis, S. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly, 42*(1), 39-47.

Tirri, K. (2001). Finland Olympiad Studies: What factors contribute to the development of academic talent in Finland? *Educating Able Children, 5*(2), 56-66.

Tirri, K., & Koro-Ljungberg, M. (2002). Critical incidents in the lives of gifted female Finnish scientists. *The Journal of Secondary Gifted Education, 12*(4), 151-163.

Wager, M. (1994). *Constructions of femininity in academic women*: Annales Academie Scientiarum Fennica Jyväskylä: Academia Scientiarum Fennica.

Women in Academia. (1999). *Report of the working group appointed by the Academy of Finland. Publications of the Academy of Finland 3/98.* Helsinki: Academy of Finland.

Kirsi Tirri is currently a Professor of Religious Education at the University of Helsinki, Finland. Internationally she has published numerous publications and research papers and is known for her work in moral development.

Author Note

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

German Olympiad Studies: Findings from a Retrospective Evaluation and from In-Depth Interviews. Where Have all the Gifted Females Gone?

Angelika Lengfelder
and
Kurt A. Heller
University of Munich

This article summarizes two studies with the German Olympians. The quantitative study utilized 235 academic Olympians; for the qualitative study 23 Olympians were interviewed about the underrepresentation of females in the Olympiad competitions. Both studies examined factors related to the development of talent. The results of these studies show that Germany has the largest national gender gap (28:1) of any of the other national Olympiad studies. We found no gender differences in ability or achievement for these Olympians, but we did find important differences in terms of motivational factors (academic self-concepts, attributions). We also found that the selection process used in Germany was flawed in its overdependence on nominations by teachers. This nomination process epitomizes the larger issue of teachers favoring males for their competence, especially in the technical areas. The article presents some important recommendations for promoting more gender equity in technical areas.

Introduction

The simple and intriguing idea of academic competitions is that they should help to identify and develop talent by offering features that Weinert and Wagner (1987) identified as stimulating, such as incitement (stimulating the participant's curiosity), options (tutoring, training programs), challenges, and incentives (rewards, chances for recognition). Do the academic Olympic competitions serve these purposes?

One of the intentions of this article is to examine differences between the male and female German Olympians. In addition, we will summarize the Olympians' views regarding the underrepresentation of women in these competitions.

Study 1 - Questionnaire - retrospective evaluation 1977 –1997

In Germany members of the Olympiad programs in math, physics, and chemistry annually conduct national searches to identify the most outstanding high school students. The endpoint of these searches culminates in the selection of a small team of semifinalists or prefinalists. These individuals are then invited to a national training camp where further selection is done to isolate the 4-6 top students (finalist's team) who are then sent to compete in the international academic competitions.

In this study participants of the finalist German Olympiad team as well as prefinalist participants and their parents were asked to complete questionnaires and instruments.

We had many clues that acceptance to the finalist team did not really discriminate between the very good (prefinalists) and the excellent achievers (finalists). We found no statistically significant differences on most measures between finalist and prefinalist participants. Similarly we found no important statistical differences among the German math, physics, and chemistry Olympians (Heller & Lengfelder, 2000).

Methods

Participants

The German academic Olympiad programs have been in existence since 1977 and have involved approximately 600 participants. We were able to track down addresses for a large number of these Olympians and were able to get overall responses from 40%. However, over the years the Olympians went to colleges and universities and then onto jobs. Some work at jobs overseas, while others occupy positions at universities in other countries. This inherent mobility over

decades makes them very difficult to find.

We labeled the 235 German Olympians in this study the evaluation sample. It consists of 100 finalists ($N_{math} = 28$, $N_{physics} = 41$, $N_{chemistry} = 31$) and 135 prefinalists ($N_{math} = 44$, $N_{physics} = 54$, $N_{chemistry} = 37$). The mean age is 27.8 years, the youngest person is 19 years old, and the oldest person is 42 years old. The majority of the participants are between 21 and 34 years of age.

Measures

Questionnaires were sent to as many former Olympiad participants and their parents as possible. The questionnaires dealt with the following research questions: (1) Do math, physics, and chemistry Olympians fulfill their high potential? (2) What factors contribute to the development of the math and science talent of the Olympians?

Results and Discussion

Rate of Female Participation

The Olympic competitions are analogous to the summer and winter Sports Olympics. Nowadays women participate in the sports Olympics, and they also compete in the academic Olympics. The difference is that there are no separate competitions for females in the academic Olympics. Suprisingly, for the academic competitions, the rate of female participation in all of the cross-national samples is shockingly low. The lowest rate (3.4%) is in the German sample compared to 10% in the American sample, 6% in the Chinese and Taiwanese samples and 9 % in the Finland sample.

To make sure that this finding was not biased, we compared the percentage of females (2 math prefinalists, 5 chemistry prefinalists, 1 chemistry finalist) in the evaluation sample (N=235) to the percentage of females that participated from 1977 to1997 and found a similar pattern (see Figure 1). The mean age is 25.8 years in the female sample (N=8) and 27.9 years in the male sample.

The present series of studies confirm earlier studies that documented a gross underrepresentation of females and of minorities among the Math Olympians (Campbell, 1996). This may be one of the most important findings as it raises many questions.

Academic achievement

We compared the academic achievement of the Olympians in their disciplines during their last two years at the German Gymnasium, which provides training comparable to college/ university preparatory courses, (see Figure 2). For the males, this analysis shows that math, physics, and chemistry Olympians had better grades (ranging from 1 very good to 6 not sufficient) in their Olympic subject ($M = 1.04$) than in their non-Olympic subjects. Their grades, however, in other subjects were still very good ($M = 1.44$); $F(1,198) = 334.44$, $p < .001$. The same pattern is found with the female participants ($M_{Olympic\ discipline} = 1.00$; $M_{non\text{-}Olympic\ discipline} = 1.49$). All participants - males and females alike - were highly successful in terms of academic achievement. In fact, 77% of the surveyed prefinalists, 82 % of the finalists, and seven

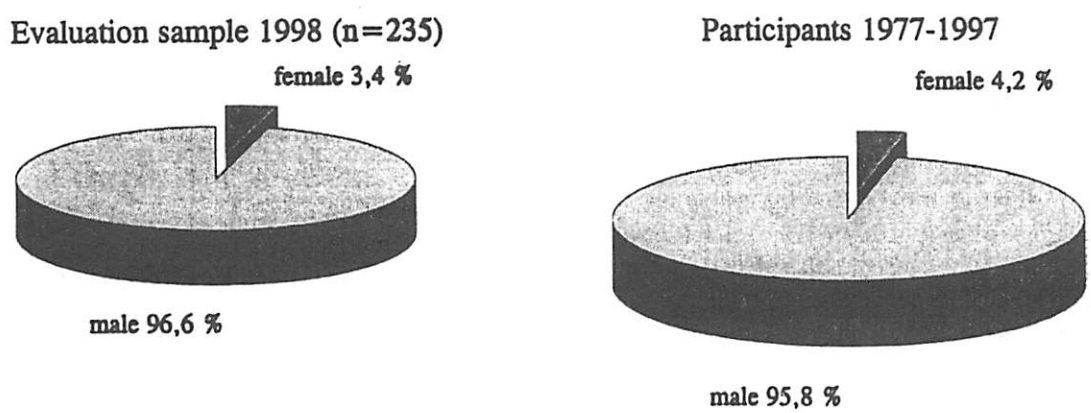

Figure 1. German prefinalist and finalist olympians: Proportion of female and male participants on math, physics and chemistry olympiads

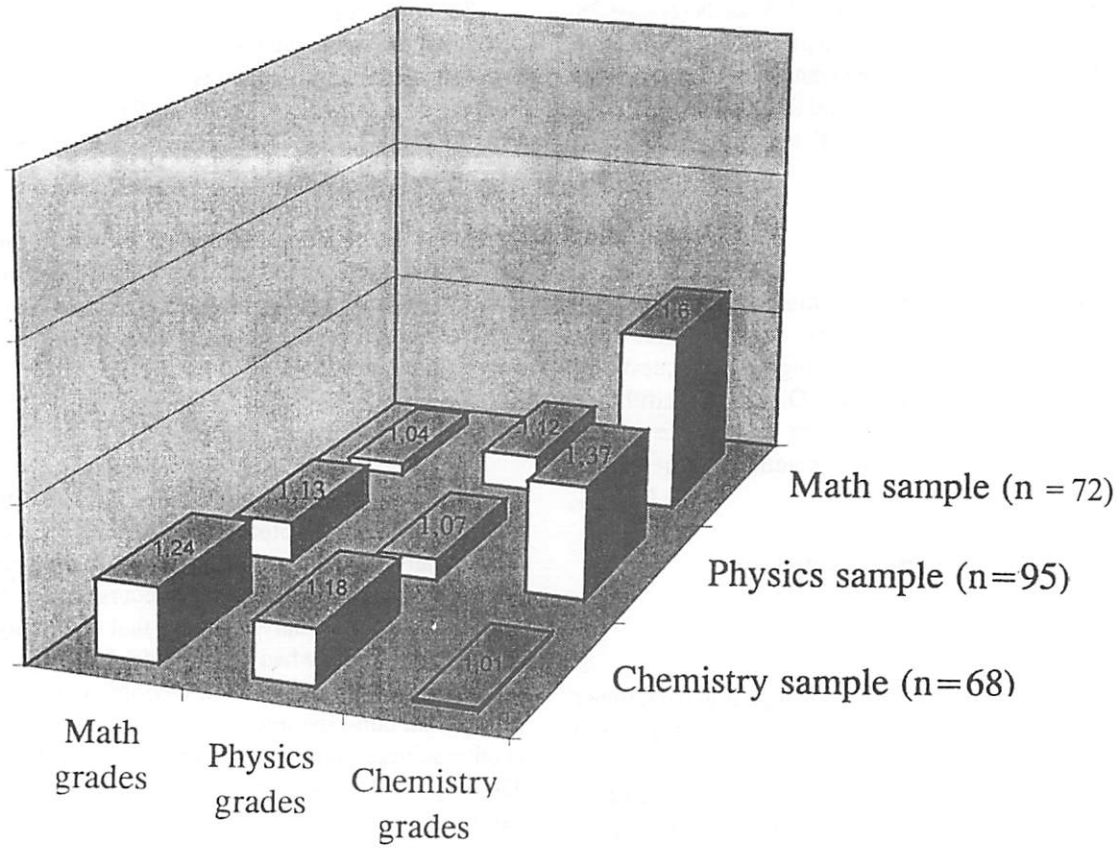

Figure 2. Grades in the final two years of the German Gymnasium

Note: 1 = highest grade; 6 = lowest grade

of the eight female participants ranked in the top ten in their final year in the German Gymnasium.

As adults, many of the German Olympians continue to work in the same discipline. The proportion whose current duties are either partly or completely related to their Olympic discipline, is remarkably high: 38% of the prefinalist participants (61% in math; 46% in physics; 24% in chemistry) and 53% of the finalist participants (60% in math; 53% in physics; 56% in chemistry).

Development of talent

With the males, among the factors contributing to the development of their talent, were a conducive home atmosphere and the availability of books at home. Both of these factors score significantly higher than stimulating school influences, (all $ts > 3.4$; all $ps < .01$). However, the female participants consider stimulating school influences more important than the home atmosphere and much more important than the availability of books at home. Additionally, many of the male participants (39%) believe that the main source for the development of their talent lies in oneself (i.e., in one's motivation, effort, initiative, curiosity, and playful desire to approach the subject), but 6 of the 8 female participants believe that a teacher's influence was the most important.

This finding that female participants consider teachers and stimulating school influences as most important corresponds with the observation made by Callahan et al. (1996) that encouragement by teachers and counselors is important when females apply for Math/Science/Technical schools. The results of this quantitative study of the German Olympians is described in more detail elsewhere (Heller & Lengfelder, 2000; Lengfelder & Übele, in press).

Study 2 - Findings from In-Depth Interviews

The cross-national quantitative studies also raise the question of why there are so few female participants in the Olympic competitions. To answer this question in-depth,

interviews were conducted with a randomly selected subsample. Some of the results of the in-depth interviews are presented here.

In the in-depth interviews we tried to explore the Olympians opinions about possible explanations for the large underrepresentation of females in Olympic competitions. One possible explanation is that the selection of the Olympiad teams could have been biased at different stages when the teams were initially assembled.

Methods

Thirty participants were randomly selected from the original sample of 235 participants, and 23 agreed to be interviewed (20 male participants and 3 female participants). We utilized a semi-standardized telephone interview of approximately 30 minutes. The interviewers were trained to use the interview guide listed in Table 1.

To understand the data we used a combined methodology. Besides statistically describing the data we also focused on qualitative data analyses. Classification schemes were developed, and single answers were also valued according to their impact, complexity, and originality.

Results

Beliefs about one's own talent and beliefs about the development of talent

Attributional patterns, such as Dweck (1996) describes, represent implicit theories of ability and provide the background for talent development. A majority of the participants thought that talent is 50% dependent on environmental stimulation and 50% genetically determined. However, one group of participants thought that talent is more than 60 % genetically determined, and another set of participants thought that talent is more than 60 % influenced by environmental factors.

Furthermore, when asked what would happen if talent lay fallow, more than 50 % thought that it is never too late to start training, while some participants even thought that talent cannot lie fallow because it would always develop on its own. But eight participants believed that negative consequences such as boredom, talent decrement, or even the development of a personality disorder would result.

Additionally, we wanted to know why some people work harder to develop their talents. Motivation and social background were named as the main sources, while only five Olympians thought that environmental stimulation was the main reason for this phenomenon. Seven of the Olympians thought that the social background, namely the influence of parents, and friends, and in some instances teachers are important sources. Six participants named one's own achievement motivation, and five participants said that what was crucial was interest, curiosity, and enjoying the subject. Three participants named socialization.

Table 1

Questions for Semi-structured In-Depth-Interviews

Implicit Theories of Giftedness and Talent

To what extent is your own talent innate? To what extent can talent in your opinion be developed? Please try to give a rating in percent!

What would happen if talents lie fallow?

How is it that some people thrive more in developing their talent.

Selection Bias

Are there special characteristics of the students who do get selected for the Olympic competitions and how do students get involved with this process?

How good, in your estimation, is the chance that gifted students will get in touch with the Olympic competition and similar competitions?

Why is the rate of female participation in the Olympic teams so extremely low?

What measures could be taken to increase the rate of female participation in the Olympic competitions?

Selection bias affecting recruitment of Olympians with special focus on the reasons for the low rate of female participation

We asked the Olympians about the experiences they had during the recruitment of the teams. Nearly half of the Olympians thought that the probability of making contact on their own with the Olympic competition was very low for any gifted student. Some thought that a better approach would be to use teacher nominations in combination with high grades and, in some cases, with one's own initiative. This enlarging of the selection criteria would give more gifted students the opportunity of participating in the Olympiad competitions.

It is not surprising that the former participants we interviewed were aware of the low level of female participation. In their judgement, the unsystematic selection process was identified as a main reason for the low rate of female participation. Most of the Olympians were convinced that the small rate of female participation did not match the actual rate of talented females.

In the commentaries (see Table 2), attitudes and expectations (on the part of parents and especially teachers) toward girls also emerged as important. Some mentioned the different attitudes of girls and boys toward competition. Others criticized the teacher's judgement as the sole source for nominations, therefore, instruction and the encouragement of girls in the classroom as well as improved dissemination policies concerning the Olympiads were seen as prime tools to interest more girls in this type of academic competition. Six interview partners (males and females) voted for delaying coeducation until an age when proper math and science self-concepts can be solidified.

Discussion and Conclusions

When participants in the questionnaire study (Heller & Lengfelder, 2000) were asked about their attributions, it turned out that all groups attributed success and failure more to ability than to effort. Accordingly, when asked about the sources of talent development, the majority of the Olympians interviewed ascribed at least 50 % or more of their talent development to genetic determination. Additionally, more than half of the interviewed sample thought that there was no risk in not recognizing talented individuals because talent would have the inherent capacity to develop on its own or could be developed later in life.

Social background and different personal goals and

Table 2

Examples of Commentaries on the Reasons for the Low Rate of Female Participation

Attitudes and Expectations towards Girls
Responsible: *stereotypical role* that girls have a special gift for languages (n = 4), and boys for sciences (n = 1)
Responsible: teacher's *expectations* (n = 2)
Responsible: *stereotypical role* girls missing self-confidence (n = 4)
Responsible: *stereotypical role* that girls are less talented than boys (n =2)
Actual Facts
The low rate of female participation in math or science "Leistungskursen" (n = 5) is a *fact*
It is a *fact* that girls are less talented than boys (n = 1)
It is a *fact* that girls are suppressed by boys in the classroom (n = 2)
It is a *fact*, that the co-education of girls and boys is counterproductive for girls (n = 1)
It is a *fact*, that girls have other non-scientific interests (n = 1)
It is a *fact*, that girls are not interested in competitions (n = 1)

motives were considered responsible for the phenomenon that some people work harder to develop their talents than others.

The majority of the sample considered the probability as too low that talented students will make contact with the Olympic competitions. Also it was critically shown that teachers' nominations are the main source of recruitment. Selection at an early stage via a combination of tests and teacher estimation as described by Feldhusen & Jarwan (1993) would certainly be an improvement, but this process would also increase the expense of the competition.

Female participation was considered to be far below the actual rate of talented females found in the German schools. This observation matches the results of other investigators that show females are underrepresented in math and science (e.g., Benbow & Stanley, 1983; Campbell & Beaudry, 1998; Lubinski, Benbow, & Stanley, 1993; Ziegler, Broome, & Heller, 1999), *but certainly not to the extreme extent that appears in the Math, Physics and Chemistry Olympiads, especially in the German sample.* Expectations (on the part of parents and especially teachers), social roles, and different attitudes toward competition as well as biased selection were isolated as main reasons why so few girls participate in the German Olympiad competitions.

Furthermore, all of the female Olympians interviewed emphasized that they thought girls would appreciate more cooperation (such as group work and projects) instead of competition. In a second follow-up study (Lengfelder & Übele, in press), the data from females in the fields of math, physics and chemistry who had never participated in final or prefinal Olympic competitions, were compared to data from the original study. The results indicate that the questioned females valued social support and solidarity higher than competition. This observation is in line with reports of the differential effects of cooperative versus competitive classroom interaction of girls and boys (Fennema &Peterson, 1987) as well as with findings on differences in the motivational orientation of girls and boys (Ziegler, Heller & Stachl, 1998).

The explanations given by our Olympians are close to the findings in the literature. First, the assumption levels of ability in boys and girls for mathematics, natural sciences, and technical subjects is not supported. Instead, comparable abilities are evident (Callahan, 1991, Beerman, Heller, & Menacher, 1992).

Second, motivational factors such as the self-concept of ability and related attributional styles (Heller & Ziegler, 1996) may prevent gifted girls from partaking in academic competitions.

Third, as girls prefer cooperation to competition (Fennema & Peterson, 1987), the way in which the Olympiads are organized with its different stages of competition and selection does not appeal to girls.

The development of female talent may not only be a function of the gifted females' distance from the mainstream (Noble et al., 1996), but also be related to teachers' interactions and to the environments within schools. It may also be more important to consider these contexts in order to help bright females in overcoming adversities. None of these factors are present in male biographies.

Some of the specific hindrances concern differences in socialization, the absence of role models (Raymond & Benbow, 1989), the missing of traditions in ritualized competitions, expectations, attention, reactions of teachers (Eccles & Blumfield, 1985), parents, and coping strategies that use mechanisms of befriending rather than competing (Taylor, et al., 2000). Considering all of these factors, it is no longer surprising that the female participants in the German sample ascribe a higher value to the influence of teachers and schools than the male participants do.

The instruction and encouragement of girls in school are prime ways to interest more girls in this type of academic competition. Also adjusting some of the organizational and recruitment features of the Olympiads may be necessary, (e.g., female teachers to accompany the teams or, as suggested by Le Maistre and Kanevsky, 1997, peer support groups). At an earlier level in education the conceptualization of coeducation in all disciplines should be critically reexamined. Furthermore, hope lies in using attributional retraining strategies to enhance girls' self-concepts of ability (Heller & Ziegler, 1996).

Positive effects of participation in academic competitions such as the Olympiad in math, physics, and chemistry have been reported. It would therefore be worthwhile to put more effort in the development and use of strategies to prepare and to attract more girls to these competitions.

References

Beerman, L., Heller, K. A. & Menacher, P. (1992).*Mathe: Nichts für Mädchen? Begabung und Geschlecht am Beispiel von Mathematik, Naturwissenschaft und Technik.* Bern: Hans Huber.

Benbow , C. P. & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: more facts. *Science,* 222, 1029-1031.

Callahan, C. M. (1991). An update on gifted females. *Journal for the Education of the Gifted.14,* 284-311.

Callahan, C., Adams, C., Bland, L, Moon, T., Moore, S., Perie, M. & McIntire, J. (1996). Factors influencing recruitment, enrollment, and retention of young women in special secondary schools of mathematics, science, and technology. In K. D. Arnold, K. D. Noble & R. F. Subotnik (Eds.), *Remarkable Women* (pp.243-261). Cresskill: Hampton Press.

Campbell, J. R. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *International Journal of Educational Research. Special Issue: Cross-national retrospective studies of mathematics olympians, 25(6),* 497-522.

Campbell, J. R. & Beaudry, J. S. (1998). Gender gap linked to differential socialization for high-achieving senior

mathematics students. *The Journal of Educational Research. 91(3)*, 140-146.

Campbell, J. R., Wagner, H. & Walberg, H. J. (2000). Academic competitions and programs designed to challenge the exceptionally talented. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of research on the development of giftedness and talent* (2nd ed., pp. 523-535). Oxford: Pergamon/ Elsevier Science.

Dweck, C. (1996). Implicit theories as organizer of goals and behavior. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action. Linking cognition and motivation to behavior* (pp.69-90). New York: Guilford Press.

Eccles, J. S. & Blumfield, P. (1985). Classroom experiences and student gender: Are there differences and do they matter? In L.C. Wilkinson & C.B. Marret (Eds.), *Gender influences in classroom interaction* (pp.79-114). Orlando: Academic Press.

Fennema, E. & Peterson, P. L. (1987). Effective teaching for girls and boys: The same or different? In D. C. Berliner & B. V. Rosenshine (Eds.), *Talks to teachers* (pp.11-125). New York: Random House.

Feldhusen, J. G. & Jarwan, F. A. (1993). Identification of gifted and talented youth for educational programs. In K. A. Heller, F. J. Mönks & A. H. Passow (Eds.), *International Handbok of research and development of giftedness and talent* (pp. 233-251). Oxford: Pergamon.

Heller, K. A. & Lengfelder, A. (2000). German Olympiad study on Math, Physics and Chemistry. Paper presented at the Aera conference meeting in New Orleans in April 2000.

Heller, K. A. & Ziegler, A. (1996). Gender differences in mathematics and the sciences: Can attributional retraining improve the performance of gifted females? *Gifted Child Quaterly, 40(4)*, 200-210.

Heller, K. A. & Ziegler, A. (1997). Gifted Females: A cross-cultural survey. In J. Chan, R. Li & J. Spinks (Eds.), *Maximizing potential: Lengthening and strenghtening stride* (pp.242-247). Hong-Kong: WCGT.

Lengfelder, A. & Übele, A. (in press). Bedingungen und Effekte der Schüler(innen)- Olympiaden in Mathematik, Physik und Chemie. - Teilnehmerinnen gesucht! *Psychologie in Erziehung und Unterricht.*

Le Maistre, C. & Kanevsky, L. (1997). Factors influencing the realization of exceptional mathematical ability in girls: an analysis of the research. *High ability studies, 8(1)*, 31-46.

Lubinski, D., Benbow, C. P., & Sanders, C. E. (1993). Reconceptualizing gender differences in achievement among the gifted. In K. A. Heller, F. J. Mönks & A. H. Passow (Eds.), *International Handbok of research and development of giftedness and talent* (pp. 233-251). Oxford: Pergamon.

Noble, K. D., R. F. Subotnik & Arnold, K. D. (1996). A new model for adult female talent development: A synthesis of perspektives from Remarkable Women. In K. D. Arnold, K. D. Noble & R. F. Subotnik (Eds.), *Remarkable Women* (pp.427-439). Cresskill: Hampton Press.

Raymond, C. L. & Benbow, C. B. (1989). Educational encouragement by parents: Its relationship to precocity and gender. *Gifted child quaterly, 33 (4)*, 144-151.

Taylor, S., Klein, L., Lewis, B., Gruenewald, T., Gurung, R., & Updegraff, J. (2000). Biobehavioral responses to stress in females: Tend and befriend, not fight-or flight. *Psychological Review, 107(3)*, 411-429.

Weinert, F. & Wagner, H. (Hrsg.) (1987). Die Förderung Hochbegabter in der Bundesrepublik Deutschland: Probleme, Positionen, Perspektiven. Bad Honnef: Bock.

Ziegler, A., Broome, P. & Heller, K. A. (1999). Golem and Enhancement: Elternkognition und das schulische Leistungshandeln in Physik. *Zeitschrift für Pädagogische* Psychologie, 13, 135-147.

Ziegler, A., Heller, K. A., & Stachl, S. C. (1998). Comparison of the general school related motivational set of average, gifted and highly gifted boys and girls. *Gifted and Talented International, 13*, 58-65.

Angelika Lengfelder is a researcher in the Department of Psychology at the University of Munich, Germany.

Kurt A. Heller is a Professor Emeritus of Psychology at the University of Munich, Germany. He has published extensively, and is one of the Editors of the International Handbook of Giftedness and Talent (2nd Edition) (Pergamon).

Author note

Correspondence concerning this paper should be addressed to Dr. Angelika Lengfelder/ Prof. Kurt A. Heller, University of Munich, Dept. of Psychology, Leopoldstr.13, D-80802 München, Germany; e-mail: lengfeld@edupsy.uni-muenchen.de or heller@edupsy.uni-muenchen.de.

We would like to thank the coordinators of the German teams in math, physics and chemistry Hans-Heinrich Langmann, Dr. Gunther Lind and Dr.Wolfgang Bünder for helping us to get in touch with the former participants.

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

Understanding Gender Inequity in America: Interviews with Academic Olympians

Annie Xuemei Feng
College of William and Mary

James R. Campbell
St. John's University

Marilyn A. Verna
St. Francis College

This study examines the underrepresentation of females in the academic Olympiad in the United States. Fifteen American academic Olympians (2 males, 13 females) participated in this interview study. Three major explanations account for the disparity of females in the academic Olympiad: (1) given the differences between males and females, the biased structure of the field of mathematics and science and academic competitions favors males; (2) societal expectation and social roles account for another explanation where the academic competition itself was colored as a more masculine arena that females have a tendency to avoid; and (3) differential socialization for males and females, starting from early on, provides a third explanation. A number of suggestions on how to ameliorate the situations are highlighted.

The retrospective studies of American Olympians is part of a larger set of International Olympiad Studies (Campbell, 1996; Feng, Campbell, & Verna, 2000; Feng, Campbell, & Verna, 2001; Verna & Campbell, 2000). A total of 218 academic Olympians were identified and participated in the survey studies. Results from these previous studies reveal that a conducive home environment and flexible, yet disciplined, parental influence, in conjunction with appropriate academic opportunities provided by schools, and the Olympians' internal motivations and interests, comprise the key variables that lead to the talent development and academic productivity of the American Olympians later in their careers.

As it occurs in the other participating nations (Campbell, 2002) in this theme issue), the female underrepresentation in the American academic Olympiad is not an exception. Among the sample of 218 American Olympians, only 11.4% of them are females (N = 25).

We aimed to understand the gender inequity issue by conducting in-depth interviews with American academic Olympians. This study serves three major purposes: (1) to explore possible explanations for gender disparity among talented female mathematics and science Olympians; (2) to provide a basis of comparison with other nations on the same issue; and (3) to find solutions that could alleviate the problem in America and possibly in other nations as well.

Review of the Literature

The existence of gender gaps in mathematics and science is not new. Extensive research has been conducted to explore gender differences in terms of academic achievement in the past decades (Baker, 1987; Linn & Hyde, 1989; Maccoby & Jacklin, 1966, 1974; and Wilder & Powell, 1989). These studies indicate that there are no significant overall differences in academic performance of girls and boys at younger ages. However, studies have indicated differential patterns of achievement favoring male students in mathematics and science starting in middle school and intensifying as students get older (Dweck, Davidson, Nelson, & Enna, 1978; Dweck, 2002; Fennema, & Ayer, 1984; Linn, & Hyde, 1989). This seems to be a robust achievement pattern over the years (Willingham, Cole, Lewis, & Leung 1997).

The gender gaps are aggravated when men and women face a choice of college majors and their future careers. Evidence shows that males tend to choose mathematics and science college majors, whereas females prefer to major in the humanities, education, psychology, and foreign

languages (Eccles, 1982, 1983; Eccles, 1994; Wigfield, & Eccles, 2002). Graduate school has already diverted a large percentage of females turned them out of the technical pipelines. In a report of the degrees granted in the United States in 1997-1998 (NCES, 2000), the percentages of female doctorates were as follows: mathematics 25.7%, chemistry 32%, physics 13.6%, and engineering 12.2%.

Numerous explanations and paradigms have been postulated for the gender disparity in mathematics, science, and engineering (see Campbell in this theme issue). To date, the underrepresentation of females in mathematics and science has been widely documented. The tendency seems to be that the higher the level of achievement, the more obvious the problem. The academic Olympians certainly represent one of the top levels in mathematics and science. Their experiences and struggles provide a better understanding of the unevenness of talent development in America.

Methods and Procedures

In this study data were collected through semi-structured phone interviews.

Participants

While we are interested in uncovering the reasons behind the underrepresentation of females in the academic competitions in mathematics and science, we are particularly interested in gaining insights into what made these female Olympians beat the odds and become the successful "minority" in these fields, and what they perceive as the hurdles that prevented a majority of their female peers from pursuing their learning and careers in the same areas. We decided to interview as many female Olympians as we could. Through E-mail or phone contact we tried to reach 25 female pre-finalist or finalists in the American sample (2 in mathematics, 13 in chemistry, and 10 in physics). We also sampled 5 male academic Olympians to serve as a comparison group (anchor the data). Finally, 15 U.S. Math, Chemistry, and Physics Olympians agreed to be interviewed (2 males, 13 females), including four who responded through E-mails. The participants were on average 27 years old ($SD = 5$), ranging from 22 to 42 by the time the interviews were conducted. There are two Asian American females in the sample; the others are Caucasians. Close to half of them (46.7%) were either working as assistant professors (one male, one female) or pursuing their Ph.D. or M.D.degrees ($N = 5$) in mathematics or science or related fields in the leading universities in the United States. Four of the Olympians were seniors at their universities majoring in chemistry or physics. Two female Olympians currently taught mathematics and physics at magnet high schools. There is one business owner who established her own software company; and one of the male Olympians is a lawyer.

Instrumentation

Based on current research literature and the discussion held by the principal investigators, we developed an interview protocol consisting of six questions (Campbell, 2002). However, in the middle of the interview sessions, a new theme emerged — balance between family life and career — and we included the question "What do you think about the issue of balancing a family and career?"

Procedures

We interviewed eleven Olympians on the phone (2 males, 9 females) and 4 female Olympians responded to our questions through E-mail. All interviews were tape-recorded and transcribed. Permission for tape recording was obtained in advance. Each member of the research team listened to the tapes once to clarify any confusion that was found in the verbatim transcripts. Open coding (Strauss, 1987) was used to organize, file, and analyze the data systematically. Each researcher derived themes from the transcripts independently, and we used the same or similar categories in the index sheet to code the data. In the data collection process, we kept a detailed interview log so that as a team we could easily retrieve the data. In addition, after each interview, new tentative themes were identified, and systematically occurring themes were built up in the questions for the following interviews. For example, one consistent concern of our female Olympians is their struggle for a balance between their career and family life. In the later interview sessions, all the female Olympians were asked about this issue. Marshall and Rossman (1995, p39) argued that qualitative researchers "should maintain the needed flexibility of research design" to accommodate changes during the course of an investigation. Content analysis was conducted to analyze the data further.

Results

Ten themes were derived from the analyses of participants' responses. We summarized these findings under four major headings (group differences, academic competition, socialization factors, and role models) where specific propositions from our Olympians' perspectives were delineated.

Group Differences

Our female Olympians note that males and females are different in a number of ways in terms of dealing with mathematics and science and the academic competitions in these areas.

> ***"The way that males communicate in mathematics and science scares a lot of females off sciences."***

One of the consistently mentioned concerns among female Olympians is that males and females have different ways of discussing and conveying their ideas about mathematics and science. More than a third of the female Olympians state that the direct and blunt way that males speak about a problem or offer criticism is very uncomfortable for females. In describing their experiences of working with a male and a female colleague in the lab, our female Olympians feel that women tend to be polite and thoughtful in a collaborative work, whereas men tend to be very blunt in making criticism. In a lab discussion a man would say, " you are wrong," or "that is stupid," which often make women feel annoyed.

"Females are more teamwork oriented; males prefer to work individually."

A majority of our female Olympians mention that females enjoy working with people; they are interested in teamwork and collaboration versus the males' preferences for individual work. However, the Academic Olympiad system is operated at the individual level, and it creates an environment where independent work and competition are stressed and collaboration is de-emphasized. One female Olympian illustrates these differences in the following way:

> Men work differently; they are not interested in teamwork ...The best research experience that I had has been working with women. There is much more questioning, much more exchange of ideas... I see women who are just as good in research as men, but they have different styles. And I think definitely male professors tend to encourage those students, including females, that work in a similar style to the way they have been told to...The male professor sees male students as more similar to himself. He tends to think male students are better. Whereas the reality is that a lot of them have styles of their own. Girls are more likely to want advice, more likely to want interaction and collaboration, [which] can be misconstrued by male professors, and even women professors, as being a sign of weakness; they are not good enough to do these things on their own. I think it is a matter of style.

"Girls or women tend to downplay their success and underestimate their performance."

It is well cited in the literature that females have a tendency to underestimate their ability and performance (Adam, 1996). This characteristic is not totally absent from our female Olympians.

One female Olympian shared her experience in the Olympiad training camp of how bright female high school students underestimated their potential in science:

> I had the experience of being there with another girl while I was there at the training program. I remember about a week and a half into the program and talking and discovering that both of us assumed we hadn't been really meant to be there; we hadn't really qualified. When one girl qualified, we realized that they needed another girl as a roommate, even if she did not qualify. Once we realized this, it was miserable. Both of us assumed we were the one who was not qualified.

The female Olympians' underestimation of their performances echoed the findings from Kramer (1991) that females tended to underestimate their ability. Interestingly, these female Olympians are extraordinarily talented females among the top in mathematics and science, yet they encountered a situation similar to what less talented females experience. It should be noted that there is a qualitative difference between the two levels of underestimation. Whereas females in the general population have a tendency to misjudge their abilities in mathematics and science, these female Olympians are quite confident about their academic competence in these demanding subjects. The real problem they faced seems to be the pressure coming from the atmosphere at the Olympiad training program. However, as one chemistry Olympian further stated, "I think women do well in the competition and like competition and thrive in competition in their own way."

"Males and females are different in pursuing their interests."

Different interests possessed by males and females have been given as another explanation for the gender gaps in mathematics and science. Several female Olympians expressed the view that females have broader interests; in addition, females want a balance for their different interests, whereas boys are more likely focused on one certain areas of interest.

One female Olympian expressed her view that "we tend to have broader interests, the interests in the humanities, and not just the narrowing interests in 100% math. Of course, the Olympians are the most focused people. The guys are pretty much exam-focused."

By contrast, male Olympians seem to perceive female Olympians' ways of dealing with their interests as having fewer interests in mathematics or science. As one male Olympian said, "given the opportunity, girls tend to be less interested in mathematics."

Academic Competition

"The connotation of competition and the structure of the academic competition favors males."

Competition has been portrayed more for its masculine characteristics; it is associated with speed, prestige, and aggressiveness. Females do not like to be identified with a

masculine image. "You don't want to be boys." Females' avoidance of the masculine image commonly associated with mathematics and science and competition turns girls off from these subjects early on despite their interest in and aptitude for these subjects.

One female Olympian shared her insights with us on how the "biased" structure of academic competition disadvantaged females:

> I really think academic competitions are like sports competitions, and a lot of the structure that constitutes the educational system is designed for the work of boys. And they are designed by the men for whom they worked.... The structures in which people do these difficult sciences are less welcoming and less comfortable places for girls, and therefore, they are not as likely to participate and also less likely to do well when they do participate.

In addition, the results of the academic Olympiad had a different impact on males and females that were involved with this academic competition. Females seem to possess a greater tendency to be affected by academic contests, and their self-esteem is more tied up with these competitions. One female Olympian confessed that she personally stopped participating in competitions in college because she got too upset when she did not do well. Given that there were so few females in the competition, she felt she must prove something. This was a common feeling among several of our female Olympians in their Olympiad experiences. This identical theme seemed to corroborate research findings that women have a greater tendency to be sensitive to criticism (Dwyer & Johnson, 1997). Moreover, females' lack of experience with the competition situation is cited as another likely explanation for the stronger emotional impact females had from the Olympiad experience.

Socialization Factors

In the analyses of the participants' narratives, the influence of various socializing agents emerged as an overarching theme addressed by all participants. These agents, including parents, teachers, peers, relatives, and the media, shaped and reinforced the existing gender-biased social roles and expectations, which, to a large extent, accounted for the current status quo of females' underrepresentation in mathematics, science, medicine, and technology. The following are the themes we derived from the Olympians' reports of their experiences.

"Parents tend to want the best for their children. They also see their children through the filter of cultural norms."

All of our participants stress the importance of the family in influencing daughters' and sons' interests and career paths. When asked if parents were less willing to accelerate math and science talents for their daughters, female Olympians unanimously responded that "willing is not the word." However, parents unconsciously raise their daughters and sons to conform to cultural norms.

What we learn from the interviews is that parents tend to foster different types of problem solving expertise in their sons and daughters without realizing it. More often than not, boys were brought up with toys such as building blocks, machines, and legos, which help build up spatial intuition, an important step for developing reasoning and thinking for mathematics and science; whereas girls received Barbies and tea-sets that gradually nurture their interests in dealing with things that are more human-oriented and develop their preferences for interaction and communication. These different types of toys orient them in different directions for development. Perhaps these early diversions account for differential interests among boys and girls.

Fortunately, most of our female Olympians report that their parents, although aware of the differences they might make in raising their daughters, encouraged them to pursue their interests in mathematics and science. Some parents openly dealt with the issue of sexism with their daughters in their adolescent years by alerting them well in advance about the societal stereotypes they might encounter. These parental efforts obviously helped the female Olympians excel in mathematics and science. One of the male Olympians we interviewed confessed that now, as a parent of a son and a daughter, he is well aware of the gender bias that parents practice. He takes practical steps to develop his daughter's spatial ability and to stimulate her interests in computers. Another Olympian who is now a mother of boy-girl twins shared her thoughts with us:

> Family socialization exists within an overall cultural context that is hard to shut out.... Families must nurture each child's talent while somehow shielding them from societal messages that may label the talent as abnormal or wrong.

"Social pressure from peers make females hesitate to pursue further mathematics and sciences."

Another obstacle for females going into mathematics and science and academic competitions arises from peer pressure. Our female Olympians clearly state that academics were not a problem for them. Indeed they loved math and science. However, they felt both socially and culturally stressed in excelling in mathematics and/or the sciences. For young adolescents, what their friends and peers perceive is crucial. In a typical American secondary school, a smart girl who is good at math or science is unpopular and almost unacceptable. Our finding, through female Olympians' descriptions of their experiences at middle or high school, supports Kerr's (1997) observation that smart girls camouflage their talents.

One female Olympian told a story of how she tried to

hide her talents in mathematics:

> I remember when I was young, I was accelerated in my classes, and this was something I hid from other middle schoolers. ...At that point I was taking calculus at the university... When I was with my peers, that was the sort of thing that I tried to keep under ... I don't know if that would be different if I had been a boy.

In American society, where a male's talent in mathematics and science is honored, our male Olympians were very fortunate in exhibiting and enjoying their talents in these hard sciences. One male Olympian told us that he did not [illegible] being thought of as someone who was intelligen[illegible]neavily in mathematics, yet he did observe the reluctance of his female classmates in "showing off" their talents in these academic areas.

Our data seem to suggest that males and females that are equally talented in mathematics or science receive different messages in their schools. Females appear to be discouraged by the covert assumption and belief of this society that doing well in mathematics and science is not woman's business.

"Teachers tend to hold lower expectations for female students in mathematics and science."

A majority of our female Olympians indicate that they had one or more excellent teachers who played a significant role in encouraging them to sustain their interests in math and science and to develop their talents and aspirations in the technological fields. Admittedly, our Olympians also point out that teachers might socialize their female students away from math and science. Subtly, but substantially, teachers might have unintentionally delivered different expectations for male and female students.

Several of our female Olympians who went to single sex schools express the view that it was a blessing for them to have been in an all-girl environment where they were not openly exposed to different expectations. However, they could not escape once they were out of that school setting. One female Olympian told us when and how she started to realize the different expectations at a state math contest:

> I had been in a math class with all girls all along, and they were all very smart girls, and our school did very well in state competitions. We placed second two years in a row in state math competition, and I remember a teacher saying when we went up to collect our awards, the teacher said, 'An all-girl team? I have never seen that before.'

The teacher's surprise and unbelievable tone in seeing an all-girls team who succeeded in these math contests delivered an unspoken but fairly clear message that females were not expected to do so well in mathematics contests.

An interesting finding we conclude from our interviews with the Olympians is that teachers' stereotypical expectations for female students are partially attributable to the misinterpretation of female students' more reticent behavior. Female students' modest way of responding, in contrast to boys' very assertive manner, might affect a teacher's judgment of a female student's true competence in mathematics and science.

Teachers might also show their different expectations in subtle ways. Indeed the over attention a teacher pays to his/her female students might damage their confidence level in mathematics and science. One female Olympian concedes that the help she received because she is a woman disabled her more than it has enabled her because it allowed her to coast at times, and it also made her believe that she could achieve less because she is only a woman who needs help.

Role Models

"There is a lack of positive female role models in math and science."

The historically lower ratio of female to male teachers in mathematics and science becomes more evident at the secondary level. It becomes more severe during the last few years in high school and then into college and graduate school. Students, even at the elementary level, are aware of this phenomenon when they are looking for a role model for their future ambitions. Our female Olympians unanimously contend that the lack of female role models in mathematics and science makes a great impact on the scarce talent pool of females in these domains. Moreover, a female role model who is successful in mathematics and science and is a charismatic person, who balances family life and career, is even rarer. It is much easier to find a charismatic female role model in other majors such as the humanities, education, and social sciences. One of our female Olympians summarizes the impact of the lack of female role models in the following way:

> For the Olympiad, I think the academic culture is very important. So is the fact that there are no or very few females who are at the top tier of theoretical physics or mathematicians in the world, [who] made an impact... It maybe not a direct one...If you are a little kid, if you don't think you can find a place in the world with your skills, then you are going to find something else to do, to a certain point. One way to judge if you are going to succeed in the world and find some place is to look for somebody like you. So having no women at the top is the problem there.

"Success means a balanced life between family life and career."

An ideal role model for our female Olympians is not only a woman who achieves the top in mathematics and

science, but also one who succeeds in her personal life; that is a charismatic woman who maintains a balance between family and career. This revelation led us to look further at the challenges our female Olympians are facing. For many of them, creating such a balance has been a major concern. In our society women often have to make a decision early in their lives about their career and whether or not to have children. This proved to be a difficult decision for many of our female Olympians. Some female Olympians admit that it is a topic they think about almost on a daily basis.

One female Olympian who was a senior and was going to pursue a M.D./Ph. D. joint degree program told us her concerns:

> With these MD/Ph.D. programs, you're committing to up to six years and approximately 8, 9, or 10 years for residency...I think it really does become an issue because I am thinking I am going to have children in the future...The nature and kind of culture in science is [such that] you really need to be immersed in science, and all the other people who we will be working with and competing against, a lot of them devote all of their time to research...I am going to spend about 8 years time in school, and I am 30 by the time I receive my Ph. D., and then a few years residency and fellowships; after that if I want a tenure-track academic position, a few years being an assistant professor before tenure,... There is a lot of career pressure all the way down the entire road, and by the time I really have a stable job and win tenure, I could be about 40 years old, and by that time, I can't plan at this time to wait that long to have children. That is something visible and is really hard to do...I am still very interested in the career goals, and I am going to pursue my MD/Ph.D. joint degree, [but] I would like to have a family and I am interested in that also. Exactly how things are going to work out in the end is a major mystery.

Several of our female Olympians expressed their concerns that in America, family was somehow a "women's issue." If a woman is to be successful in mathematics and science, she will have to face the academic pressure from her male colleagues. In addition, she has to struggle to manage a series of responsibilities if she wants a family. However, not many universities have a good support system for women who are expecting or who have children. As one female Olympian describes, "Seemingly there is an anti-family epidemic, particularly in academics. The much bigger responsibility of females for a family prevents many women from going to the very top."

Discussion

The dialogues we had with these Olympians indicate that males and females are different in a number of ways. The analyses of Olympians' narratives show that females enjoy team work and prefer communication and interaction at work. This result is in agreement with findings from numerous studies that show females are more likely to be person-oriented rather than thing-oriented (Eccles, 1983; Harding & Surtoris, 1984; Lubinski, Benbow, & Morelock, 2000). Yet females' preference for collaborative work and human interaction was discordant with how the mathematics and science world is structured where "rugged individualism" and manipulation of things are encouraged. Academic competitions in these fields underscore the male-oriented structure that make females a more vulnerable group of candidates for an event such as academic Olympiad. Moreover, females' preference for collaboration and teamwork is often interpreted as their weakness, which reinforces the social expectations for females in mathematics and science. Scarcity of females in math and science is likely attributed to the way these subjects are structured.

Our study also indicates that males and females tend to develop different dimensions of interests across the years. This result confirms findings made by Lubinski, Benbow, and Morelock (2000) who also found different interests among gifted male and female students. Admittedly, all of the female participants have strong interests in their quantitative fields. Yet what we know from the Olympians is that the spectrum of interests between males and females is different. Females tend to have a broader interest band, whereas males tend to be more focused. In addition, the analyses point out that females have a stronger desire to balance their various interests as compared to males' more focused interests in a certain field. Females' more diversified interests versus males' more focused interests partially explains why more males stay and sustain their interest in mathematics and science.

The competitive characteristics of the academic Olympiad, the way that the competition is conducted, and the masculine connotation of the mathematics and science Olympiad could all be related to the scarcity of female Olympians in the event. Academic competition has been specifically cited as a reason for women's avoidance of choosing a field such as engineering and science that is perceived by many women as involving direct personal competition in order to succeed (Holloway, 1993). Our female Olympians did participate in the most competitive academic activities, and they did choose mathematics and/or science as their major and possible future career. However, their narratives demonstrate that competition was the part that they enjoyed the least. What is most appealing to female Olympians is the experience of meeting with other high caliber peers and being able to work in a challenging environment. Our results, substantiate the claim that females do not like being in competition. Our data also indicates that, psychologically, females are more strongly affected by the results of academic competition than their male counterparts, a fear which might divert them away from the Olympiad.

Nevertheless, the analyses of the Olympians' narratives reveals that females' avoidance or dislike of competition does not necessarily mean that females are innately less competitive than males. Such a tendency is likely attributable to females' lack of experience with competitive events. Moreover, girls appear not to be cultivated in the way boys are toward competition.

Differential socialization has been perceived as a crucial factor in accounting for gender inequities in mathematics and science. Socializing agents at different stages of a child's life have passed on stereotypical messages concurring with the cultural norms and social expectations of society. The reasons that our female Olympians finally beat the odds and thrived in mathematics and science lay in their families' encouragement and support for the pursuit of their interests in these fields. Most female Olympians had some excellent teachers who valued their talents in mathematics and science and provided them with more opportunities to succeed in these fields. To some of them, a segregated environment of the single-sex school seemed to cushion them during the critical time of adolescence where social pressure was lessened and stereotypical expectations were minimized.

However, underneath the successful path of each female Olympian lay unique, yet similar, stories of females' struggles in a biased society to pursue their interests and to actualize their talents in mathematics and science. The female Olympians faced tremendous social pressure while dealing with academic pressure in the most competitive field of scholarship. The struggle and pressure accelerates as the academic tiers grow. This is particularly devastating in the top universities where a majority of these female Olympians currently work or study. Moreover, many of them have arrived or are arriving at an age of having families and children, which adds more challenges to their career. Striving for a balance between family life and a career in mathematics or science creates another challenge for these brilliant women. Unfortunately, these same challenges might turn off other talented women from pursuing careers in mathematics and science. Subsequently, there exists a scarcity of positive female models in these fields, a deficiency causing a large impact on talented young girls in middle school or even as early as elementary school.

Conclusions and Recommendations

Our interview study with Olympians reveals that the underrepresentation of females in the academic Olympiad is a complicated issue. We derived three major explanations from our dialogues with the Olympians. Given the group differences between males and females in preference of working style, communication, and interest, the fields of mathematics and science and academic competitions are structured in a way that favors males. The dissonance of females' preferences with the structure of the field contributes to the disparity of female mathematicians and scientists in general. Societal expectations and social roles account for another explanation where the academic competition itself is colored as a more masculine arena that females have a tendency to avoid. Differential socialization for males and females, starting from infancy, provides a third explanation.

The female Olympians in our study are facing a triple challenge by handling academic and social pressure from the field and females' inclination toward a balance between career and family life.

Despite the abundant data our Olympians provided, a limitation of our interview study is that we derived possible explanations for female scarcity from those females who stayed in mathematics and science and who excelled in the academic Olympiad. Those bright females who did not participate in these competitions and who avoided the field deserve further investigation. Their unique views on the same issue will surely enrich the current findings of this study.

Nevertheless, our Olympians gave us many valuable ideas on how to ameliorate the situation. Their suggestions include: (a) broaden the goals of the academic programs, de-emphasizing the competition aspect; (b) encourage girls to take more math and science courses and facilitate their interests in these subjects; (c) positively recognize females' achievement in mathematics and science; (d) hold high expectations for both male and female students in school settings and honor their talents equally; (e) make the qualifying exams for the Olympiad accessible at the grass root level; (f) inform parents about the impact of differential child-rearing practices on their children, particularly their daughters; (g) inform teachers about the differences between their male and female students, and recognizing these differences; let girls learn in their own way, (h) to provide academic programs for young women that will be supportive; (i) reform the format of the Olympiad program, including team work, by allowing group collaboration; (j) lobby the media to discourage advertising promoting gender bias messages and visual stereotypes; and (k) re-examine the structure of mathematics and science in general, redefining what accounts for the best mathematics and science spirit, and what are the alternative ways that these subjects can be done.

References

Adams, C. M. (1996). Gifted girls and science: Revisiting the issues. *The Journal of Secondary Gifted Education.* 447-458

Baker, M. A. (Ed.). (1987). *Sex Differences in Human Performance.* Chichester, England: Wiley.

Campbell, J. (1996). Early identification of mathematics talent has long-term positive consequences for career contributions. *Journal of Educational Research, 25*(6), 497-521.

Campbell, J. (2002). Gender Inequity Among Academic Olympians Across the Globe: Theoretical Paradigms.

Journal of Research in Education, 12, 75-79.

Dweck, S. C. (2002). The development of ability conceptions. In A. Wigfield, & J. S. Eccles (Eds.) *Development of achievement motivation.* (pp. 91-120). CA: Academic Press.

Dweck, S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex Differences in learned helplessness II. *The contingencies of evaluative feedback in the classroom and III. An experimental analysis. Developmental Psychology, 14*, 268-276.

Dwyer, C.A. & Johnson, L. M. (1984). Grades, accomplishments, and correlates. In W.W. Willingham, & N.S. Cole (Eds.), *Gender and fair assessment* (pp. 127-156). NJ: Lawrence Erlbaum Associates.

Eccles, J. (1982). *Sex differences in achievement patterns.* Paper presented at the Annual Meeting of the American Educational Research Association.

Eccles. J. (1983). Expectancies, values and academic behavior. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W.H. Freeman.

Eccles, J. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly, 18*, 585-609

Feng, A. X., Campbell, J. R, & Verna, M. A. (2000). *American Math, Physics and Chemistry Olympians: What makes them so talented?* Paper presented at the 8th international meeting of Chinese American Education and Development Association, New York.

Feng, A. X., & Campbell, J. R., & Verna (2001). The talent development of American Physics Olympians. *Gifted and Talented International, 16* (2). 108-114.

Fennema, E., & Ayer, M. J. (Eds.). (1984). *Women and education: Equity or equality?* Berkley, CA: MrCutchan Publishing Corp.

Harding, J., & Surtoris, M. (1984). *An object relations account of the differential involvement of boys and girls in science and technology.* Paper Presented to the International Symposium, Interests in Science and Technology Education, Kiel, Germany.

Holloway, M. (1993). A lab for her own. *Scientific American, 18*(8), 17-27.

Kerr, B. (1997). *Smart girls (revised edition). A new psychology of girls, women, and giftedness.* AZ: Gifted Psychology Press.

Kramer, L. R. (1991). The social constructs of ability conceptions: An ethnographic study of gifted adolescent girls. *Journal of Early Adolescence, 11*, 340-362.

Linn, M., & Hyde, J. (1989). Gender, Mathematics and Science. *Educational Researcher, 18*(8), 17-27.

Lubinski, D., Benbow, C., & Morelock, M. (2000). Gender differences in engineering and the physical sciences among the gifted: An inorganic-organic distinction. In K. Heller, F. Monks, R. Sternberg, & R. Subotnik (Eds.), *International handbook of giftedness and talent. (2nd ed.)* (pp. 633-648). Oxford, UK: Pergamon.

Maccoby, E. (1966). *The development of sex differences.* Stanford, CA: Stanford University Press.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Marshall, C., & Rossman, G. (1994). *Designing qualitative research, 2nd edition.* CA: Sage Publications.

National Center for Education Statistics. (2000). *Digest of education statistics.* Washington, D. C. U. S. Department of Health, Education, and Welfare, Education Division, National Center for Education Statistics. (pp. 298-305)

Strauss, A. L. (1987). *Qualitative analysis for social scientists.* New York: Cambridge University Press.

Verna, M. A., & Campbell, J. R. (2000). *Career orientation of American Chemistry Olympians.* Paper presented at the Annual Meeting of American Association of Educational Research, New Orleans.

Wigfield, A., & Eccles. J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.) *Development of achievement motivation.* (pp. 91-120). CA: Academic Press.

Wilder, G. Z., & Powell, K. (1989). *Sex differences in test performance: A survey of the literature (CB Rep. No. 89-3; ETS RR-89-4).* New York: College Entrance Examination Board.

Willingham, W. W., Cole, N. S., Lewis, C., & Leung, S. W. (1997). Test performance. In W. W. Willingham & N. S. Cole (Eds.), *Gender and Fair Assessment*. NJ: Lawrence Erlbaum Associates Publishers.

Annie Xuemei Feng is a researcher at the College of William and Mary, Williamsburg, Virginia.

James Reed Campbell is a Professor at St. John's University, Jamaica, New York.

Marilyn Ann Verna is an Assistant Professor at St. Francis College, Brooklyn, New York.

Author Note

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

American Chemistry Olympians Achieve the Highest Level of Equity

Marilyn Ann Verna
St. Francis College

Annie Xuemei Feng
College of William and Mary

The female American Chemistry Olympians consider ability to be a more important factor for success. There is no gender difference for the effort factor. A significant difference is also observed between the American chemistry males and females with regard to the conducive home atmosphere. The females view the home as a place of educational productivity where their parents recognized their talent and provided stimulation to foster its development. Parents provided educational books, magazines, and association with inspiring peers and relatives. Parents actively encouraged their daughter's success in the technical area.

Description of the American Chemistry Olympiad Program

Many countries around the world sponsor academic contests as a way of identifying and developing their most talented science and math students. One of these international competitions is the Chemistry Olympiad. This competition began in 1968 among European countries. The United States commenced participation in 1984. Over 50 countries send representatives to this competition. The students compete in written and laboratory examinations that include problems in analytical, organic, inorganic, physical and biochemistry. Points are awarded for each section of the competition by a panel of international judges. Students are ranked according to the total number of points they receive.

The selection of students begins at the local level. The local high schools select their top chemistry candidates. Each section of the United States designates at least 5 students to take the initial nationwide exam sponsored by the American Chemical Society. Based on the results of this exam, 20 students are invited to spend two weeks at a chemistry study camp at the Air Force Academy in Colorado Springs. Four students are then selected from the camp participants to represent the United States in the International competition. In the eighteen years that the United States has participated, the competitors have won 15 gold medals, 26 silver medals, and 19 bronze medals. Included in these totals were the 1999 winning medalists. Of the 199 high school age competitors representing 51 countries, Lisa Carlivati won a silver medal, Wei Ho, another female, and Alexander Ho each won a gold medal, and Timothy Jones had the highest individual score in the competition earning him the top gold medal. Thus, the U.S. team finished first worldwide (American Chemical Society, 1999).

The goals of the United States National Chemistry Olympiad (USNCO) are multipurpose. USNCO views the program as a means of investing in the future of young chemists. The program is a way to stimulate all young people to achieve excellence in chemistry, to recognize and challenge the knowledge and skills of outstanding chemistry students, and to recognize the achievement of the teachers of these students and the importance of the school environment in which they learn.

Objectives

In this theme issue of the Journal of Research in Education, 11 international samples of academic Olympians are represented. The sample with the most females is that of the American Chemistry Olympians. Some of the Europeans gender ratios are as high as 95:1 (males: females); the Asian samples are 17:1 and 15:1; the US Chemistry Olympians had a ratio of 6:1. This study attempts to determine why this technical subject in America attracts so many females. How did the American Chemical Society accomplish this level of equity?

Design and Procedure

Subjects

This study was conducted with United States high school Chemistry Olympians from the years 1984-1998. During this time period the number of United States participants in the Olympiad program has been 300 Olympians. One hundred seventy-seven Olympians were located and sent the survey and questionnaire materials. Of the 177 surveys sent, 133 had stable addresses where the materials could be delivered. The final number of participants, consisting of students and/or parents, in the study was 93 (70%). The chemistry Olympians (70 males and 13 females) ranging in age from 16 to 34, and 102 parents (52 fathers and 50 mothers). Ninety-five percent of the students were born in the United States. The ethnic composition of the students was 86% Caucasian and 14% Asian.

Instrumentation and Methods

Each student was asked to respond to the Self-Confidence Attribute Attitude Scale Form O (SaaS) (Campbell, 1994) that measures the student's attribution level. The SaaS contains 29 Likert scale items ranging from strongly disagree to strongly agree. Principal Components analyses was conducted on all scales. This survey produced two distinct scales, effort and ability. Olympians who viewed success as an outgrowth of effort scored higher on the effort scale. Those Olympians who linked ability as a reason for success scored higher on the ability scale.

Each student and his parents were asked to respond to the Family Influence Scale (14 Likert scale items), and the School Influence Scale (11 Likert scale items) (Campbell, 1996) which measures the student's opinion of the importance of positive family influences and school hindrances on the development of one's talent. The respondents rate each statement as (a. none; b. almost none; c. little; d. some; e. much; f. great). Sample statements of the Family Influence Scale are: "Everyone in the family was an avid reader." "Being accelerated or taking advanced courses." Sample statements of the School Influence Scale are: "Not enough challenge." "I knew more than many of my teachers." Alpha reliability coefficients were calculated for each of the scales: effort (α =.96), ability (α =.93), family influence (α =.95), and school influence (α =.90).

Each student and parent was asked to respond to open-ended questions on the written questionnaire. These questions asked about possible contributing factors to the development of the Olympian's talent. The following questions were asked. "Describe any influence that helped to develop your talent/your child's talent." "Were there any negative school experiences you can describe?" "What did the Olympiad experience mean to you/your child?" "What training experiences were the most valuable?" "Did the intensity of the Olympiad experience 'burn you out?'" "Have you/your child experienced any hostility because of your giftedness?" " Have you ever been accused of being an elitist?" "What other important issues or factors do you consider important for the development of talent?"

Information regarding the parents' occupational and educational background was collected from the Olympian.

Results and Discussion

Characteristics of Achievement

The Chemistry Olympians seem to believe that ability is needed in order to perform well in a subject area. Simpson and Oliver (1990) studied gender differences in science achievement. They found that although males achieved higher scores and possessed a more positive attitude than females, the females were significantly more motivated to achieve in science. In this study the Chemistry Olympians expressed the view that effort is linked to ability as a reason for achievement. Although they may have the necessary intellect to succeed, the Chemistry Olympians believe in working hard and employing self-discipline to their endeavors. There is a small significant difference between the males and females in the study with regard to ability (see Table 1). The females consider ability to be a more important factor for success than the males. There is no difference for the effort factor. This finding is opposite to Wu and Chen's (1999) findings of Taiwanese Chemistry and Physics Olympians. Wu's Asian population attributes their achievement more to effort than to ability.

Competition is another aspect that our Olympians explored. Our female Olympians are as competitive as the male Olympians. However, they realize that in society females tend to be more humble, are less competitive, are more reticent, down play their achievement, and try to be an equalizing factor in society. Whereas the males tend to volunteer more, tend to be more aggressive, and hierarchical in nature. One young female states that in her view 95% of the participants in the Olympiad program were males, and that the program seemed to be designed by men to appeal to 16- or 17-year old boys.

School Influence

Positive school influences helped in the development of the gifted student's achievement. Some parents reported that the very good teachers were demanding and well prepared. Science fairs, special school for the sciences, and summer programs for advanced studies were a plus for their education.

Many math and science teachers recognized and encouraged the talents of female students. One female Olympian attended a single-sex high school. In order for her to enroll in AP physics or AP chemistry classes she had to attend the boy's high school. In those classes the males treated her as an equal. She had a wonderful teacher who

Table 1

Means and Standard Deviations of the Study's Variables

Variable	Mean	Standard Deviation	Scores	
			Minimum	Maximum
Extracurricular Activities	3.84	2.34	1	9
Effort (EFF)	3.29	.45	1	5
Ability (ABIL)	2.98	.40	1	5
Failure Due to Lack of Effort (EFAIL)	3.53	.60	1	5
Success Due to Effort (ESUC)	3.23	.79	1	5
Success Due to Ability (ASUC)	2.22	.89	1	5
School Hindrances (Neg. Sch.)	2.33	.84	1	6
Family Influence (Fa. Inf.)	4.79	.78	1	6

recognized her talent and encouraged her to study chemistry. Another female contestant was so intellectually stimulated in her high school science classes that upon college graduation she taught for two years in a magnet school for math and science.

Unfortunately, more negative aspects than positive were reported. Students felt impaired by the school environment. The schools provided neither a challenging environment nor high level courses. They said the elementary school teachers lacked knowledge of subject matter especially in math and science, and an inability to motivate these students. In the 3rd grade, one Olympian came home to his mom and said, "I'm not getting to use my brain." The parents switched the child to a private school. Other Olympians complained that there was no flexibility in class work. The Olympians viewed their classes as rigid, boring, almost prison-like, and lacking competition. One Olympian was so bored in his English class that this inspired him to study for the Chemistry Olympics.

Several of these gifted students had to hide their intelligence, especially the girls. Their intelligence was viewed as being fresh. Being on top wasn't "cool." In many instances the teachers resented, embarrassed, and belittled the gifted student. Parents recalled the mean-spiritedness of the mainstream parents and teachers. In high school these gifted students raised the curve in class tests. Students and parents resented this. A mainstream parent asked why the gifted student was not advanced into college. Parents cited the following actions taken by the teachers and students:

- A computer teacher erased his son's program so he would have to start again.
- The teacher said he would take him out of the honors class because he was sloppy.
- The teacher said the child needed humility.
- Another teacher said that the child's job was to help others.
- More time was spent with the poorer student because the teacher did not know what to do with the gifted student.
- Students would make fun of the young female if her paper was not perfect.

Family Influences

To overcome the school situation, the home atmosphere offered positive influences. The Olympians and their parents cited a strong family emphasis on education. Olympians attribute development of their talent to older siblings and friends. Exposure to friends and family members who were interested in the sciences was a common factor. Many of the parents, themselves, were highly intelligent and were

scientists. Curiosity and self-motivation were forces in seeking knowledge for these students. Students were encouraged to take responsibility for their own learning and be independent learners at an early age. The parents were not in favor of skipping grades. Instead, they provided horizontal enrichment rather than vertical, except for math. Parents provided excursions to stimulating places, visits to historical places, concerts, art exhibits, and opportunities to partake in extra-curricula activities. Security and relaxation in the home was provided.

We observed a significant difference between the males and the females with regard to the conducive home atmosphere. The females view the home as a place of educational productivity more than the males. The females feel that their parents recognized their talent and offered more stimulation to foster the development of this talent. Parents provided educational books, magazines, and association with inspiring peers and relatives. One young woman Olympian said that when her parents realized that she was interested in the Chemistry competition, her mom found a woman chemistry friend who was willing to tutor her. They would review chemistry exams and discuss questionable answers.

Parents actively encouraged their children's success. Many parents were instrumental in getting the high schools to initiate a gifted program or more challenging course for this population of students.

The Olympiad Experience

The Olympians and parents expressed their views on the entire Olympiad experience. Their experience in this competition had a more positive than negative impact on their lives. Sixty-six percent of the Olympians believe they would not have succeeded without this experience; it did make them more aware of the educational opportunities available to them and helped them in accepting their talent. The Olympians believe that the experience increased their confidence and gave validation to their abilities, and provided encouragement for the future. It was an exciting experience. They enjoyed meeting interesting people, making international acquaintances, and lasting friendships in some cases. It offered the gifted students a chance to communicate with students equal to them in ability and interests, and people smarter than they. One female Olympian believes that the Olympiad experience was a very enriching educational experience. She learned a tremendous amount about herself through the intellectual challenge that she had never known before. It was a maturing experience. The Olympians are pleased that this gave them formal recognition for all their hard work. Overall, the Chemistry Olympians have been very productive in their professional academic productivity (Verna, Campbell, & Feng, 2002).

With regard to the chemistry material studied, an emphasis was placed on organic chemistry with individual lab sessions that were not available in their high schools. Two weeks at the camp was so intense that the work equaled one semester of chemistry, one semester of inorganic chemistry, and one semester of calculus. They enjoyed the training experience and the opportunity to work with others. They felt challenged beyond their limits.

Receptiveness of Society to Highly Gifted Students

Positive comments about the community were expressed. The schools were receptive to parents volunteering their time in school. In fact, one school asked that the parents devote one hour per week for each of their children attending the gifted program. The parents viewed this as a means of learning about what the school expected, what courses were appropriate for their children, and who were the excellent teachers. This also gave them the opportunity to voice their opinions about gifted programs. After three years, with the parents' persistence, a calculus course was added to the curriculum. In some schools, the initiation of a gifted program was started because of vocal parents. Many administrators were supportive of their efforts and expanded existing programs. One parent was proud to state that her son was given a standing ovation when honored for achievement in high school.

Conclusion

Parents Speak Out

Parents expressed their opinions on topics that address the importance for the development of their children's talent. They recommend developing supportive social structures where learning is valued. The Olympiad parents recommend creating an intellectual atmosphere in the family and urge this orientation be extended to the child's peer group. They believe that schools should offer gifted programs to assist families who do not know how to deal with the giftedness of their children.

The parents stressed that gifted children need free time to pursue their own interests at their own pace. They did not smother the child. It is important not to structure every minute of the child's day. It is during these down times that creativity blossoms.

They also emphasized that every child, not only gifted children, needs to be read to in the beginning years. Children should be brought up to be active participants in their activities rather than expecting to be entertained. All children need a learning environment rich in stimuli for intellectual and moral growth. They need to investigate the properties of materials, observe nature, play word games, engage in conversations, and do things with adults. Children need exposure to museums, concerts, play, diversified vacations, age appropriate television, and other stimulating activities. There should be a multitude of material, more than they can absorb. Children need to be exposed to a variety of experiences so that they can find out what they like and enjoy doing. In this way they will be challenged to extend

themselves, ask questions, and investigate more deeply into related topics.

The Olympiad parents regard eating meals with the family as important. This is a time to express values, expectations regarding school and church, and alternatives.

Teachers

Interested teachers can make a world of difference. They can provide self-confidence and cultivate curiosity. When dealing with exceptional students, such as Olympians, teachers should not be afraid to say they do not know something. They should take that opportunity to learn with their students. Teachers should have rigorous standards, demand the most from their students, and recognize, respect, and develop each student's strengths.

References

American Chemical Society (1999). *16th summary report September 1999 United Stated National Chemistry Olympiad.* Washington, DC: ACS Education Division.

Campbell, J. R., (1996). Developing cross-national instruments: Using cross-national methods and procedures. *International Journal of Educational Research, 25*(6), 485-496.

Simpson, R. D., & Oliver, J. S. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science Education, 74*(1), 1-18.

Verna, M. A., Campbell, J. R., & Feng, A. X. (2002). Chemistry Olympians' academic development and productivity. *Tempo, 22*(3), 1, 15-22.

Wu, W. T., & Chen, J. D. (1999). *A follow up study of Taiwan physics and chemistry Olympians: The environmental influences.* Paper presented at the 13th World Conference of the World Council for Gifted and Talented Children, Istanbul, Turkey.

Marilyn Ann Verna is an Assistant Professor at St. Francis College, Brooklyn, New York. She is known for her research with gifted students.

Annie Xeumei Feng is a researcher at the College of William and Mary, Williamsburg, Virginia.

Author Note

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

Korean Gifted Girls and Boys: What Influenced Them to be Olympians and non-Olympians?

Seokhee Cho
and
Hye-Joo Lee
Research Center for the Gifted and Talented
Korean Educational Development Institute

This article summarizes three interrelated studies. The first examines data from 27 Korean Olympians and their parents; the second is a qualitative study of female Olympians and equally capable females; the third study deals with parents of gifted Korean children (152 boys; 168 girls). The results of these studies show that cultural beliefs, values, and orientations largely influence Korean parents. The parents operate within gender-appropriate frameworks, apply gender stereotypic beliefs, and have gender-typed academic areas. Most parents want their "pretty little girls" to become proficient in gender-safe appropriate areas such as music, art, or the humanities. They do not want their talented daughters to pursue careers in math, computers, and science. We make recommendations to rectify these inequities.

Introduction

There are very few Korean female Math and Science Olympians. One hundred thirty eight Korean Olympians have been awarded 179 medals since 1988. Among them there were only 9 females, all of whom were graduates of the Science High Schools (SHS). These schools are self-contained special schools for the scientifically gifted. There are currently 16 SHS in Korea with a total enrollment of 1,500 students. The percentage of female students in these elite schools has increased steadily from 0.0% (1983-1987) to 0.9% in 1988 to 35.4% in 2001. Correspondingly, the number of female Olympians increased from 1 in 1995 to 3 in 2000. Even though the number of female Olympians has been increased, still it is too small to represent the number of female students with extraordinary talent in math and science.

There are several variables that have been found to effect math and science achievement. Regarding the training of scientists at SHS, Min's (1999) multiple regression analyses showed that female students' self-efficacy was influenced the most by their parents' encouragement (B=.218) and initiative to develop their talents in math and science (B =.204). Meanwhile, male students' confidence was determined most by their liking of physics (B=.264) and by their teachers' encouragement (B=.075).

This study focuses on the beliefs, attitudes, and nurturing behaviors that parents and teachers use with the female students who participate in math and science competitions. Some of these competitions lead directly to technical careers. Three studies were carried out. The first was conducted to find out if the parents and teachers of the Olympians used positive or negative beliefs, attitudes, and nurturing behaviors differentially. The second study was done to compare the early development of two of the female Olympians with two equally talented female non-Olympians. In the third study the parents of young gifted boys and girls were surveyed about their gender specific beliefs and nurturing behaviors in the math and science areas.

Review of Literature

There are several variables that influence the achievement of elite male and female students. They include biological variables, personality variables, educational and social variables, and cognitive variables (Campbell & Beaudry, 1998). Among these variables our study focused on the social and educational variables that we believe are related to gender inequities. We examined parents and teachers beliefs and attitudes because we viewed them as central. The fact that gender differences in math and science are decreasing these days is clear evidence that they are not due to biological causes, rather it is more likely that they originate from social and psychological causes. Indeed we believe it is a question of nurture and not nature; that is, gender differences are not a matter of females lack of capabilities but of a lack opportunities (Rebhorn & Miles,

1999).

In view of this, parental beliefs and expectations for children may play a critical role in determining children's attitudes toward math or science (Campbell, Wagner, & Walberg, 2000). Parental gender stereotypic beliefs may play a vital role. Two lines of reasoning may be considered. First, stereotyping gender roles gives males advantages by gender typing mathematics and science. Parents act along gender-appropriate or non appropriate lines and in so doing help to create gender stereotypes. They do this by initially providing gender-appropriate toys and then by encouraging behaviors that fit these roles.

Second, parents have lower expectations for a girl's achievement in these areas. This low expectation downplays a girl's interests in these subjects. These processes are reversed for boys. Parents encourage their sons to excel in mathematics and science and then to exhort them to select careers in these areas. Expertise is typically not a static condition (endstate). To maintain proficiency in an area, one must undergo a process of continual development (Sternberg, 2000); therefore, even gifted individuals who have talent continually need to develop more expertise. If they do not, they stop being recognized as gifted, or they become gifted "has-beens." Consequently, gifted children must be encouraged to study hard in order to build records of achievement even from the beginning of their academic careers. Finally, parents of female gifted children do not want their daughters compared to gifted males in these academic areas.

Interaction between teachers and students in classrooms also plays an important role in determining students' attitude and achievement in math and science. There are several studies dealing with the interaction of teachers with male and female students (Eccles, 1982; Fennema, 1983). These studies show that male students receive more attention from teachers. In other words, teachers have more interaction with male students, especially more superior male students. Male students are also more active in learning situations where answers are expected. As students get older, there is a change in the way teachers deal with male and female students. In elementary school, female students receive more attention from teachers. On the contrary, in high schools, male students get more attention from teachers. Stage, Krienberg, Eccles and Becker (1985) finds that male students get more questions, feedback, individualized instruction, and encouragement. Lafrance (1991) suggests that these classroom interactions cause female students to keep quiet and become inactive in classroom situations.

Study I: Beliefs and Attitudes of Olympians' Parents and Teachers

Purpose

The purpose of the first study was to find out to what extent teachers and parents are responsible for the limited number of Korean female Olympians. The research questions are as follows:

1. Do parents' beliefs and attitudes about girls' talent in math and science contribute to the small number of female Olympians? Do the parents of male and female Olympians differ in their beliefs and attitudes about how extraordinary talent is developed in these areas? If they are different, how are they different?

2. Are teachers' beliefs and attitudes about girls' talent in math and science responsible for the lack of recognition or encouragement given to talented Korean females?

Methods

The Republic of Korea has produced 138 Olympians from 1988-2001 including 9 female Olympians. However, only 75 Olympians had "valid" addresses because some of the students are fulfilling their military service in the army or are abroad for further studies. Furthermore, in the 13 years since the inception of the program, many of the addresses are no longer valid. As in all of the national studies, the mobility of this talented sample made it very difficult to locate the Olympians. Two questionnaires were utilized: one for the Olympians and another for their parents. Both instrument packets were mailed to the Olympians. We received data back from 27 Olympians (mean age 21.2 years; the youngest 17 and the oldest 31 years old) including 4 female Olympians. Gender differences were analyzed with the SPSS program. Five Olympians had earned their doctoral degrees, and 22 are still working on their doctoral degrees. Out of 27 respondents, only three were regular high school graduates and the rest were Science High School graduates. All the female respondents were SHS graduates. Parents were found to be above the middle level in terms of socio-economic status (education, career, and monthly income).

Results

1. **Most helpful experiences:** Regarding the most helpful experiences in their development, the male and female Olympians responded that self-learning through books was the most helpful experience, and the second most helpful experience was interaction with friends, alumni, and parents. Both regular and special programs in schools and private tutoring were not viewed as helpful experiences. This response shows that students' self-learning and social support contributed a great deal compared to any other provision by the schools ($X^2=26.543, p<.01$).

2. **Possible reasons for the small number of female**

Olympians: Olympians were asked to select from six explanations and write freely why there were so few female Olympians. Male Olympians' responses were evenly spread across the six explanations. However, all the female Olympians responded that the lack of support or encouragement from teachers, parents, and society were the main reasons for the sparse number of female Olympians. Writing freely, most of the Olympians pointed out that there were few female students at the SHS. Some of the Olympians believed that very few parents would encourage "their pretty little daughters" to become mathematicians or scientists.

3. **Parents' beliefs or nurturing behaviors:** There were no gender differences for the Olympians or their parents in their beliefs about nurturing creativity and inquiry skills or in providing opportunities to exercise inquiry skills and interact with professional scientists.

   However, parents of females had much stronger views about persistence and abilities. Obviously, they had emphasized these issues with their daughters. The parents of males had more neutral beliefs regarding these gender issues. When considering the possibility of having successful careers in science, the females thought that their parents strongly encouraged them to become scientists, whereas the males thought that their parents had neutral beliefs about this issue (t =-5.156, p <.001). The females and their parents both believed that women can become successful scientists, while male Olympian parents expressed a neutral view on this issue (t =-4.85, p <.001).

4. **Teachers' beliefs and nurturing behaviors:** Olympians rated their high school teachers in terms of their beliefs, attitudes, and nurturing behaviors. Since 24 Olympians were graduates from the SHS and 3 from regular high schools, the data on teachers were mostly about those at the SHS. Regarding teachers' belief in the capability of female students, male and female Olympians saw their teachers differently. The female Olympians were negative in their perceptions about their teachers' belief and attitudes regarding their abilities and attitudes (t =2.670, p <.05). They thought that their teachers saw females as less persistent, less capable of doing challenging scientific investigations. Male and female Olympians had different views regarding teachers' expectations about the possibility of their launching successful technical careers (t =3.770, p <.001). The male Olympians believed that their teachers had positive beliefs about the feasibility of their having productive scientific careers (m =4.87), while female Olympians had negative perceptions (m =2.75).

   Regarding teachers' nurturing behavior towards competitions, the males perceived a neutral stance on the part of their teachers; the female students perceived little encouragement from their teachers (t =2.123, p <.05).

5. **Environmental factors that facilitated talent development:** Olympians were asked to check the variables that facilitated the development of talent in math and science. Female Olympians considered such factors as an encouraging atmosphere at home (t =-7.567, p<.001), acceleration (t =-3.063, p <.05), relative's influence (t =-2.569, p <.05), family members as avid readers (t =-2.140, p <.05), and fathers' discovery of talent (t =-2.061, p <.05). Male Olympians considered some specific books that were at home more strongly than female Olympians did (t =2.299, p<.05). These findings show that female Olympians were more influenced by the interaction with other people, especially within their extended family compared to their male peers.

6. **Academic factors that facilitated talent development:** There were significant differences between males and females (m =4.75) in regard to academic factors that facilitated the development of their talents. Females were more apt to believethat their attendance in private schools facilitated development of their talent in math and science, while males responded that their study in private institutions facilitated the development of their talent in math and science, while males (m =1.65) did not (t =-5.406, p <.001). In Korea these private schools pay more attention to each individual and have less students in each class. They also provide more individualized programs and more encouragement and support for female students. The female Olympians used them to prepare for their entrance examinations to SHS because most of the regular junior high schools do not provide special programs for the gifted in science and math.

Study II. Interview with Olympians and Non-Olympians

Purpose

The purpose of the second study was to find out how and why the Olympians and non-Olympians decided to participate or to terminate their participation in competitions like the Olympiad despite having similar capabilities. The research questions were: "Why did you continue to participate or why did you give up participating in these

competitions?"

Methods

Some male and female Olympians were randomly selected and contacted. They were asked to recommend their friends who possessed similar levels of capability but were not Olympians. Four competent individuals were selected and agreed to be interviewed: 1 male Olympian (MO1), 1 female Olympian (FO2), and 2 female non-Olympians (NFO1, NFO2). All were SHS graduates. The interview questions focused on how and when their talents were developed and the educational and environmental conditions that facilitated or hindered this development. Interviews with each person were carried out by 5 doctoral students for 5 hours, tape recorded, transcribed and analyzed qualitatively.

Results

Early identification and providing challenging experiences: Two Olympians (MO1, FO2) emphasized the importance of gaining insights in problem solving in their academic area. They believe that these understandings were critically important in winning this competition. They also mentioned that this knowledge can be acquired through intensive experiences in math and sciences over a long duration of time. The parents of MO1 and FO2 tried to identify their area of talent when they were young. They started participating in national competition when they were in Grades 4 and 5. NFO1 and NFO2 started participating in competitions only in Grade 7. NFO1 enjoyed math and science when she was young, but her parents did not provide any advanced training experiences until Grade 9. By then she was enrolled in one of the SHS schools, and it was the first time she considered math and science seriously. During the preschool years, NFO2 went for training in English and in music (violin and piano). All of these experiences were generally considered appropriate for girls. In Grade 9 she decided to continue her study in math because it was easy for her. However, she did not know that she had special talent in this discipline because she had only participated once in a national competition.

Continuous provision of extraordinarily challenging experiences: FO2's mother was very active in providing challenging experiences for her daughter in computer science. For example, her mother searched for a university student who had already been one of the Olympians in computer science so that he could be a tutor for her daughter in Grade 7. She continued to attend the same elementary school even after her family moved to an adjacent city. She remained enrolled in this elementary school because it offered challenging programs in computer programming. In middle school she continued to participate in computer programming activities and competitions every year. This accumulated training helped her to be proficient in computer programming.

Keeping balance between schoolwork and the preparation for the Olympiads: In SHS, NFO1 and NFO2 abandoned participating in Olympiad competitions despite encouragement by their teachers. They stopped competing because they realized that they would need to sacrifice their school work and thereby get lower grades. They knew that to do well in the competitions they would need to concentrate on preparing. They believed they could not do both jobs. On the contrary, MO1 and FO2 who started competing when they were young did not have to sacrifice their school work. They were able to keep up with their regular school assignments and still prepare for the competitions. In short, those students who started in Grade 4 could keep the balance between school work and preparing for the Olympiad. NFO1 decided to give up participating in the Olympiad because she did not want to take the risk of being rejected from the universities because of low GPA. Furthermore, if she did not win a place on the Olympic team her acceptance to a university could be jeopardized. FO2 did not need to invest too much time in preparing for the competition. That enabled her to continue to participate in the Olympiad.

Teachers' belief in female students' lack of perseverance: There were some female students in most of the Olympiad training camps. FO2 was one of the best in computer programming at SHS. However, her teachers did not recommend her for the National Olympiad training camp. Her Olympian tutor was also a SHS alumni. He strongly recommended that she be nominated. Finally she was accepted to participate in the Olympiad in Grade 10 and was awarded a Silver medal. FO2 reasoned her teachers believed that girls did not have the perseverance or the needed ability to deal with such challenging projects; therefore, they did not recommend girls for the Olympiad training camp.

Female students' lack of a network: Female students had difficulty in obtaining information about the Olympiad program. NFO1 needed specific information about the competition, but her male peers at the library and dormitory would not give it to her. In the SHS boarding school, there were separate dormitory buildings for male and female students. Actually, a network of information about the training camp and the competition existed within the boys' dormitory. Many decisions were made in the male students' dormitory at night. Sometimes this information was delivered to the female students the next day, and sometimes it was withheld entirely. As a minority in SHS, female students had difficulty getting information from this male network. MO1 pointed out the fact that significant information regarding the Olympiad was shared only among male students and was passed down from one year to the next primarily among the male students. Female students did not have such a network since there were no female Olympians in the beginning and very few female Olympians in any academic area.

Study III. Beliefs and Nurturing Behaviors of Parents of Young Gifted Children

Purpose

The purpose of the third study was to discover how parents perceive their young children in terms of talent development. The research questions are as follows: Do parents have gender specific beliefs in their child's talent area regardless of their child's aptitude or IQs? Do parents of young gifted children have gender specific expectations for their children's future career? Do parents of young gifted children differ in encouraging their children to participate in competitions?

Methods

Questionnaires were sent to parents of 153 gifted boys and 168 gifted girls (ages 3 to 11). The children were identified as gifted by professional school psychologists on the basis of IQ tests (Wechsler Intelligence Scales: WISC-R or WPPSI) and creative problem solving tests in math, science, language arts, and social studies. The mean IQ's of the gifted children was 139.15. There were no significant differences between boys and girls for any IQ or other measure including parents' SES (education, career, and monthly income).

Results

Parents had gender specific beliefs and perceptions about their children's aptitude and future careers: For careers, parents significantly preferred gender appropriate careers. The parents of boys wanted them to pursue careers in science ($X^2 = 31.337, p < .001$) or computers ($X^2 = 5.333, p < .05$). The parents of girls wanted their daughters to work in literature ($X^2 = 14.727, p < .001$), art ($X^2 = 7.364, p < .01$), and the humanities ($X^2 = 5.818, p < .05$). In terms of children's aptitude area, significantly more boys' parents believed that they were talented in math ($X^2 = 11.521, p < .01$) and science ($X^2 = 11.919, p < .01$). The girls' parents saw their daughters' talent in language arts ($X^2 = 25.583, p < .001$). On the basis of this information we conclude that these parents aligned their expectations along cultural lines and not on the basis of aptitude score information.

Parents' effort to identify and develop talents: We asked the parents to consider whether they would work to develop their children's talent in the following areas: reading, English, science, math, sports. The parents of the girls wanted to develop their talents in English and reading. The boys' parents wanted to put their effort into identifying and developing talents in science and math. Significantly, more parents of boys expressed the highest priority for tutoring in science ($X^2 = 5.556, p < .05$). Meanwhile significantly more parents of girls expressed the highest priority of tutoring in music ($X^2 = 9.800, p < .01$) and art ($X^2 = 4.263, p < .05$).

Parents' provisions for maximizing their children's capabilities: Parents were asked to indicate how much support they would provide to develop their child's potential (from 1-no support; to 6 -strong support). Boys' ($m = 4.13$) and girls' ($m = 4.52$) parents were not significantly different from each other in their degree of support for the development of their children's potential.

Parents' perception about their children's interest in competitions: Parents were asked to estimate their children's interest in competing (from 1 -Strong interest; to 5 -No interest). Both parents of boys and girls responded that their children were interested at the average level or not very interested in competing. However, parents of boys thought that their gifted boys were more interested in competing ($m = 3.75$) than the parents of gifted girls ($m = 4.16$) ($t = -2.078, p < .05$).

Parents' interest in children's scoring high in competitions: Parents were asked how strongly they were interested in their children's getting high scores in competitions (from 1 -Strong interest; to 5 -No interest). All the parents responded that they were interested in their children scoring at the average level in such competitions. There were no significant differences between parents of boys and girls.

Implications

This article addresses the reasons for the scarcity of Korean females in mathematics and science, first among the academic Olympians and, by implication, among career scientists. Our focus was on national high school competitions in these areas. The variables considered were parental and teachers' beliefs about females' abilities and aptitudes. We were particularly interested in the nurturing behaviors of parents and teachers.

Based on the three studies reported above, it is possible to conclude that the most logical explanation for the small number of Korean academic Olympians is the parents' gender-specific beliefs about talent and their gender typing of careers in the technical areas. Korean parents encourage their sons or daughters to participate in activities in different fields: daughters more in literature; sons more in math and science. These orientations explain why many parents do not provide challenging mathematics and science experiences or resources for their daughters in middle and high school years. These parental beliefs also explain why so few female students enroll in the SHS schools. Furthermore, our studies show that even girls enrolled in these schools might not be able to participate in academic Olympiad programs because their parents do not provide the needed training during their developing years.

Teachers' beliefs about female abilities are not found to be significantly different from the beliefs about male students except when it comes to participation in competitive activities such as the Olympiad programs. Teachers believe that females would not succeed in such competitions.

To solve these inequities, parents of young gifted girls must break with tradition and recognize mathematical and

scientific talent when it exits. In such cases they should seek ways to develop these talents when their daughters are young. This process will funnel more gifted girls into the Science High Schools. It will also assure that more women become scientists. In this way more Korean females will have the opportunity to make mathematical or scientific contributions. In essence, the academic Olympiad programs are designed for these ends.

References

Campbell, J. R., & Beaudry, J. S. (1998). Gender gap linked to differential socialization for high-achieving senior mathematics students. *Journal of Educational Research, 91*(3), 140-147.

Campbell, J. R., Wagner, H., & Walberg, H. (2000). Academic competitions and programs designed to challenge the exceptionally talented. In K. Heller & F. Monks & R. Sternberg & R. Subotnik (Eds.), *International handbook of giftedness and talent (2nd Ed.)*. Oxford, UK: Elsevier.

Eccles, J., Kaczala, C., & Meece, J. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development, 53*, 322-339.

Fennema, E. (1983). *Final report, research on relationship of spatial visualization and confidence to male/female mathematics achievement in grades 6-8.* Washington, DC: National Science Foundation.

Lafrance, M. (1991). School for scandal: Different educational experiences for females and males. *Gender and Education, 3*(1), 3-13.

Rebhorn, L. S., & Miles, D. D. (1999). High-stakes testing: Barrier to gifted girls in mathematics and science. *School Science and Mathematics, 99*(6), 313-319.

Stage, E., Kreinberg, N., Eccles, J., & Becker, J. (1985). Increasing the participation and achievement of girls and women in Mathematics, Science, and Engineering. In S. Klein (Ed.), *Handbook for achieving sex equity through education*. Baltimore, MD: The Johns Hopkins University Press.

Sternberg, R. J. (2000). Giftedness as developing expertise. In K. Heller & F. Monks & R. Sternberg & R. Subotnik (Eds.), *International handbook of giftedness and talent (2nd Ed.)*. Oxford, UK: Elsevier.

Seokhee Cho and Hye-Joo Lee are both researchers at the Research Center for the Gifted and Talented in the Korean Educational Development Institute, Korea.

Author Note

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

How Do the Theoretical Paradigms Fit the Findings of the International Olympiad Studies?

Sharon O'Connor
SUNY at Westbury

Diversified theories and frameworks have been presented to explain the gender gaps found in mathematics and the technical sciences. This article summarizes the explanations in the following schema: Biological (genetic & hormonal), Nature & Nurture, and Socialization. Within these generalized schema, eight paradigms are presented together with the extensive databases that have been used to support them. Within these frameworks the quantitative and qualitative data from the international Olympiad studies are analyzed to support or refute the different paradigms. These analyses support some of the socialization frameworks. However, consensus emerges that gender gaps remain steadfast in both the United States and several other countries. Consequently researchers need to continue to explore additional variables and measures in order to eradicate gender inequity in the hard sciences.

A critical concern that continues to plague national and international researchers is the steadfast gender gap in mathematical and science achievement at the secondary level (NCES, 2001). Although diversified theories and frameworks have been present since the early 1970's (Benbow & Stanley, 1980; Campbell, 1994; Eccles, 1983; Feingold, 1992; Hedges & Nowell, 1995; Kerr, 2000; Lubinski & Benbow, 1992; 1994; Plomin, 1990; 1997; Stanley, Keating, & Fox, 1974), researchers continue to wrangle with the "array of predictor variables" that influence high mathematics and science achievement – a necessity for productivity (Dewey, 1916). However, consensus exists among twenty-first century researchers that the gender gap continues to endure and "alarmingly" transcends culture and time.

This closing article summarizes the aforementioned International Olympiad Studies within the following schema: Biological (genetic or hormonal differences), Nature & Nurture, and Socialization Factors. Within this generalized schema eight paradigms are presented together with the extensive databases that have been used to support them. Within these frameworks the quantitative and qualitative data from the international Olympiad Studies are analyzed to either support or refute these paradigms.

Deficit Theories

The John Hopkins Research Team, led by Julian Stanley, did groundbreaking work with their studies of mathematically precocious youth and found that young males scored significantly higher than females on the mathematics section of the SAT (Keating & Stanley, 1972; Stanley, 1973; Stanley, Keating, & Fox, 1974). Between 1972 and 1983, the John Hopkins Research Team tested over 65,000 adolescents, specifically 12-13-year olds. Culling their sample from a national search that comprised over 40,000 students, they selected seventh graders (prior age 13) who scored at/or above 700 on the SAT. The results were alarming. The gender disparity was 13 males for every female. Although the John Hopkins Research Team concluded that males dominated mathematical reasoning ability, they also stated the reasons for this disparity were unclear, but they discounted socialization and course preparation influences.

Nature vs. Nurture

Plomin believes the field of genetics research has long neglected the contribution of familial genes to human intelligence, specifically "giftedness" (Plomin, 1977). He further suggests that research geneticists take a more balanced view of the psychological quandary, "*is it nature or is it nurture that influences intelligence.*" Current research (Snyderman & Rothman, 1988) acknowledges the role of inheritance in IQ (Plomin, 1990) and behavior (Zirpoli & Melloy, 2000), and the acknowledgement of the reciprocity and contributing influences of both nature and nurture (Bandura's "Social Learning Theory," 1997) has become widely accepted.

Plomin's (1990) family studies show the average correlation of IQ scores for parents, offspring, and siblings to be $r = .45$. He utilized adoption and twin studies because they provide a natural experiment to disentangle genetic and environmental influences. Interestingly, Plomin found

that each predictor carried similar influence. In the study of related pairs, genetic influence accounted for half (r =.22) the variance. Similarly, in the study of unrelated pairs, family influence (nurture) contributed to half the variance (r =.22); hence he deduced it is a complexity of factors that contribute to IQ.

Variability Extremes

Hedges and Nowell (1995) performed an extensive secondary analysis of six large data sets that analyzed student achievement over a period of 32 years. The data sets include Project Tales, National Longitudinal Study 1972, National Longitudinal School Youth, The High School & Beyond Study, National Educational Longitudinal Study 1988, and The National Assessment of Educational Process. Due to large sample sizes (some in excess of 25,000 cases), effect sizes were used. The results show that all the effect sizes were small except for those associated with vocational aptitude scales (mechanical reasoning, electronics information, and auto and shop information). These differences were profound with 8 to 10 times as many males as females scoring in the top 10%. The researchers attribute this disparity to the high variability of test scores by males. More males than females were in the upper tails of the distribution - ratios of 1:3 to 3:4 in the top 10%, while more females than males were in the lower tails. Hedges and Nowell believe these results elicit mounting concern. Although the gender gaps appear to narrow at times, there is little indication that variance ratios are changing with time.

Males Exhibit Greater Variability at the Extremes

Feingold (1992a, 1992b) attributes gender gaps to the fact that males exhibit more variability than females. He found males to be consistently more variable in quantitative reasoning, spatial visualization, spelling, and general knowledge. He also believes that such variability exists in a number of intellectual traits. However, some of these differences may be due to cultural factors because males may be the extreme in some countries and females the extreme in other countries.

Socialization

The Eccles Model is the most comprehensive theoretical paradigm to date. It is the synthesis of promising decision, achievement, and attribution theorists (Atkinson, 1976; Fennema & Sherman, 1977; Macoby & Jacklin, 1974; Parsons, Kaczala, & Meece, 1982; Weiner, 1972) and is largely based on perceptions of students as they make their own important decisions. Over a four-year period (1975-1979) Eccles conducted a large scale cross-sectional /longitudinal study of the ontogeny of students' achievement beliefs, attitudes, and behaviors. Data were collected from students, parents, and teachers in many forms, such as school records, standardized achievements tests, observations, and questionnaires. Attitudinal scales were factor analyzed and produced three distinct loadings, two of which (subjective task value and self-concept in ability) were significant for females in the regression models. Eccles concluded that student decisions are the product of complex interactions among 11 sets of factors, and each specific decision is not dependent on any one variable but is the result of numerous interactions that are reinforced over time.

Females Camouflage Talent

Kerr (2000) examined how boys and girls are treated at different grade levels. Initially, young girls with exceptional talents are encouraged by parents and teachers to excel. However, at the onset of dating and into adulthood, females with the most talent receive different messages from society. They learn to camouflage their intellectual gifts so that they can be acceptable by their female peers and males who may want to date them.

Work Adjustment

Lubinski, Benbow & Morelock (2000) provide a multifaceted theory to explain the traits of gifted females. A key ingredient is personal preference in the work area termed "job satisfactoriness." Job satisfactoriness includes meeting the females' expectations and utilizing, thus nurturing, her abilities and interests. These interests include outside forces such as lifestyle and family. These researchers found that gifted females have more balanced career goals and equity between marriage, family, and career. Thus Lubinski, Benbow & Morelock's (2000) theory of work adjustment (TWA) is dependent upon "job satisfactoriness" and nurturing these abilities "satisfaction."

Differential Socialization

Campbell's (1994) differential socialization paradigm is a framework for understanding the complexities that underlie gender inequity. After a thorough review of gender literature, Campbell isolated 325 socio-psychological variables that present themselves in multivariate forms. Socio-psychological variables are the products of social and psychological forces such as the family processes, attributions, self-concepts, school-related variables, and so on. Social agents (teachers, parents, other children, and the media) reinforce these inequities. Hundreds of these micro-inequities occur over macro times with numerous social agents (macro social agents) that eventually produce observable gender gaps and gender stereotypes.

Fitting Paradigms to the Data in the Olympiad Studies

With the presentation of the eight paradigms, it is now possible to see how well the data from the international

Olympiad studies fit these theoretical conceptualizations.

Finland Olympians

Although Tirri's research (2001) on Finnish Olympians supports prior research on socialization issues for gender disparities in science, such as inequitable parental, peer, and teacher influences towards females, her case studies of 26 Finnish Olympians support an internal link that Tirri names "inner drive in learning." Tirri states that this common factor, this "inner drive in learning," was evident in the fact that both genders were independent learners and sought intellectual challenges such as the Olympiads, in spite of social inequities.

Her research also supports social biases at home and in school: males receive more encouragement from friends and family to choose a career in science and to choose their professional career early in comparison to their female counterparts. Males also identify early reading and more activities in mathematics and science at home than their counterparts. Although Tirri states that female Olympiads believe their teachers' encouragement was an influential agent, this finding may be due to socialization factors. Gifted women tend to underestimate their abilities (Siegle & Reis, 1998), and gifted males tout their own persona for high achievement (Linn, 1989). Tirri further cites the importance of "partner choice" (a major social factor/influence) in the lives of gifted Olympiads.

American Olympians

Similar results are seen in the Western Hemisphere. Feng, Campbell, and Verna (2001) conducted a qualitative study on 15 American Math, Physics, and Chemistry Olympians and concurred that gender disparities are due to socialization factors. They found that gender inequities were in accordance with socialization factors and cultural stereotypes: female Olympians were underestimating their performance, de-emphasizing their intellect, devaluing and fearing verbal confrontation and/or competitions in mathematics and science, and fearing negative social pressures from peers for excelling in the hard sciences. Similarly the authors found that females perceive bias from their teachers and parents. These Olympians cite a lack of positive female roles in mathematics and science.

German Olympians

The 23 in-depth interviews from the German Olympiad studies (2001) conducted by Lengfelder and Heller find similar findings with the Tirri Finnish Olympiad studies. Although the authors acknowledge the strong socializing factors of a "conducive home environment," many of their Olympiad participants cite the main source of their talent, "motivation, effort, initiative, and curiosity" as genetic. One-third of their participants believes that talent is 60% genetically predetermined. All the participants attribute their successes more to ability than effort.

Lengfelder and Heller also support previous research on social biases at home and at school and that parental and teacher expectations were higher for males than their female counterparts. Furthermore, this male dominated random sample (20 males and 3 females) cited unsystematic selection (categorical social bias) as the main reason for low female participation in the German Olympiads. Most Olympiad participants were convinced that the small rate of female participation was inconsistent with the actual rate of talented females in Germany.

Lengfelder and Heller's interviews also indicated another disparity in terms of gender perceptions of solidarity and competition. Female participants clearly value collective social support and solidarity on a higher threshold than individual competition - unequivocally a socialization factor.

Korean Olympians

Cho's research (2001) on what influences Korean girls and boys to choose to become Olympians was primarily due to socialization factors within the home and school. Although parents nurtured and supported participation in mathematics and science and subsequent competitions (1988-2000), parental perceptions of their children's talent in these disciplines are strikingly different. More than 70% of the parents stated that cognitive talent for their male children was located within the domains of mathematics (32%), language arts (19%), and science (22%), whereas their female children overwhelmingly excelled in the area of language arts (55%) followed by negligible influences of mathematics (12%) and art (6%) talent. Similarly parents encouraged a career in science (44%), law (16%), and the medical sciences (11%) for their sons and supported a career in law (25%), humanities (18%), medical sciences (16%), and literature (10%) for their daughters.

Female participants also reported a lack of early and continuous training in mathematics and science in school and not enough information on the Olympiads. An undeniable fact is the lack of female participation in the Olympiad competitions – only 8 females within a 12-year range. There was wide consensus among the female Olympiads that lack of social support and too few female participants were strong influences in Korean girls preferring a safe path to the university as opposed to competing in Olympiads.

Conclusions

In conclusion, although diversified theories still exist as to "What's in the R^2 for gifted male and female achievement in mathematics and science?" all of our current researchers cite "socialization factors" as the main reason for the dearth of female Olympiads across the globe.

Although some social scientists still believe there are genetic and hormonal differences, such as the "Deficit

Theory" (Stanley, 1973; Benbow & Stanley, 1975), none of the authors within this theme issue concurred with this framework.

A second school of thought claims a combination of nature and nurture (Plomin, 1990;1997; Feingold, 1992a, 1992b; Hedges & Nowell, 1995). In recent qualitative studies, Tirri (2001), Lengfelder & Heller (2001) disclosed a similar reciprocity between the development of talent.

However, all of the Olympiad Studies authors acknowledged socialization factors (Eccles, 1982; Kerr, 2000; Lubinski, Benbow & Morelock, 2000; Campbell, 1994) as the culprit for the gender gap in mathematics and science. Furthermore, the current Olympiad Studies show these gender inequities transcend time and cultures. Studies conducted in the Far East (Cho, 2001), Europe (Lengfelder & Heller, 2001), Scandinavia (Tirri, 2001), and, the United States (Feng, Campbell, & Verna, 2001) evince that negative social milieus and stereotypes consistently reinforce gender inequities.

There is an undeniable consensus among past and present researchers that gender disparities in student achievement in mathematics and science remain steadfast. Present researchers also show that these gender gaps are an international concern. Thus it is incumbent upon twenty-first century researchers to explore additional variables and measures in order to eradicate gender inequity in the hard sciences (O'Connor & Miranda, 2001; O'Connor, Miranda & Beasley, 1999).

References

Atkinson, L. (1976). Update on attitudes and other affective variables in learning mathematics. *Review of Educational Research, 46*, 293-311.

Bandura, A. (1997). *Self-efficacy in changing societies.* Cambridge, UK: Cambridge University Press.

Benbow, C., & Stanley, J. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 1262-1264.

Campbell, J. R., & Beaudry, J. S. (1998). Gender gap linked to differential socialization for high-achieving senior mathematics students. *Journal of Educational Research, 91*(3), 140-147.

Campbell, J. R. E. (1994). Differential socialization in Mathematics achievement: Cross-national and cross-cultural perspectives. *International Journal of Educational Research, 21*(7), 667-743.

Cho, S. (2001). *Korean gifted girls and boys in science and math: What influenced them to be Olympians and non-Olympians?* Paper presented at the Annual Meeting of the American Educational Research Association.

Dewey, J. (1916). *Democracy and education.* New York: Macmillan.

Eccles. (1983). Expectancies, values and academic behavior. In J. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W.H. Freeman.

Eccles, J. (1982). *Sex differences in achievement patterns.* Paper presented at the Annual Meeting of the American Educational Research Association.

Feingold, A. (1992a). The greater male variability controversy. *Review of Educational Research, 62*(1), 89-90.

Feingold, A. (1992b). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research, 62*(1), 61-84.

Feng, A. X., Campbell, J. R., & Verna, M. (2001). *In-depth interviews with American Olympians.* Paper presented at the Annual Meeting of the American Educational Research Association.

Fennema, E., & Sherman, J. (1977). Sex related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal, 14*(1), 51-71.

Hedges, L., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*, 41-45.

Keating, D. P., & Stanley, J. C. (1972). Extreme measures for the mathematically gifted in mathematics and science. *Educational Researcher, 1*(3-7).

Kerr, B. (2000). Guiding gifted girls and young women. In K. Heller & F. Monks & R. Sternberg & R. Subotnik (Eds.), *International handbook of giftedness and talent. (2nd ed.)* (pp. 649-657). Oxford, UK: Pergamon.

Lengfelder, A., & Heller, K. (2001). *German Olympiad Studies: Findings from in-depth interviews.* Paper presented at the Annual Meeting of the American Educational Research Association.

Lubinski, D., Benbow, C., & Morelock, M. (2000). Gender differences in engineering and the physical sciences among the gifted: An inorganic-organic distinction. In K. Heller & F. Monks & R. Sternberg & R. Subotnik (Eds.), *International handbook of giftedness and talent. (2nd ed.)* (pp. 633-648). Oxford, UK: Pergamon.

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math/science pipelines. *Current Directions in Psychological Science, 1*, 61-66.

Lubinski, D., & Benbow, C. P. (1994). The Study of Mathematically Precocious Youth: the first three decades of a planned 50-year study of intellectual talent. In R. F. Subotnik & K. D. Arnoldk (Eds.), *Beyond Terman: contemporary longitudinal studies of giftedness and talent* (pp. 255-281). Norwood, NJ: Ablex.

Maccoby, E. (1966). *The development of sex differences.* Stanford, CA: Stanford University Press.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

National Center for Education Statistics. (2001). *Highlights from the third international mathematics and science study-Repeat* (TIMSS-R). (NCES 2001-07). Washington, D.C.: U. S. Government Printing Office.

O'Connor, S., & Miranda, K. (2001). *The linkages among*

family structure, self-concept, effort, and performance on mathematics achievement of American high school students by race. Paper presented at the Annual Meeting of the American Educational Research Association.

O'Connor, S., Miranda, K., & Beasley, M. (1999). *The linkages among the home environment and academic self-concepts of mathematics and science achievement of intact and nonintact family structures of American high school students.* ERIC DOC# 436 314.

Parsons, J., Kaczala, C., & Meece, J. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development, 53*, 322-339.

Plomin, R. (1990). The role of inheritance in behavior. *Science, 248*, 183-188.

Plomin, R. (1997). Genetics and intelligence. In N. Colangelo & G. Davis (Eds.), *Handbook of gifted education (2nd ed.)* (pp. 67-74). Boston, MA: Allyn and Bacon.

Siegle, D., & Reis, S. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly, 42*(1), 39-47.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media, and public policy.* New Brunswick, NJ: Transaction Books.

Stanley, J. C. (1973). Accelerating the educational progress of intellectually gifted youths. *Educational Psychologist, 10*, 133-146.

Stanley, J. C., Keating, D. P., & Fox, L. H. (Eds.). (1974). *Mathematical talent: Discovery, description and development.* Baltimore, MD: John Hopkins University Press.

Tirri, K. (2001). *Actualizing talent in science: Case studies of Finnish Olympians.* Paper presented at the Annual Meeting of the American Educational Research Association.

Weiner, B. (1974). *Achievement motivation and attribution theory.* Morristown, NJ: General Learning Press.

Zirpoli, T. J., & Melloy, K. J. (2000). *Management: Applications for teachers and parents.* Saddle River, NJ: Prentice Hall.

Sharon O'Connor is an Assistant Professor at the State University of New York (SUNY) at Westbury and is known for her research in Germany and with national data bases.

Author Note

You may find further information about the international Olympiad studies at www.olympiadprojects.com.

Teacher Research

Editor: Barbara Beyerbach
SUNY Oswego

With the "No Child Left Behind" legislation, there is a call for all federally funded education programs to be based on 'scientific research'. How 'scientific' is defined in this legislation has been critiqued as being positivist, narrow, and favoring experimental research. Yet within the educational research community, there is increased emphasis on the importance of context, teacher voice, and researcher-practitioner partnerships in addressing concerns in today's schools.

The articles in this section trace the philosophical roots of teacher research and exemplify the knowledge that can be gained when educators systematically investigate questions of practice to improve learning for their students. In the first article, Hendricks reviews the work of Lawrence Stenhouse, illustrating how his work from the 70's and 80's laid the groundwork for teacher research today. She illustrates how many of the questions and struggles Stenhouse raised are still alive today.

The Cuccaro article interrogates the impact of discipline referrals, examining who gets referred, how frequently, and with what impacts. His work exemplifies the connections between gender, class, special education label and referrals, illustrating clearly that referrals are connected to those attributes. Cuccaro questions the effectiveness of referrals and generates alternatives that might be more effective in meeting the needs of 'these' students.

In the final article, "Improving Novice Teachers' Communication Skills with Urban Families" Peter McDermott and Julia Johnson Rothenberg illustrates how professors can use action research to reflect on and assess the impact of their own teaching on preservice teachers' thinking. The design is rich in that coursework and fieldwork are integrated, and pre/post survey data reveals growth in preteachers' understanding of how to interact with and involve urban families. Much current research on community schools suggests that this is critical in meeting the needs of urban children.

Together these articles provide a rich collage of the work that is being done by practitioners to improve learning for those students most in need. Their work is systematic, recursive, contextualized, and focused on real the real challenges of leaving no child behind.

If you would like to submit a manuscript for this section please contact:

Barbara Beyerbach
Professor
Oswego State University
Dept. of Curriculum & Instruction
100 Poucher Hall
Oswego, NY 13126
(315)341-2650
beyerbac@oswego.edu

A Review of the Work of Lawrence Stenhouse: Questions, Ambiguities, and Possibilities

Cher C. Hendricks
University of West Georgia

This article traces the work of Lawrence Stenhouse, an educator in Britain whose work with teachers in the 70s and 80s positively impacted the action research movement there. Stenhouse's innovative ideas about the role of the teacher in the research process, the nature of authority in educational research, and teachers' emancipation through knowledge can inform our current study of action research in education. Stenhouse's early death, which occurred just as he was making great strides in the study of practitioner research, left many of the questions and ambiguities he was struggling with unresolved. This article describes those questions and ambiguities and offers suggestions

Emancipation through knowledge

My interest in Lawrence Stenhouse, a British educator whose work in the teacher-as-researcher movement was cut short by his early death in 1982, originated when I stumbled across his article entitled "What Counts as Research?" published in 1981 in the *British Journal of Educational Studies*. Having just completed my second semester teaching an action research course for Master's degree students—and against my will at first, I must add, having been trained as a quantitative researcher—I was intrigued by Stenhouse's ideas about the role of the teacher in the research process. Spending two semesters with teachers engaged in research into their own practices was enough to transform my thinking about the nature of research, the role of the practitioner in

Teacher Research

Editor: Barbara Beyerbach
SUNY Oswego

With the "No Child Left Behind" legislation, there is a call for all federally funded education programs to be based on 'scientific research'. How 'scientific' is defined in this legislation has been critiqued as being positivist, narrow, and favoring experimental research. Yet within the educational research community, there is increased emphasis on the importance of context, teacher voice, and researcher-practitioner partnerships in addressing concerns in today's schools.

The articles in this section trace the philosophical roots of teacher research and exemplify the knowledge that can be gained when educators systematically investigate questions of practice to improve learning for their students. In the first article, Hendricks reviews the work of Lawrence Stenhouse, illustrating how his work from the 70's and 80's laid the groundwork for teacher research today. She illustrates how many of the questions and struggles Stenhouse raised are still alive today.

The Cuccaro article interrogates the impact of discipline referrals, examining who gets referred, how frequently, and with what impacts. His work exemplifies the connections between gender, class, special education label and referrals, illustrating clearly that referrals are connected to those attributes. Cuccaro questions the effectiveness of referrals and generates alternatives that might be more effective in meeting the needs of 'these' students.

In the final article, "Improving Novice Teachers' Communication Skills with Urban Families" Peter McDermott and Julia Johnson Rothenberg illustrates how professors can use action research to reflect on and assess the impact of their own teaching on preservice teachers' thinking. The design is rich in that coursework and fieldwork are integrated, and pre/post survey data reveals growth in preteachers' understanding of how to interact with and involve urban families. Much current research on community schools suggests that this is critical in meeting the needs of urban children.

Together these articles provide a rich collage of the work that is being done by practitioners to improve learning for those students most in need. Their work is systematic, recursive, contextualized, and focused on real the real challenges of leaving no child behind.

If you would like to submit a manuscript for this section please contact:

Barbara Beyerbach
Professor
Oswego State University
Dept. of Curriculum & Instruction
100 Poucher Hall
Oswego, NY 13126
(315)341-2650
beyerbac@oswego.edu

A Review of the Work of Lawrence Stenhouse: Questions, Ambiguities, and Possibilities

Cher C. Hendricks
University of West Georgia

This article traces the work of Lawrence Stenhouse, an educator in Britain whose work with teachers in the 70s and 80s positively impacted the action research movement there. Stenhouse's innovative ideas about the role of the teacher in the research process, the nature of authority in educational research, and teachers' emancipation through knowledge can inform our current study of action research in education. Stenhouse's early death, which occurred just as he was making great strides in the study of practitioner research, left many of the questions and ambiguities he was struggling with unresolved. This article describes those questions and ambiguities and offers suggestions

Emancipation through knowledge

My interest in Lawrence Stenhouse, a British educator whose work in the teacher-as-researcher movement was cut short by his early death in 1982, originated when I stumbled across his article entitled "What Counts as Research?" published in 1981 in the *British Journal of Educational Studies*. Having just completed my second semester teaching an action research course for Master's degree students—and against my will at first, I must add, having been trained as a quantitative researcher—I was intrigued by Stenhouse's ideas about the role of the teacher in the research process. Spending two semesters with teachers engaged in research into their own practices was enough to transform my thinking about the nature of research, the role of the practitioner in

the research process, and the ways in which action research puts teachers in charge of their professional development. After reading more works by Stenhouse, I realized that the shift in my thinking mirrored Stenhouse's own journey from the culture of traditional academic research to a culture of classroom modes of inquiry by teachers. The questions he struggled to answer were the questions I, too, was grappling with: How can collaboration occur between academic researchers and teacher researchers so that research is useful to practitioners and adds to the knowledge base? How do we define rigor in action research studies? How do we alter the status of teacher researchers so that they are on equal footing with academic researchers?

Although Stenhouse's work ended just as the recent educational action research movement was beginning, there is much to be learned from his early work with teacher-researchers, particularly for those of us teaching and conducting action research studies in the United States where we still remain several years behind our colleagues from abroad. Stenhouse's work in Britain in the 70s and early 80s has had a lasting impact on the action research movement there, and by considering his experiences with teachers, his struggle to define teacher research, and the questions he left unanswered about the process of teacher research, we can begin to move toward a clear definition of educational action research as we posit answers, based on our experiences in the study and pursuit of action research, to Stenhouse's questions.

Stenhouse and Emancipation through Knowledge

The primary theme in Stenhouse's work was the idea of emancipation—both of students and teachers—through knowledge. Rudduck (1995) explains that as a high school student, Stenhouse became aware of his ability to understand the world around him by considering the knowledge he possessed, exercising his personal judgment, and using various modes of inquiry. This awareness occurred for Stenhouse, Rudduck explains, because he was educated in a system where knowledge and the power associated with knowledge were a right for all rather than a privilege of a few. In this system, Stenhouse was encouraged to use knowledge to make sense of his world so that he could become an authority in his own right. When Stenhouse became a teacher several years later, it quickly became apparent to him that the system he was teaching in was substantially different from the one in which he was educated. The system in which he taught placed teachers in charge of knowledge and thus in a position of power over students (Rudduck). Stenhouse was able to shift this power structure when he began working on the Humanities Curriculum Project (HCP), a funded study that supported secondary humanities teachers working with adolescent students.

In his work with teachers on the Humanities Curriculum Project, Stenhouse challenged the traditional roles of authority in the educational system and he attempted to change the power structure in the educational process so that students had more control over their learning. As Stenhouse worked with HCP teachers, his goal was to help them develop in students the ability to understand controversial moral and intellectual issues that occur due to social conditions and human acts (Stenhouse, 1983a). This goal, however, could not be reached using the standard objectives-based curriculum model, which stressed the attainment of pre-established learning outcomes. Stenhouse (1985a) explained his opposition to objectives-based curriculum stating, "Knowledge cannot be reduced to behaviors. In particular it cannot be expressed in terms of pre-specified performances for it is the function of knowledge, as opposed to mere agglomerations of facts, that it does not determine behavior but liberates it" (p. 77). Rudduck (1988; 1995) suggests that for Stenhouse the objectives-based curriculum model signified academic researchers' lack of trust in classroom teachers because it provided a teacher-proof curriculum. Thus, establishing learning outcomes not only stifled students' ability to use judgment and inquiry methods to advance their understanding but it also diminished teachers' ability to use their experience and professional judgment to help students reach that very goal.

Stenhouse and Research

It was in Stenhouse's work with the Humanities Curriculum Project that he first started to question the role of academic research in improving education. Questioning the power relationship that put teachers in a position of authority over students led to a questioning of the power structure that placed academic researchers, who were influencing the ways teachers taught and students learned, in a position of authority over teachers and schools. Stenhouse believed that studying, developing, and experimenting with curricula was the task of teachers, not academic researchers. Rudduck (1988) provides greater insight on Stenhouse's views stating,

> He saw a curriculum development project not as a convenient means of regimenting teachers in a different set of routines, but as a way of extending their individual and communal power. A curriculum project was not a solution worked out by others and offered to teachers, who had merely to apply it; it was a diagnostic and experimental tool, designed to help teachers examine some of the fundamental problems of schooling. In its framework and materials, a curriculum project gave support for trying out and evaluating new approaches in a spirit of inquiry. Teachers, not curriculum packages, are the agents of change, and the function of curriculum projects is to service the professional learning of teachers by offering specifications teachers can evaluate by testing them in their own classrooms. (pp. 31-32)

Working with teachers on the Humanities Curriculum Project revealed to Stenhouse that not only *could* teachers investigate their own practice but they *should* be in charge of studying what does and does not work in the classroom.

Working from this premise, Stenhouse and his colleagues established the Centre for Applied Research in Education (CARE) after their work on the HCP ended. The Centre's purpose was to make teachers part of the educational research community, which Stenhouse believed was imperative if research were to apply to the practice of education. Over several years, CARE became a place where teachers and academic researchers collaborated on projects that focused on teachers' refinement of their teaching practices (Stenhouse, 1983a). It was during his time working with CARE that Stenhouse's writings became increasingly critical of the academic research tradition. During this period, Stenhouse attempted not only to define practitioner research but he also attempted to find its place in both the academic researchers' and practitioners' worlds. Several ideas emerged in Stenhouse's work between 1978 and 1982, as well as in his writings that were published posthumously, including criticism of the academic research tradition, the role of theory in pedagogy, and academic authority in research.

Stenhouse (1979a) criticized the positivist academic research tradition because it was based on the assumption that teaching and curricular strategies could be tested by applying the same quantitative methods that are used in agriculture experiments. One weakness of applying the positivist research tradition to educational settings was the necessity of having a random sample to increase external validity, a condition Stenhouse (1979a; 1981) explained, which was impossible to achieve in educational studies. A second problem associated with positivist research was the notion of control over context. Stenhouse was hesitant to accept results from traditional academic research studies that began with the phrase, "Other things being equal" because he realized that in real educational settings things are never equal (Stenhouse, 1979b, p. 5). Positivistic research was concerned with establishing general laws and theories, but real classroom situations, Stenhouse (1981) asserted, were too complex for general theories to be useful. Stenhouse (1983a) described the problem with generalization stating, "...few such generalizations take account either of the professional biographical development of teacher and student or the crucial contextual and temporal variables. Hence, at this level of action, research can offer only relatively insecure hypotheses, principles, and theories" (p. 212). Stenhouse made it clear that his argument against positivist research methods in education had less to do with random sampling issues and the use of statistical procedures than it did with its lack of application to the practice of education (see Stenhouse, 1985b).

In Stenhouse's view, removing context from educational research made research irrelevant to teachers and thus to educational practice. The theories developed in traditional academic research, Stenhouse (1981) explained, were based on studies that were difficult to replicate or impossible to apply in the classroom. In addition, it seemed to Stenhouse that there was little real desire by traditional academic researchers to see their theories applied and tested in actual classrooms. Academic research, Stenhouse (1985b) suggested, had been made inaccessible to teachers because it was written in a language laden with jargon that only other academic researchers understood (Stenhouse, 1981). A second issue related to the inaccessibility of academic research relates to the notion of authority. Stenhouse (1981) remarked that academic researchers tended to view teachers as theoretically innocent. Stenhouse believed, though, that teachers merely lacked experience with educational theory and thus the confidence to relate theory to their practice.

The notion of authority in educational research and the structure of power in educational research were of utmost importance to Stenhouse. In fact, one of his most radical ideas, based on his concept of teachers' emancipation through knowledge, was that practitioners need not justify themselves or their actions to academic researchers. Instead, the burden was on academic researchers to justify themselves and their work to practitioners (Stenhouse, 1981). Stenhouse clearly saw the disparity of status between teachers and researchers, and he was concerned that it would prevent both parties from pursuing collaborative studies and engaging in discourse about theory and practice (Stenhouse, 1983a). This problem, Stenhouse speculated, was even more serious in the United States than in Britain because he felt that American researchers held themselves in a higher esteem than did British researchers and they held teachers in a lower esteem. In an address to practitioners, Stenhouse (1985c) appealed to teachers to use academic research to supplement and enrich their judgment, stating that teachers—not researchers—were in the position to critically appraise academic research. This appeal was at the heart of Stenhouse's concept of how education could best be served by research: academic researchers provided theory derived in the traditional academic (statistical) method and teachers tested that theory in the real world context of their classrooms. Stenhouse's position was that both researchers and teachers had to be involved in the research process if there were to be any relationship between educational practice and educational theory, maintaining that, "...practitioners [are] in a position to send theorists back to their drawing boards with confidence. That authority should rest with teachers" (Stenhouse, 1983a, p. 214).

Questions, Ambiguities, and Possibilities

As Kemmis (1995) explains, CARE, under Stenhouse's leadership, became well-known as a place where new ideas emerged that were thought of as creative, innovative, unconventional, radical, and even sacrilegious. Kemmis sees Stenhouse as an important transitional figure who, at the time of his death, was on the verge of major breakthroughs in the study of theory, practice, and research in education. Being in this position, however, also meant that Stenhouse left many unanswered questions. For example, Stenhouse

articulated that educational research should be a collaborative effort between academic researchers and teachers, and although he saw their roles clearly delineated—academic researchers provided educational theories and teachers tested those theories in the context of the classroom—Stenhouse had not resolved how the collaboration would work. A major issue centered on the fact that without training and guidance teachers could not understand the language of traditional research, which made it impossible for them to apply theories in practice. Further, although Stenhouse made it clear that the role of teachers was to test academic researchers' theories and send them back to their drawing boards, he had not determined how this might be accomplished. An obvious solution would have been for teachers to disseminate their work through publication, which Stenhouse (1981) acknowledged in "What Counts as Research" but then suggested in the same article that it was sufficient for teachers to share their work with a critical community who could scrutinize the work and offer suggestions for improvements. The ambiguity here is that if teachers are only disseminating the results of their research to those in their critical community, how can the research truly impact educational theory? It is not possible to send academic researchers back to their drawing boards without disseminating the results of practitioner studies in a way comparable to the dissemination of traditional educational research studies.

Another unresolved issue in Stenhouse's work relates to the actual process of practitioner research. Stenhouse (1981) expressed the need for research to be systematic, planned, and self-critical, but he did not suggest ways this could be achieved. An analysis of Stenhouse's writings indicates that he was struggling with ways to define practitioner research and in fact he seemed to have a difficult time even deciding on terminology to use for practitioner research calling it at times illuminative research (Stenhouse, 1979c), case-study research (Stenhouse, 1979a; 1985d; 1985e), teacher research (Stenhouse, 1981, 1983a), and action research (Stenhouse, 1985f). Rudduck (1995) explains that regardless of the terminology used to describe practitioner research, Stenhouse consistently maintained that its goal was the advancement of teachers' professional understanding of their practice. To reach that goal, Stenhouse suggested that teachers use 'descriptive' research methods that focused on the study of cases rather than the study of samples, and he suggested that the two methods that could be used in this pursuit were ethnographic methods based on participant observation and historical methods based on interviews (Stenhouse, 1979b; 1979c). Stenhouse (1979c) also stressed the need for rigor in practitioner research stating, "I want to see that academic standards are maintained in this new paradigm through the development of adequate academic conventions" (p. 6). However, the standards Stenhouse suggested were limited, focusing only on practitioners making their data available to others for verification.

Practitioner research is still referred to in a number of ways— teacher self-study, teacher-as-researcher, and a variety of labels associated with action research (collaborative action research, educational action research, classroom action research, etc.). And though it is not necessary to limit ourselves to one term when referring to practitioner research, it is necessary to define and describe standards for its practice. Standards must go beyond the delineation of limited methods for data collection and the provision that data must be made available for verification, a conclusion Stenhouse would have come to if his work had continued. Theories of educational action research have come a long way in the years since Stenhouse's work with teachers. Many of the standards established in qualitative research methodology, such as collecting multiple forms of data and triangulating data sources to increase validity and credibility, have been applied to action research. Also, there seems to be agreement on the basic cycle of the action research process, which includes reflection to identify a research focus, generating research questions, collecting and analyzing data, and then continuing in the process of ongoing action planning and reflection.

Still, there remains some ambiguity about the role of theory in the action research process, an area where discourse leading to clarity is needed. I, like Stenhouse, believe that a knowledge and understanding of theory are critical when conducting action research studies, and so the teacher researchers I supervise must complete a review of the literature on their topics even before they formulate research questions. Further, whereas Stenhouse suggested that teachers share their research results with a critical community, I encourage teachers to submit their work for publication or presentation because I believe that action research conducted by teachers in the context of real classrooms can substantially contribute to the educational knowledge base. I am discouraged, however, by the trend of many journals to solicit research by teachers but then only publish collaborative research conducted by academic researchers and teachers that is so full of academic jargon that it fails to be useful to most teachers. Research that claims to be action research but contains p-values and t-tests confirms first that there are still many in education who don't know what action research really is, and second that academic researchers are continuing to use jargon to maintain their positions of authority in educational research. I see this as the teacher being brought into the world of the academic researchers "on a temporary ticket" (Rudduck, 1988, p. 36), and this kind of unnatural collaboration does not serve teachers or help them study their teaching in meaningful or sustainable ways.

Of course, it may be that published teacher research doesn't really look like teacher research for a number of reasons. Journal editors may not be receiving enough submissions from teachers, and if this is the case, the burden falls on those of us who facilitate action research studies to help the teachers we work with in the publication process. Another problem may be that the quality of teacher research is so poor that journal editors are reluctant to publish the

studies they receive from teachers. This situation may be exacerbated by the fact that we have yet to establish firm standards for action research. This is an area where academic researchers and teacher researchers need to work together. Keeping in mind what we can take from Stenhouse on teachers' emancipation through knowledge, the responsibility of creating standards should not fall on academic researchers because that means that once again we are placed in a position of authority over teachers. Instead, teachers and academic researchers should work together to establish acceptable standards for action research, focusing perhaps on issues related to validity, credibility, and the role of theory.

The intention of this article was to describe the impact of Lawrence Stenhouse's work on the action research movement. Though Stenhouse's greatest influence occurred in Britain, there is little doubt that its ripples were felt here, too. Many of the action research theorists that have contributed to our practice—Stephen Kemmis, Clem Adelman, Susan Noffke, John Elliot—were influenced by the work of Stenhouse. Great strides have been made in the study and use of action research in education in the years since Stenhouse's death, but one question he asked remains unanswered:

> Could we have an educational science?...could we have a study of educational phenomena which opted neither for the common language of education nor for the language of social science theory, but instead for a theory which related directly to educational practice? Not a sociology, nor a psychology, but a pedagogy? (Stenhouse, 1981, p. 108)

I think that most of us who have been engaged in research with teachers would say that not only is this a possibility, but we are on our way to seeing this happen. Still, teachers and academic researchers must define this educational science and establish its standards focusing on questions such as 1) What is the role of theory in action research? 2) How can the results of studies by teachers be used to inform theory and inform practice? 3) How do we alter the authority structure in educational research so that academic researchers and teacher researchers value and use each other's work? And, of course, there are a number of methodology questions related to the process of action research that could be added to this list. As teachers and academic researchers work together to answer these questions, I hope that we can remember two quotes from the work of Stenhouse. First, "The virtue of humanity is diminished in man when judgement [sic] is overruled by authority" (Stenhouse, 1983b, p. 163). We are best served in our pursuit of a collaborative educational research tradition if we as academic researchers can let go of our ownership and authority in research and share with teachers the responsibility of informing education. This leads to a second important assertion from Stenhouse, "It is teachers who, in the end, will change the world of the classroom by understanding it" (in Rudduck, 1988, p. 35).

References

Kemmis, S. (1995). Some ambiguities in Stenhouse's notion of 'the teacher as researcher': Towards a new resolution. In J. Rudduck (Ed.), *An education that empowers: A collection of lectures in memory of Lawrence Stenhouse* (pp. 73-114). Clevedon, Avon [England]: Multilingual Matters LTD.

Rudduck, J. (1988). Changing the world of the classroom by understanding it: A review of some aspects of the work of Lawrence Stenhouse. *Journal of Curriculum and Supervision, 4,* 30-42.

Rudduck, J. (Ed.). (1995). *An education that empowers: A collection of lectures in memory of Lawrence Stenhouse.* Clevedon, Avon [England]: Multilingual Matters LTD.

Stenhouse, L. (1979a). Using research means doing research. In H. Dahl, A. Lysne, & P. Rand (Eds.), *Spotlight on educational problems* (pp. 71-82). Norway: Oslo University Press.

Stenhouse, L. (1979b). Case study in comparative education: Particularity and generalization. *Comparative Education, 15*(1), 5-10.

Stenhouse, L. (1979c). The problem of standards in illuminative research. *Scottish Educational Review, 11*(1), 5-10.

Stenhouse, L. (1981). What counts as research? *British Journal of Educational Studies, 29,* 103-113.

Stenhouse, L. (1983a). The relevance of practice to theory. *Theory into Practice, 22*(3), 211-215.

Stenhouse, L. (1983b). *Authority, education, and emancipation.* London: Heinemann Educational Books.

Stenhouse, L. (1985a). The objectives model: Some limitations. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 75-79). London: Heinemann Educational Books.

Stenhouse, L. (1985b). The psycho-statistical paradigm and its limitations 2. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 25-30). London: Heinemann Educational Books.

Stenhouse, L. (1985c). Reporting research to teachers: The appeal to professional judgment. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 40-41). London: Heinemann Educational Books.

Stenhouse, L. (1985d). How research can contribute to the improvement of teaching. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 49-51). London: Heinemann Educational Books.

Stenhouse, L. (1985e). The case-study tradition and how case studies apply to practice. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching:*

Readings from the work of Lawrence Stenhouse (pp. 52-55). London: Heinemann Educational Books.

Stenhouse, L. (1985f). Action research and the teacher's responsibility for the educational process. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 56-59). London: Heinemann Educational Books.

Cher Hendricks is an Assistant Professor of Educational Research at the University of West Georgia. Her work focuses on collaborating with teachers and administrators on classroom and school-wide action research studies.

Author Note

Address correspondence to: Cher Hendricks, Assistant Professor, Educational Research, University of West Georgia, Carrollton, GA 30118. Phone: 770.836.4441. E-mail: cchester@westga.edu

---

Discipline Referrals: A Wealth of Information

Carlo Cuccaro
James E. Lanigan School
Fulton, New York

Student discipline problems continue to be a great concern for educators today. This study examined the discipline referrals in a small rural elementary school for an entire school year. An attempt was made to determine types, frequency, and dates of student misbehaviors as well as type and frequency of administrator intervention. Also, relationships between discipline problems and student achievement were considered. The results indicated that the vast majority of discipline referrals were filed on a very small number of students. Also, there seemed to be a strong correlation between discipline problems and special education and remedial education students. In addition, almost without exception, the discipline problems were dealt with in a "reactive" manner. Lastly, there was a substantial "upswing" in referrals filed as the school year progressed. Problem areas as well as implications are discussed.

Introduction

I am a school psychologist in an elementary school of approximately 480 students in grades Kindergarten through 6th. Recently, I had an "eye-opening" experience as I walked into the principal's office one afternoon and was struck by how crowded it was. It was not filled with teachers, staff members, or parents. Rather, it was crowded with students. Two students sat in front of the secretary's desk. Another student sat next to the door marked "Storage" while two other students stood in front of the teachers' mailboxes. Also, I heard a discussion going on inside the principal's office between students and the principal about "good choices" and "consequences for behavior." I didn't have to ask why the students were sent to the office. I knew the answer. They were all sent there for disciplinary reasons.

In this descriptive research study, data regarding student misbehavior, namely written discipline referrals and information from a discipline logbook, were collected and analyzed. The purpose of this study was to develop a precise description of the total number of written discipline referrals, number of students receiving referrals, referral rates per student, and number of referred students who received special education or Academic Intervention Services (AIS). Also, the types and frequency of student misbehaviors and administrative action, as well as the frequency of written referrals by month were described. Lastly, the most frequent locations for behavior problems were determined. It is the hope of this researcher that the results of this study can be shared with teachers and staff to inform them about the "state" of discipline problems within the school. Also, the information may prove useful in developing future professional development opportunities for teachers and staff.

Overview of Relevant Research

School discipline is a major concern for educators, parents and students alike. For too many of us, "a daily, low – level nastiness and disorder turn schools from communities into obstacle courses or even combat zones" (Educational Testing Service, 1998, pg.33). Of course, dealing with the problems and proposing solutions is an ongoing challenge.

One method of dealing with student's misbehavior is the time-honored tradition of sending them to the principal's office. When a student is sent to the principal's office for a disciplinary infraction, the principal must determine the next discipline step for that student. The consequences include, but are not limited to, the principal talking to the student, detention, call to parents and out-of-school suspension. The latter is normally reserved for repeat offenders or serious infractions such as fighting or bringing a weapon to school. Suspension and expulsion appear to be an ineffective way of dealing with this behavior, as they do not appear to be deterrents (Bock, Tapscott, & Savner, 1998). In most cases, suspension is seen as a "break" or "vacation" from school. Also, there's a strong relationship between failing grades and suspension (Bock, Tapscott, & Savner, 1998), which makes student time out of school via suspension and away from instruction all the more troubling.

Another option to out of school suspension is in–school suspension or ISS (Eggleston, 2001). ISS students remain in school but they are isolated from their classmates in a room to work on their regular assignments. Some forms of ISS have been found to be effective. These programs have a

Readings from the work of Lawrence Stenhouse (pp. 52-55). London: Heinemann Educational Books.

Stenhouse, L. (1985f). Action research and the teacher's responsibility for the educational process. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse* (pp. 56-59). London: Heinemann Educational Books.

Cher Hendricks is an Assistant Professor of Educational Research at the University of West Georgia. Her work focuses on collaborating with teachers and administrators on classroom and school-wide action research studies.

Author Note

Address correspondence to: Cher Hendricks, Assistant Professor, Educational Research, University of West Georgia, Carrollton, GA 30118. Phone: 770.836.4441. E-mail: cchester@westga.edu

Discipline Referrals: A Wealth of Information

Carlo Cuccaro
James E. Lanigan School
Fulton, New York

Student discipline problems continue to be a great concern for educators today. This study examined the discipline referrals in a small rural elementary school for an entire school year. An attempt was made to determine types, frequency, and dates of student misbehaviors as well as type and frequency of administrator intervention. Also, relationships between discipline problems and student achievement were considered. The results indicated that the vast majority of discipline referrals were filed on a very small number of students. Also, there seemed to be a strong correlation between discipline problems and special education and remedial education students. In addition, almost without exception, the discipline problems were dealt with in a "reactive" manner. Lastly, there was a substantial "upswing" in referrals filed as the school year progressed. Problem areas as well as implications are discussed.

Introduction

I am a school psychologist in an elementary school of approximately 480 students in grades Kindergarten through 6th. Recently, I had an "eye-opening" experience as I walked into the principal's office one afternoon and was struck by how crowded it was. It was not filled with teachers, staff members, or parents. Rather, it was crowded with students. Two students sat in front of the secretary's desk. Another student sat next to the door marked "Storage" while two other students stood in front of the teachers' mailboxes. Also, I heard a discussion going on inside the principal's office between students and the principal about "good choices" and "consequences for behavior." I didn't have to ask why the students were sent to the office. I knew the answer. They were all sent there for disciplinary reasons.

In this descriptive research study, data regarding student misbehavior, namely written discipline referrals and information from a discipline logbook, were collected and analyzed. The purpose of this study was to develop a precise description of the total number of written discipline referrals, number of students receiving referrals, referral rates per student, and number of referred students who received special education or Academic Intervention Services (AIS). Also, the types and frequency of student misbehaviors and administrative action, as well as the frequency of written referrals by month were described. Lastly, the most frequent locations for behavior problems were determined. It is the hope of this researcher that the results of this study can be shared with teachers and staff to inform them about the "state" of discipline problems within the school. Also, the information may prove useful in developing future professional development opportunities for teachers and staff.

Overview of Relevant Research

School discipline is a major concern for educators, parents and students alike. For too many of us, "a daily, low – level nastiness and disorder turn schools from communities into obstacle courses or even combat zones" (Educational Testing Service, 1998, pg.33). Of course, dealing with the problems and proposing solutions is an ongoing challenge.

One method of dealing with student's misbehavior is the time-honored tradition of sending them to the principal's office. When a student is sent to the principal's office for a disciplinary infraction, the principal must determine the next discipline step for that student. The consequences include, but are not limited to, the principal talking to the student, detention, call to parents and out-of-school suspension. The latter is normally reserved for repeat offenders or serious infractions such as fighting or bringing a weapon to school. Suspension and expulsion appear to be an ineffective way of dealing with this behavior, as they do not appear to be deterrents (Bock, Tapscott, & Savner, 1998). In most cases, suspension is seen as a "break" or "vacation" from school. Also, there's a strong relationship between failing grades and suspension (Bock, Tapscott, & Savner, 1998), which makes student time out of school via suspension and away from instruction all the more troubling.

Another option to out of school suspension is in–school suspension or ISS (Eggleston, 2001). ISS students remain in school but they are isolated from their classmates in a room to work on their regular assignments. Some forms of ISS have been found to be effective. These programs have a

strong foundation, well-trained, effective instructors and an evaluation component (Sheets, 1996). The prevalence of successful ISS programs is unknown. These programs need to be differentiated from ineffective ISS practices such as leaving students largely unsupervised with the occasional "drop-in" visit from the classroom teacher or having a classroom or cafeteria aide sit with the child in more of a "baby sitting" capacity.

Other methods of dealing with students' misbehavior once they arrive in the principal's office include verbal reprimand, calling the child parent(s), and detention. According to Herzog (2000), the use of punishment has limited and negative effects on students including,

- Short-term effectiveness
- Fails to teach prosocial alternative behaviors
- Does little to change student affect and thinking that is underneath the behavior
- Can be harmful to student – educator relationships
- Can produce student resentment and revenge

In almost all schools today, discipline referrals serve to document what the student's misbehavior was as well as how the behavior was dealt with. These discipline referrals can also serve to increase accountability and documentation on the part of teachers and administrators, provide information to parents of the infraction and consequences, and create a "paper-trail" for particular students. The use of discipline referrals to manage and monitor disrupted behavior in schools is documented in the research (Sugai, Sprague, Horner, Walker, 2000).

Office discipline referrals represent an event in which (a) a student engaged in the behavior that violated a rule or social norm in the school, (b) the problem behavior was observed or identified by a member of the school staff and (c) the event resulted in a consequence delivered by administrative staff who produced a permanent (written) product defining the whole event (Sugai, et al, 2000, p. 96).

Discipline referrals are a barometer, not only of student behavior, but also the discipline systems within the school (Sprague, et. al 2000). Existing research indicates that office discipline referrals can be useful in identifying discipline patterns of students with and without disabilities (Wright & Dusek, 1998), for identifying improvements in school – wide systems (Taylor-Greene, et al, 1997) and staff training (Tobin, Sugai & Colvin, 2000, as cited in Sugai, et al, 2000).

Data Collection

Office discipline referrals were collected for the 2000-2001 school year within the elementary school in which the researcher is employed as a school psychologist. These referrals are standard in many schools and generally contain the student's name, date, type and location of infraction, and disciplinary action taken by administrator. The classroom teachers usually complete these referrals.

Using information from the written referrals, a computer spreadsheet was created with the following categories:

- Date of infraction
- Student name
- Student gender
- Student grade
- Classroom teacher
- Location of infraction (i.e. bus, classroom, lunchroom)
- Referral problem (i.e. insubordination, assault, bus disruption)
- Referral resolution (i.e. call home, counseling, suspension)
- Special Education services received
- Academic Intervention Support services received

Data from each written referral was entered into the spreadsheet and multiple analyses were made including number of referrals per student, number of referrals per month, number of referrals from various locations, frequency of particular misbehaviors, frequency of particular resolutions, correlations between number of referrals and provision of either Special Education or Academic Intervention Support services.

Also, the researcher looked at the discipline logbook that is maintained by office staff. This logbook contained the names, dates and types of infractions for students who were sent to the office, but for whom written referrals were not completed. In general, the students listed in the logbook were sent to the office for misbehavior outside of the classroom (i.e. bus, recess) or in special areas (i.e. music, physical education, art). These infractions in the logbook were in addition to those misbehaviors documented with written discipline referrals. For the 2000 – 2001 school year, however, due to changes in clerical staff within the office, the logbook was not maintained for the entire school year. Instead, it was only kept from 2/01 through 6/01.

Findings

During the 2000 – 2001 school year, there were 218 written discipline referrals filed in the principal's office. Those referrals involved 74 different students. There were approximately 480 students enrolled in the school, indicating that slightly more than 15% of the student body received at least one written discipline referral. Of the 74 students referred, there was great disparity among total number of referrals per student

There were 39 students who received written referrals one time each. These students made up 53 % of the overall students referred, however, their referrals represented only 18% of the total referrals. On the other hand, there were nine students who received seven or more each written referrals. This group of students made up only 12% of the students referred, yet, their referrals (95) represented 44% of the total written referrals.

The cumulative folders of children who received 7 or more written referrals were reviewed to determine how many of these students received Special Education Services or Academic Intervention Services (AIS). Of the 9 students

Table 1

Type and Frequency of Misbehaviors

Misbehavior	Number of Reported Occurrences
Insubordination/class disruption (i.e. refusing to follow teacher directions, shouting out in class)	85
Bus disruption (i.e. refusing to remain seated, shouting on the bus)	78
Assault of a student/fighting	26
Harassment of a student (i.e. name calling, verbal threats)	14
Leaving class without permission	6
Stealing	4
Assault/harassing staff	3
Possession of a weapon	2
Total	218

who received seven or more written referrals, 6 were classified as special education students while 2 received AIS services. This indicated that almost 90% of students who were frequently referred to the office for behavior problems had concomitant academic difficulties.

Of the 218 written discipline referrals analyzed, there was a wide range of infractions that resulted in the student being sent to the office in the first place. The most frequently reported misbehaviors are reported below (See Table 1):

After the students are sent to the office, it is up to the principal to decide on a disciplinary action. In many cases, more than one disciplinary action was taken. For example, the principal might talk to the student about the problem as well as call the child's parent(s). The frequency of disciplinary actions for the 2000-2001 school year is reported in Table 2.

Written referrals were grouped according to date of infraction. The number of written referrals ranged from 7 during the month of September to 47 during May.

Of the 218 written referrals, 119 (54%) were for misbehavior in the classroom, special areas (i.e. art, physical education, music) or hallway. Seventy-eight written referrals (36%) were for bus misbehavior. Twenty-one referrals (10%) were from the cafeteria/recess.

For many children who were sent to the office, there were no written discipline referrals completed. Instead, they were simply sent to the office for misbehavior. For these students, upon their arrival, or shortly after they met with the principal, the clerical staff recorded the student's name, misbehavior, referral source and consequence in a "Discipline Logbook." During the five-month period that the logbook was maintained by office staff (2/01-6/01), there were 185 behavioral incidences requiring administrative intervention. During that same period, there were 144 written referrals filed in the office. That means that during the five months between February and June 2001, there were 329 documented discipline problems (144 written referrals + 185 entries in the discipline logbook). Considering that there were approximately 84 school days during that period, that means that on average, 4 students were sent to the principal's office per day. By breaking these discipline problems down by location, we see that, *for the months of February through June 2001 alone*, there were a *total* of 180 referrals for in-class behavior (55%), 54 for bus problems (16%) and 95 for cafeteria/recess (29%).

Conclusions

This descriptive research study took place in a small rural elementary school in the northeast. The school has approximately 480 students in grades kindergarten through 6th. In this study, data regarding student misbehavior, namely written referrals and information from a discipline logbook, were collected and analyzed in an attempt to describe the types and frequency of discipline problems, most common locations of misbehaviors, types and frequency of administrative action, etc. The findings ranged from obvious to thought provoking to surprising.

First, this school is not unlike elementary schools across the country that are dealing with problems of student insubordination, disruptions in the classroom, cafeteria, and on the bus, and verbal and physical aggression. On a daily basis, students are sent to the office for behaviors like answering teachers back, failing to comply with teacher requests, shouting out in class, refusing to remain seated on the school bus, pushing classmates at recess, and fighting. For a small number of these students, it did seem that a visit to the principal's office served as a deterrent. After all, there were 39 students who were referred to the principal's office only one time each. However, for students with seven or more discipline referrals, being sent to the office seemed to

Table 2

Type and Frequency of Administrative Disciplinary Actions

Discipline Action	Frequency
Discuss problem/"counseling"	172
Office time-out	80
Call home	47
In-school suspension	25
Bus suspension	15
Out of school suspension	12
Parent conference	12
Total	363

have little, if any, impact. Students with chronic behavior problems need more than being sent to the principal's office again and again. These students require highly individualized and targeted support such as extra academic support and adult attention, increased access to rewards, and a functional behavioral assessment (Sugai, et al, 2000). Also, the fact that those students who received either special education or academic intervention support (AIS) services were referred most often to the principal's office was not surprising. Students who have learning difficulties also often (but certainly not always) present behavioral challenges to teachers including inattention, poor work completion, hyperactivity and insubordination. At this school, there seems to be a rather strong correlation between academic difficulties and student discipline problems. In light of this, one cannot help but wonder that if student misbehavior(s) increases as academic difficulties and frustration increases, what would happen if we supported student learning in ways to decrease frustration and increase the students' sense of competence?

One of the more compelling findings in this project was the difference in referral rates throughout the school year. Why, for example, were there 15 written discipline referrals for the first two months of the year, yet 79 referrals during April and May? It does not seem that student "Spring Fever" alone would account for a greater than five-fold increase in referrals. The difference might be due to a combination of factors including increased outside activities (due to warmer weather) during recess and gym, greater assumption of lead teacher responsibilities by student teachers and decreased tolerance on the part of adults as the school year "winds down."

One of the more significant findings in this project was the realization that two separate methods of documenting discipline problems existed at this school. First, there is the "formal" referral involving completing a written discipline form. For the most part, classroom teachers completed these. In addition, there was a discipline log completed by office staff when a student is sent down to the office but there is not a written referral. In these cases, it was generally the adult aides who work in the cafeteria and at recess, or the teachers in the special areas (i.e. art physical education, music) who sent students to the principal's office. Cafeteria and recess in particular seemed to be problematic areas as misbehaviors in these areas accounted for almost 3 out of every 10 students sent to the principal's office.

As a result of this research project, several things have already occurred at this school. The findings were shared with both the faculty and student teachers. Also, the researcher met with cafeteria aides to discuss discipline during lunch and recess and to brainstorm ideas to decrease problem behaviors. Lastly, all students in grades 2 through 6 were surveyed in an attempt to get a better sense of their perceptions of lunch and recess. These surveys have been collected and they will be analyzed.

This research project led to the development of 4 broad goals. First, there should be uniform and consistent method of documenting *all* discipline problems at this school. Second, all staff, including cafeteria and transportation workers, should be provided with professional development opportunities in the area of effectively dealing with student misbehavior(s). Third, students with chronic behavior problems should be identified and alternative strategies of dealing with them should be developed. Finally, a more comprehensive, positive and proactive method of dealing with discipline problems should be implemented.

References

Bock. S.J., Tapscott, K.E. & Savner, J.L., (1998). Suspensions and expulsions: Effective management for students? *Interventions in School Climate and Clinic, 34*(1), 50-52.

Educational Testing Service (2000). Beyond assumptions. In F. Schultz (Ed), *Education: Annual Editions* (pp. 114-117). Connecticut: Dushkin.McGraw-Hill.

Eggleton, T. (2001). *Discipline in schools.* Oklahoma. (ERIC Document Reproduction Service No. ED. 451 554).

Herzog, S (2000, November). Punitive damage: In search of school discipline that helps instead of hurts. *The National PTA Magazine*, 12-14.

Sheets, J (1996). Designing an effective in-school suspension program to change student behavior. *NASSP Bulletin, 80*(579), 86-90.

Sugai, G., Spague, J.R., Horner, R.H., & Walker, H.M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders, 8*(2), 94-101.

Taylor-Greene, S., Browne, D., Nelson, L., Longton, J., Gassman, T., Cohen, J., Swartz,

J., Horner, R.H., Sugai, G., & Hall, S. (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education, 7.* 99-112.

Wright, J.A., & Dusek, J.B. (1998). Research into practice: Compiling school base rates for disruptive behaviors from student disciplinary referral data. *School Psychologist Review, 27*(1), 138-147.

Carlo Cuccaro is a school psychologist at James E. Lanigan School in Fulton, New York.

Improving Novice Teachers' Communication Skills with Urban Families: Professor-Action Research

Peter McDermott & Julia Johnson Rothenberg
The Sage Colleges

Research evidence strongly indicates that family involvement improves children's academic achievement and self-esteem. The authors integrated ideas and strategies for improving novice teachers' communication with families into undergraduate methods courses. The classwork presented basic theories and concepts about the importance of family involvement in children's education. Later specific assignments required the students to communicate and interact with urban families. Using pre/post survey data we found significant growth in students' understanding that schools have the responsibility for initiating and maintaining communication with families; the students also improved in their knowledge of how to conduct family conferences, showed authentic interest in family involvement, and used a variety of strategies for communicating with families. The communication skills of new urban teachers can clearly be improved, but teacher preparation programs must be devoted to it in philosophy, time, and energy.

Teachers in the past assumed that parents supported their efforts in helping children learn - parents contributed to school events, supported teachers by helping with homework and discipline, and regularly attended parent-teacher conferences. But parental support of children's education has become increasingly problematic, especially in low-income schools, with both parents and teachers thinking that the teachers or parents do not care, depending on whom one hears. A good part of this problem is that social and cultural differences between teachers and urban families often serve as boundaries that interfere with effective communication.

For many years the authors have been teacher educators in a liberal arts college with an elementary education program. As such, we have been particularly concerned about helping new teachers become more effective in communicating with low-income, urban families. In this study we report the results of our work to prepare new teachers for better communication with the low-income urban families of the children they will teach.

Studies on family involvement in children's education have consistently indicated its importance in children's academic achievement and self esteem (Comer, 1986; Eccles & Harold, 1996). Comer, Haynes, Joyner, and Ben-Avie (1996) have demonstrated how parent involvement, even in the most poverty stricken schools, can improve a building's climate and children's academic performance. Delpit (1992) has explained that families can serve as cultural informants for interpreting children's classroom behaviors, and this information can be used to improve instruction. After investigating four Texas schools where children scored well on statewide tests, Lopez, Scribner, and Mahitivanichcha (2001) found that these schools made tremendous investments in helping families manage social and economic obstacles in their everyday lives. When the causes of their daily problems were ameliorated, the families became more involved in their children's education and more trusting of teachers and administrators in their children's schools.

Researchers have examined why some families become involved in their children's education and why others do not. McCarthey (2000) found family involvement in education is influenced by a variety of factors including culture, income, language, and the parent(s)' perceptions of school and family responsibilities. Hoover-Dempsey and Sandler (1997) identified three psychological factors contributing to whether parents become actively involved in their children's education: 1) Most important is the family's perceptions of their role and responsibilities in their children's education. Middle class parents feel that they should collaborate with their children's school, but low-income families often feel alienated from public institutions believing schools are responsible for teaching the subject areas and not parents. 2) Parents' feelings of efficacy contribute to their involvement in school; parents who believe they can make a

Eggleton, T. (2001). *Discipline in schools.* Oklahoma. (ERIC Document Reproduction Service No. ED. 451 554).

Herzog, S (2000, November). Punitive damage: In search of school discipline that helps instead of hurts. *The National PTA Magazine*, 12-14.

Sheets, J (1996). Designing an effective in-school suspension program to change student behavior. *NASSP Bulletin, 80*(579), 86-90.

Sugai, G., Spague, J.R., Horner, R.H., & Walker, H.M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders, 8*(2), 94-101.

Taylor-Greene, S., Browne, D., Nelson, L., Longton, J., Gassman, T., Cohen, J., Swartz,

J., Horner, R.H., Sugai, G., & Hall, S. (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education, 7.* 99-112.

Wright, J.A., & Dusek, J.B. (1998). Research into practice: Compiling school base rates for disruptive behaviors from student disciplinary referral data. *School Psychologist Review, 27*(1), 138-147.

Carlo Cuccaro is a school psychologist at James E. Lanigan School in Fulton, New York.

Improving Novice Teachers' Communication Skills with Urban Families: Professor-Action Research

Peter McDermott & Julia Johnson Rothenberg
The Sage Colleges

Research evidence strongly indicates that family involvement improves children's academic achievement and self-esteem. The authors integrated ideas and strategies for improving novice teachers' communication with families into undergraduate methods courses. The classwork presented basic theories and concepts about the importance of family involvement in children's education. Later specific assignments required the students to communicate and interact with urban families. Using pre/post survey data we found significant growth in students' understanding that schools have the responsibility for initiating and maintaining communication with families; the students also improved in their knowledge of how to conduct family conferences, showed authentic interest in family involvement, and used a variety of strategies for communicating with families. The communication skills of new urban teachers can clearly be improved, but teacher preparation programs must be devoted to it in philosophy, time, and energy.

Teachers in the past assumed that parents supported their efforts in helping children learn - parents contributed to school events, supported teachers by helping with homework and discipline, and regularly attended parent-teacher conferences. But parental support of children's education has become increasingly problematic, especially in low-income schools, with both parents and teachers thinking that the teachers or parents do not care, depending on whom one hears. A good part of this problem is that social and cultural differences between teachers and urban families often serve as boundaries that interfere with effective communication.

For many years the authors have been teacher educators in a liberal arts college with an elementary education program. As such, we have been particularly concerned about helping new teachers become more effective in communicating with low-income, urban families. In this study we report the results of our work to prepare new teachers for better communication with the low-income urban families of the children they will teach.

Studies on family involvement in children's education have consistently indicated its importance in children's academic achievement and self esteem (Comer, 1986; Eccles & Harold, 1996). Comer, Haynes, Joyner, and Ben-Avie (1996) have demonstrated how parent involvement, even in the most poverty stricken schools, can improve a building's climate and children's academic performance. Delpit (1992) has explained that families can serve as cultural informants for interpreting children's classroom behaviors, and this information can be used to improve instruction. After investigating four Texas schools where children scored well on statewide tests, Lopez, Scribner, and Mahitivanichcha (2001) found that these schools made tremendous investments in helping families manage social and economic obstacles in their everyday lives. When the causes of their daily problems were ameliorated, the families became more involved in their children's education and more trusting of teachers and administrators in their children's schools.

Researchers have examined why some families become involved in their children's education and why others do not. McCarthey (2000) found family involvement in education is influenced by a variety of factors including culture, income, language, and the parent(s)' perceptions of school and family responsibilities. Hoover-Dempsey and Sandler (1997) identified three psychological factors contributing to whether parents become actively involved in their children's education: 1) Most important is the family's perceptions of their role and responsibilities in their children's education. Middle class parents feel that they should collaborate with their children's school, but low-income families often feel alienated from public institutions believing schools are responsible for teaching the subject areas and not parents. 2) Parents' feelings of efficacy contribute to their involvement in school; parents who believe they can make a

difference in their children's education are more likely to participate in school activities than those who feel ineffective. 3) Finally, some schools are more welcoming than others - the extent to which schools make families feel valued contributes to parent participation.

Although research indicates that low-income, urban schools have great difficulty with family involvement, some of the problems seem reversible, at least from a teacher educator's perspective. Baker, Kessler-Sklar, Piotrokowski, and Parker (1999) discovered that urban teachers often have limited knowledge of what parents do at home to help children in school. Pianta, Cox, Taylor and Early (1999) found that most teacher communication with low-income families consists of "low intensity" letters and flyers with little face-to-face interaction with the parents. Moreover, as the number of African American and Latino children increased in a school, fewer "high intensity" teacher contacts with families ever took place. Linek, Razinski, and Harkins (1997) found that many urban teachers possess a "We -Them" attitude toward urban parents and do not view them as collaborators in the children's education. Valdez (1996) found that even well-meaning educators do not recognize the impact of family beliefs and values about schooling; consequently, many parent education programs designed for Mexican American immigrants do more harm than good because the programs do not build on the families' cultural capital.

There is significant research showing that teachers in low-income schools can learn to become more effective when they understand and respect the social and cultural experiences of the children they teach. Au are Mason (1981) found that when teachers' conversation styles match that of the community, children are more able and eager to participate in classroom activities. Heath (1983) discovered that children will achieve more when their home language patterns and values for literacy resemble that of the school. Cazden (1988) showed that teachers who are familiar with children's conversational styles, including the uses of silence, are more successful in their instruction than teachers who are not. Recently, Nieto (1999) and Bloom, Katz, Slosken, Willet and Wilson-Keenan (2000) have emphasized the importance of establishing respectful and warm social relationships between classroom teachers, their students, and families.

In this study we examined whether we could better prepare our undergraduate students to communicate effectively with low-income urban families. We realized that there would be many hurdles to overcome, not the least of which is that we only worked with our students for one academic year. With few exceptions, our students are young, (around 21 years of age), mostly white and of Western European descent, with little knowledge or experience interacting with African American, Latino, and new immigrant families in our urban community. Moreover, we feared that the students, as has been suggested in other teacher education programs (Cochran-Smith, 2000), may have unconsciously adopted negative stereotypes about low-income families that are so prevalent in the U.S. cultural experience. With these caveats in mind, our methods courses were integrated with concepts and teaching strategies emphasizing the importance of building family involvement in their classroom instruction.

Method

In an important study of teacher research, *Inside Outside: Teacher Research and Knowledge* (1993), Marilyn Cochran-Smith and Susan Lytle studied several major teacher research projects – Philadelphia's project START (Student Teachers as Researching Teachers), the Philadelphia Writing Project, the Prospect Center and School. They began with Shulman's descriptions of two paradigms of research on teaching: process-product research and what Shulman calls studies of "classroom ecology" (Shulman, 1986). They developed a typology of teacher research with two more broad categories: empirical research based on data collection from journals, classrooms, oral inquiries, interviews with appropriate documentation; and conceptual research, which is theoretical and analytical work on life in classrooms or on the research and inquiry itself. The authors distinguish this research from that done *on teachers,* as "subjects of research conducted by university-based scholars." This fine work led us to ask the question of whether professors should do research on ourselves, hopefully being both models for classroom teachers and learners from their research, a community of research/teachers.

It is crucial for professors to study their own teaching. This seems obvious, although probably not common. If we encourage teachers to study their own classrooms for authentic research on real teaching and learning problems, should we not do the same?

Partly as a result of reading the work done by Cochran-Smith and Lytle and others, (e.g. Bullough & Pinnegar, 2001) one of the authors has recently studied and written about the concept of professors studying their own teaching, both as a model and as a research endeavor on its own (Rothenberg, 2001). While we have consistently applied research to portions of our teacher education program (Rothenberg, McDermott, & Gormley, 1999), this study seemed an opportunity for *us* to conduct a teacher-professor action research. As such, we were the professors and we gathered the data from several sources: pre-post Likert scale surveys, student essay assignments, student journals, and mentor teacher reports. In four undergraduate classes over one year, we integrated concepts pertaining to the importance of family involvement and strategies for improving it in urban classrooms.

The subjects are designated *students* in their college classrooms and *novice teachers* in the elementary practica classrooms. Students participated in practica in the local urban school district every morning from 9-11, after which they had their Methods classes, often within the school buildings. These practica allowed them to experience first-hand the extent of family involvement in local urban schools. Thirty undergraduate elementary education majors

participated and provided our data. Three students were African American and two were Latina (16.7%). All of the students participated in four teacher education courses.

In the fall semester we modeled effective conversations with parents, including ways for planning and conducting conferences. Parents *are* their children's first teachers, and we reiterated the value of interactive family involvement throughout many course discussions. All the courses emphasized the inclusion of children with various challenges in regular classrooms of like-aged children. Videotapes were used that presented parents in classroom settings. In Social Studies Methods, students were required to demonstrate skills in working with parents by initiating contact with a family, attending a PTA meeting, and interviewing a parent to discuss their goals for their child's learning in school. In the language arts courses we described and explained the value of home experiences that support children's literacy development.

In the spring semester students were required to complete specific assignments requiring interaction with low-income urban families. We borrowed from the work of Edwards, Pleasants, and Franklin (1999) who recommend that new teachers need to learn about the social and cultural life of urban families, and that one way of doing this is by interviewing urban parents and writing their "stories." Their stories focused on learning activities at home with special emphasis on children's literacy history, parents' beliefs about their children, and the parents' perceptions about their schools. We thought this assignment would help the students develop an understanding of children's learning experiences at home and the extent to which the family's and school's expectations for learning matched. A second assignment required the students to write a newsletter to urban families about their child's first week in school. This assignment was generated from Power's (1999) compendium of strategies for parent involvement. This assignment was developed because students need to develop a strategy for communicating *frequently* with families, especially about good classroom news and activities of their children. A third assignment required the students to design a potential parent involvement plan for including families in their children's literacy learning activities. Although this assignment was hypothetical, we thought that new teachers needed to articulate family involvement plans; if they did not develop plans now during their preservice education, it would be less likely that they would construct one later when they were actually teaching. A fourth and final assignment required the students to attend a community meeting, preferably at a local urban school, and write their reflections about parent participation and contributions to the discussion.

A Likert scale survey asked students to rank the extent to which they agreed or disagreed with 25 propositions about parental involvement. We constructed these survey items first by examining recent papers on family involvement in urban schools (Allexsaht-Snider, 2000; Gunn-Morris, 2000). Then we developed items from the course work designed for impact on our students. The conceptual categories were the following: teacher collaboration with families, strategies for conducting family conferences, making families feel comfortable in school buildings, activities for encouraging family involvement, and overall attitudes and values about working with families and their communities. The survey items elicited students' perceptions about their knowledge and attitudes in working with low-income families (Appendix A contains the questionnaire items).

In September of the academic year (2000-2001) the questionnaire was administered. It was re-administered again in late October and then for a final time at the end of March. The Mann-Whitney U was employed to analyze the responses to these 25 Likert-scale items.

Results

Frequency rankings on the pre and post surveys revealed that the students changed in their self-reportsabout involving parents in children's education. After only eight weeks of instruction we found significant differences in their self-reports on two questionnaire items. The first of these addressed strategies for making parents comfortable in school ($p < .045$). The second pertained to involvement of children in family conferences ($p < .024$).

Table 1 shows the results of analyses from the Mann Whitney U when students' answers on the survey were compared between September and March. The results of this analysis showed differences on seven of the questionnaire items; two of these items were the same as the October results (items 19 and 21 as seen in Table 1). Significant differences were found on the students' knowledge about knowing how to conduct parent conferences, demonstrating authentic strategies for communicating with parents, knowing how to make parents feel comfortable when they come to school, communicating with minority families, involving children in parent conferences, realizing schools' responsibilities for involving families, and knowing community resources to help parents who have children with behavioral difficulties. Table 1provides the specific Mann Whitney U scores for each of these questionnaire items.

Discussion

The study supports aspects of our thinking about teacher preparation - ideas for family involvement can be effectively integrated throughout methods courses. Students in these courses showed significant gains on seven of the 25 items on the survey. They reported improved knowledge and confidence in how to conduct family conferences, display interest in children's families, communicate effectively with urban families, and make parents feel welcomed and valued as partners in their children's learning.

The study corroborates our experience and knowledge about teaching and learning. Years ago as new faculty, the methods courses were quite separate from the practica

experiences novice teachers received in schools. Over the years we gained the knowledge that students learn best by doing. The more teaching strategies from methods courses could be integrated into the practica experiences, the more effective teachers in childhood education would become. Now methods courses are not only integrated with student practica, but the courses are actually held in the urban schools.

Before beginning this project, a major concern was the successful integration of parent involvement ideas and activities into our methods courses. There was already so much our students needed to learn, that we were concerned about overloading our social studies and language arts courses with more assignments and topics. However, integrating parent involvement ideas into social studies and literacy courses, was far easier than we anticipated. The students reacted favorably, too. They enjoyed preparing classroom newsletters for families. Many of them were already highly skilled with computers, and designed layouts for their newsletters or used templates in their word processing programs. Similarly, the undergraduates welcomed the opportunity to interview parents about their children's learning experiences at home. Cooperating teachers shared anecdotes about how much they liked the family newsletters that the student teachers were using. Some even said that they felt it easier to talk with families because of the efforts of our students at improving the flow of communication with parents. Often the activities involving parents were new to the cooperating teachers. It should also be noted that there were no negative reports.

We have learned that knowledge and strategies for communicating with urban families can grow and improve through educational experiences. As with most teacher action research, this study is population specific.

However important these limitations are, data from several sources, including teachers in the schools, indicate strong progress in these students for working with parents effectively. The more barriers that we can remove between schools and the families of children our students teach, the better our learning communities will become.

References

Allexsaht-Snider, M. (2000). *Families, schools and*

Table 1

Items with Significant Change on the Pre/Post Surveys between September and March

Questionnaire Item	U Score	W Score	Z Score	*p* Value
10. I know an effective model for conducting parent conferences (September-March)	204	1332	-4.51	.000
18. I know a variety of strategies for demonstrating authentic interest in a family's education goals for their child. (September-March)	303	1231	-3.22	.001
19. I know how to make parents feel comfortable in my school and classroom. (September-March)	385	1151	-2.20	.027
20. I have strategies for communicating effectively with low-income minority families. (September-March)	303	1234	-3.25	.001
21. I know how to effectively involve children in family conferences. (September-March)	348	1189	-2.67	.0074
22. Schools have the responsibility for effectively involving families in children's education. (September-March)	371	1166	-2.48	.013
25. I am knowledgeable about community resources available to parents of children with behavioral difficulties in my classroom. (September-March)	406	1131	-1.88	.059

communities. Paper presented at the American Educational Research Association meetings, New Orleans LA.

Au, K.H. & Mason, J.M. (1981). Social organizational factors in learning to read: The balance of rights hypothesis. *Reading Research Quarterly, 17*(1), 115-52.

Baker, A., Kessler-Sklar, S., Piotrokowski, C. & Parker, F. (1999). Kindergarten and first-grade teachers' reported knowledge of parents' involvement in their children's education. *The Elementary School Journal, 99*, 367-379.

Bloom, D., Katz, L. Solsken, J. Willet, J. & Wilson-Keenan, J. (2000). Interpellations of family and classroom literacy practices. *Journal of Educational Research, 93*, 155-163.

Bullough, R.V. & Pinnegar,S. (2001). Guidelines for quality in autobiographical forms of self-study research. *Educational Researcher, 30*, 13-21.

Cazden,C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, N.H.: Heinemann Publications.

Cochran-Smith, M. & Lytle, S. (1993). *Inside/Outside: Teacher research and knowledge*. New York: Teachers College Press.

Cochran-Smith, M. (2000). Blind vision: Unlearning racism in teacher education. *Harvard Educational Review, 70*, 157-190.

Comer, J. (1986). Parent participation in the schools. *Phi Delta Kappan, 67*, 442-446.

Comer, J., Haynes, N., Joyner, E., & Ben-Avie, M. (Eds.)(1996). *Rallying the whole village: The Comer process for reforming education*. NY: Teachers College Press.

Delpit, L. (1992). Education in a multicultural society: Our future's greatest challenge. *Journal of Negro Education, 61*, 237-249.

Eccles, J.S., & Harold, R.D. (1996). Family involvement in children's and adolescents' schooling. In A. Booth & J.F. Dunn (Eds.), *Family school links: How do they affect educational outcomes*? (pp. 3-34). Mahwah, NJ: Lawrence Erlbaum.

Gunn-Morris, V. (2000). *Preparing for family involvement in education: A critical need for preservice teachers*. Paper presented at the American Educational Research Association, New Orleans, LA.

Heath, S.B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press.

Hoover-Dempsey, K. & Sandler, H. (1997). Why do parents become involved in their children's education. *Review of Educational Research, 67*, 3-42.

Linek, W., Rasinski, T., & Harkins, D. (1997). Teacher perceptions of parent involvement in literacy education. *Reading Horizons, 38*, 90-106.

Lopez, G., Scribner, J. & Mahitivanichcha, K. (2001). Redefining parental involvement: Lessons from high-performing migrant-impacted schools. *American Educational Research Journal, 38*, 253-289.

McCarthey, S. (2000). Home-school connections: A review of the literature. *The Journal of Educational Research, 93*, 145-153.

McDermott, P., Gormley K., Rothenberg, J., Hammer, J. (1995b) The influence of classroom practica experiences on student teachers' thoughts about teaching. *Journal. of Teacher Education, 46*, 3, 184-191.

Neito, S. (1999). *The light in their eyes: Creating multicultural learning communities*. NY: Teachers College Press.

Pianta, R., Cox, M., Taylor, L. & Early, D. (1999). Kindergarten teachers' practices related to the transition to school: Results of a national survey. *Journal of Teacher Education, 100*, 71-86.

Power, B. (1999). *Parent power: Energizing home-school communication*. Portsmouth, NH.Heinemann.

Rothenberg, J.J. (2001). *What do I know about my own teaching and how do I know it?* . Paper presented at the Northeastern Educational Research Association. Kerhonksen NY.

Rothenberg, J.J., McDermott, P.C., Gormley, K. (1999). Preparations and reflections: Teaching in a wider world. Paper presented at the meetings of the American Educational Research Association, Montreal Quebec, April 1999. *Resources in Education*. ERIC Access No. ED 431 718.

Shulman,L. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M.C. Wittrock (Ed.) *Handbook of research on teaching* (3rd ed.) New York: MacMillan.

Valdez, G. (1996). *Con respecto: Bridging the distances between culturally diverse families and schools: An ethnographic portrait*. NY: Teachers College Press.

Peter McDermott is an Associate Professor in the Education Division of The Sage Colleges where he teaches graduate and undergraduate courses in reading and the language arts. For the last three years he has taught in Kazakstan, Central Asia, where he has been a volunteer with the Open Society Institute and the International Reading Association's Reading and Writing for Critical Thinking Project.

Julia Johnson Rothenberg is Professor of Education at The Sage Colleges, Troy and Albany, New York. She has had long standing research and teaching interests in multicultural teaching and learning, as well as specific interests in teaching in low income schools, both in the United States and abroad.

Methodological Issues

Editors: David Shannon
Auburn University

Isadore Newman
University of Akron

We are pleased to present another section regarding methodological issues. The purpose of this section is to provide an opportunity to learn about and generate discussion about a variety of methodological issues as they pertain to educational researchers. This section includes two articles that address important issues facing educational researchers as they propose to conduct research and determine the appropriate procedures to use for the analysis of the data.

As educational researchers, we most frequently conduct research involving human subjects and must adhere to ethical and legal standards governing such research. The first article, "Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers," written Clements and Myrick, discusses the implications of legislation pertaining to research involving the use of human subjects. Drawing upon their experiences with the Institutional Review Board (IRB) and those they have had as researchers, the authors discuss implications regarding the necessary training and education for educational researchers, impact on the type and quality of educational research and responsibilities of professional organizations and journals promoting educational research.

The second article, "Possible Type VI Errors in Current Research Practices," written by Newman, Frauss, Newman, and Brown, emphasizes the importance of aligning the research procedures with research questions. The authors identify an inconsistency between between the study's research question and the analytic technique and/ or research design used in the study as Type VI errors and illustrate five different practices in which such errors are likely to occur and offer practical and statistical advice for addressing such errors.

We hope these initial articles have been valuable and are looking forward to expanding this section in future issues. To do so, we need your input. What types of issues would you like to explore? We look forward to discussing these issues with you and exploring them in future issues. If you would like to submit a manuscript for this section, or simply discuss an idea, please contact:

David Shannon
Professor
Auburn University
4036 Haley Center – EFLT
Auburn, AL 36849-5221
(334) 844-3071
shanndm@auburn.edu

Isadore Newman
Distinguished Professor
University of Akron
Zook Hall - 424
Akron, OH 44325-4208
(330) 972-6955
isadore@uakron.edu

Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers

Andrea D. Clements and Pat W. Myrick
East Tennessee State University

The recent requirement for education on the protection of human subjects mandated by the federal government has the potential to significantly impact research in all fields, including education. At this point, all key personnel involved in research utilizing human subjects are either required or strongly recommended (depending on their type of institution) to be trained, tested, and certificated, verifying that they have completed education relating to the protection of research participants from harm. A history of the requirement, clarification of when and to whom it applies, and possible implications are discussed. Due to this additional hurdle there is the risk that

Methodological Issues

Editors: David Shannon
Auburn University

Isadore Newman
University of Akron

We are pleased to present another section regarding methodological issues. The purpose of this section is to provide an opportunity to learn about and generate discussion about a variety of methodological issues as they pertain to educational researchers. This section includes two articles that address important issues facing educational researchers as they propose to conduct research and determine the appropriate procedures to use for the analysis of the data.

As educational researchers, we most frequently conduct research involving human subjects and must adhere to ethical and legal standards governing such research. The first article, "Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers," written Clements and Myrick, discusses the implications of legislation pertaining to research involving the use of human subjects. Drawing upon their experiences with the Institutional Review Board (IRB) and those they have had as researchers, the authors discuss implications regarding the necessary training and education for educational researchers, impact on the type and quality of educational research and responsibilities of professional organizations and journals promoting educational research.

The second article, "Possible Type VI Errors in Current Research Practices," written by Newman, Frauss, Newman, and Brown, emphasizes the importance of aligning the research procedures with research questions. The authors identify an inconsistency between between the study's research question and the analytic technique and/ or research design used in the study as Type VI errors and illustrate five different practices in which such errors are likely to occur and offer practical and statistical advice for addressing such errors.

We hope these initial articles have been valuable and are looking forward to expanding this section in future issues. To do so, we need your input. What types of issues would you like to explore? We look forward to discussing these issues with you and exploring them in future issues. If you would like to submit a manuscript for this section, or simply discuss an idea, please contact:

David Shannon
Professor
Auburn University
4036 Haley Center – EFLT
Auburn, AL 36849-5221
(334) 844-3071
shanndm@auburn.edu

Isadore Newman
Distinguished Professor
University of Akron
Zook Hall - 424
Akron, OH 44325-4208
(330) 972-6955
isadore@uakron.edu

Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers

Andrea D. Clements and Pat W. Myrick
East Tennessee State University

The recent requirement for education on the protection of human subjects mandated by the federal government has the potential to significantly impact research in all fields, including education. At this point, all key personnel involved in research utilizing human subjects are either required or strongly recommended (depending on their type of institution) to be trained, tested, and certificated, verifying that they have completed education relating to the protection of research participants from harm. A history of the requirement, clarification of when and to whom it applies, and possible implications are discussed. Due to this additional hurdle there is the risk that

research protocols will be modified to reduce the use of human subjects, or be abandoned altogether. However, potential benefits of the requirement are also outlined.

Requirements for Education on the Protection of Human Research Subjects: Implications for Educational Researchers

On September 17, 1999 Jesse Gelsinger died following complications of a gene therapy experiment conducted at the University of Pennsylvania. Since that day, Mr. Paul Gelsinger, Jesse's father has worked diligently within private and legislative arenas to try to understand how the death of his son could have been prevented. His inquiries led to Congressional testimonies and court appearances seeking justice for not only his son, but for many other research participants who had been harmed while participating in research.

Unfortunately the Gelsinger case is not isolated. History has laid the foundation for where we as researchers find ourselves today. The biomedical sciences have been guilty of a number of atrocities, many widely publicized, though usually after the fact, however, the non-medical arena is not untarnished. Research in the behavioral and social sciences includes such examples as Stanley Milgram's (1994b) obedience experiments. In attempting to demonstrate that individuals will all supposedly obey given the right conditioning, Milgram (1994a) summed up his position by saying, "It is ironic that the virtues of loyalty, discipline and self sacrifice that we value so highly in the individual are the very properties that create destructive organization engines of war and bind men to malevolent systems of authority" (pp. 261-262). While supposedly well intentioned, Milgram's studies are now used as examples of unethical research because of the potential harm they caused his subjects; not physically but mentally.

It is a faint gray line at best, that is drawn between the behavioral and biomedical fields of research.

> Some biomedical research pertains to behavior (e.g., in psychiatry, neurology, or epidemiology), and many of the methods used in behavioral research, such as observation and the questioning of subjects, are also used in biomedical research. Research may be designed to evaluate the behavioral changes that result from a biomedical intervention (e.g., lessening of depression after taking a particular medication or changes in psychiatric disorders following hemodialysis) or to examine physiological responses to behavioral interventions (e.g., lowering of blood pressure through biofeedback or weight loss through hypnosis) (OHRP, 2001).

How short-sighted are the statements — *We are not cutting anyone. There is no risk in simply asking questions.* What are commonly accepted as the *soft risks* can be just as harmful as physical injury. A breach of confidential information, embarrassment due to a connection between a research subject and his or her responses, loss of social standing, or loss of relationships can be psychologically harmful. The fact that subjects can be harmed but not physically injured, or physically injured but not necessarily harmed is true for both biomedical and non-medical fields.

History of the Requirement

The need for protection of human research subjects has been apparent in recent years with well publicized shutdowns of all research at institutions such as The University of Alabama at Birmingham (UAB) in January, 2000 (Hansen & Spencer, 2000), Duke University in April, 2000, and Johns Hopkins in July, 2001 (Curry, 2001) for a period of time due to human subjects protection issues. Individual researchers have also been suspended due to human subjects research violations (Robeznieks, 2001a) as well, some of them quite noteworthy. These cases underscore the oversight and importance being placed on human subjects currently. The progression of federal laws, regulations, and policies relating to education on the protection of human research subjects are shown in Table 1. The purpose of this article is to discuss these regulations and policies and their implications for educational research.

NIH Education Requirement

In October 1999, the Secretary for the Department of Health and Human Services, Donna Shalala, accepted a recommendation made by the Commission on Research Integrity to extend the education requirement for the responsible conduct in research (RCR) to include those conducting federally-supported Public Health Service (PHS) research ("DHHS/NIH announce new initiatives", 2000). Additionally, as a part of the increased vigilance with regard to the treatment of human subjects, the National Institutes of Health (NIH) imposed new regulations for education of individuals involved in human subjects research.

The NIH policy reads:

> Beginning on October 1, 2000, the NIH implemented a policy requiring education on the protection of human research participants for all key personnel submitting NIH applications for grants or proposals for contracts or receiving new or non-competing awards for projects involving human research participants (NIH: Office of Extramural Research, 2000, June 5).

The June 5, 2000 Notice OD-00-039 goes on to define key personnel as "all individuals responsible for the design and conduct of the study." The background for this measure

Table 1

History of Federal Requirements for Education on the Protection of Human Research Subjects

Date	Policy	Requirement	Penalty for Noncompliance
6/13/2000	65 Fed. Reg. 37137	Office for Human Research Projections (OHRP created with three divisions: Policy and Assurance, Compliance Oversight and *Education and Development*	
10/1/2000	NIH Education Requirement Notice OD-00-039 dated June 5, 2000	All key personnel submitting grant applications to NIH	Funding not awarded
8/15/2000	Office of Research Compliance and Assurance {ORCA} IL #01-15 Online Educational Materials: Update	Veterans Health Affairs (VHA) researchers (PIs, Co-PIs, and Co-Investigators) wishing to conduct VHA-funded research must provide documentation of education and/or training obtained over the previous three years on human subjects protection	Research will not be approved and may not be initiated
12/2000	Initiated Federalwide Assurance (FWA)	"research investigators complete appropriate institutional educational training *before* {italics added} conducting human subject research	Not tested
9/5/2001	NIH Education Requirement Clarification Notice OD-01-061	Extended education requirement to key personnel at consortium institutions or performance sites	Funding not awarded

was described as part of a new initiative to "strengthen government oversight of medical research." A later NIH notice (NOT-OD-01-061) clarified that the "education requirement also applies to key personnel at consortium institutions or performance sites if they are participating in research that involves human subjects" (NIH: Office of Extramural Research, 2001, September 5).

Further clarification of the NIH policy stated "All new and non-competing grant applications submitted [to NIH] beginning October 1, 2000, must *describe* [italics added] the education all 'key personnel' have received in the protection of human subjects" (Federal Grants News for Colleges and University, July/August 2000, p. 1). At a meeting of the Council on Governmental Relations (COGR), Dr. Wendy Baldwin, Deputy Director of Extramural Research at NIH, discussed the various announcements noting additionally

that the policy notice required investigators to *describe* [italics added] the completed education before funds would be awarded. Adding to the mounting list of changes, DHHS announced the transfer of the Office for Protection from Research Risks (OPRR) from NIH to a new, *elevated* office, the Office for Human Research Protections (OHRP). Its three divisions, Policy and Assurance, Compliance Oversight and *Education and Development*, would report directly to the Assistant Secretary for Health (65 Fed. Reg. 37137, June 13, 2000).

45 CFR 46

Many researchers may feel comfortable that, because they are not receiving NIH funding or conducting any otherwise federally funded research, they are exempt from the education requirement. However, institutions which receive any type of Department of Health and Human Services (DHHS) support are governed by the DHHS Regulations for the Protection of Human Subjects, commonly known as Title 45 Code of Federal Regulations Part 46 (45 CFR 46), and all research conducted at these institutions is classified as federally-supported by federal regulations. Subpart A of 45 CFR 46 is known as the "Common Rule," and has been adopted by all branches of the federal government involved in human subjects research. It is within these regulations that Institutional Review Boards (IRBs) are mandated for each institution.

Federalwide Assurance (FWA)

Most sizable institutions come under what is known as an assurance. Formerly most had (some still do) a Multiple Project Assurance (MPA) which ensured that all research conducted at that institution would comply with OHRP guidelines. MPAs were written specifically for each institution. In December 2000, OHRP developed what is known as a Federalwide Assurance (FWA), which provides that not only the institution, but, like the MPA, that each individual researcher will do no harm and conduct research following ethical standards. The difference between the FWA and the MPA, is that the FWA is uniform across all institutions. The FWA identifies an institutional official and establishes a partnership between the federal government and the institution to guarantee that all research, not just funded research, will be assessed prior to initiation to verify that human research participants will be protected. While MPAs are still honored, they are converted to FWAs as they expire (usually a term of three years). Educational training is covered in the FWA, which will now be the standard for all institutions with assurances. The wording in the FWA terms is

> OHRP *strongly recommends* [italics added] that the Institution and the designated IRB(s) establish educational training and oversight mechanisms (appropriate to the nature and volume of its research) to ensure that research investigators, IRB members and staff, and other appropriate personnel maintain continuing knowledge of, and comply with, relevant ethical principles, relevant Federal Regulations, OHRP guidance, other applicable guidance, state and local laws, and institutional policies for the protection of human subjects. Furthermore, OHRP recommends that a) IRB members and staff complete relevant educational training *before* [italics added] reviewing human subject research; and b) research investigators complete appropriate institutional educational training *before* [italics added] conducting human subject research (Federalwide Assurance of Protection for Human Subjects, 2002).

For the sake of consistency, institutions conducting federally-supported research and covered under FWAs (or MPAs) review all research the same, regardless of the source of funding or whether, in fact, the project is funded or not. It is for this reason that policies are applied in *blanket-fashion* to all researchers at an institution. The benefits are that this consistency serves to simplify the research protection process by eliminating necessity for determining which research falls under which guidelines and also promotes public trust.

Veteran's Administration Interpretation of the Requirement

While *strongly recommends* is not a mandate, it has been interpreted as such by the Veteran's Administration. In an August 15, 2000 memorandum, the Office of Research and Development (ORD) established a new requirement related to investigator education on human subjects protection for Veterans Health Affairs (VHA) researchers (PIs, Co-PIs, and Co-Investigators) wishing to conduct VHA-funded research. ORD gave January 1, 2001, as a deadline for completion and documentation of education and/or training obtained over the previous three years (Office of Research Compliance and Assurance [ORCA] IL #01-15 Online Educational Materials: Update). Within the same day, John R. Feussner, Chief Research and Development Officer for the Department of Veterans Affairs, issued the same statement and added, "This documentation must be on file at the local office *prior to* [italics added] submission and review of any new proposals..." (J. R. Feussner, personal communication, August 15, 2000, p. 1).

Affiliation with VA research facilities has elevated the education requirement from a strong recommendation to a requirement for many institutions. Our institution and others that are affiliated with the VA are bound by Fuessner's edict for training prior to submitting research proposals. While non VA-affiliated institutions may choose to disregard the *strong recommendation*, in most cases regarding the safeguarding of human subjects we have promoted erring on the side of caution. It would seem that meeting that requirement would be on the side of caution.

Modes of Institutional Compliance

Let us assume for the sake of argument that institutions are bound by the requirement for educating key personnel in human subjects protection, as we consider ourselves to be at our institution. How are institutions meeting the educational requirement for so many researchers? Initially there were education modules online, provided by NIH, which anyone could utilize and be *certificated* as having been trained. These were provided because of the short timeline to have researchers trained (June 5, 2000, the announcement was issued that researchers would have to be trained, and all were to have met that requirement by October 1, 2000). However, it was up to each institution to begin developing their own program for educating investigators and key personnel in the protection of human subjects and be able to verify this for Federal auditors. A variety of methods have been designed including in-house online modules, workshops, and seminars lasting several hours or several days in order to train everyone from students to experienced researchers. Education is necessary, but not sufficient to meet Federal requirements. Researchers and other key personnel must also be evaluated somehow in order to verify that they understand critical factors involved in the protection of human subjects. Our institution, which includes a regional university campus, medical college campus, and VA medical facility, provides a three hour training, two brief online modules, and a written test on which the minimum passing score is 85%.

Research dollars flow in the billions, supporting both non-medical and medical research, yet without human subjects, dollars would be worthless. Therefore it is to the advantage of researchers and institutions to protect subjects both out of ethical obligation as well as to maintain public trust in the research community so that in the future individuals will be willing to participate in research. Education in human subject protection is important not only for the new researcher, but for the experienced investigator as well. An adult serious about receiving a quality education concerned for their future vocation would not enroll in a college or university that did not offer an accredited program. By the same token, it is unlikely that a potential research participant would be willing to engage in research being conducted by researchers who were either employed by an institution or who themselves had failed to *earn* and *maintain* the public trust. Long gone are the days when the directives of an individual whose name is followed by M.D., Ph.D., Ed.D., Sc.D. or any other D., were followed without question. With the failures of human subject research presented alongside media advertisements seeking research enrollees, potential subjects are more informed and perhaps more skeptical than in the past.

Informing Researchers

Part of the dilemma faced by institutions which are responsible for the oversight of research performed under their watch is that people conducting research may be unaware that what they are doing would be classified as research. At some institutions there are still many researchers who don't realize they are required to seek IRB approval or exemption of their research, much less that there are education requirements (or at least *strong recommendations* for education) for the research personnel. The degree to which the new education requirements are emphasized and publicized obviously varies from institution to institution. However, there is probably no institution at which every researcher or potential researcher is fully aware of the new requirements. The necessary education really begins at the point of defining research, which according to 45 CFR 46 is defined as:

> (d) *Research* [italics in original] means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. Activities which meet this definition constitute research for purposes of this policy, whether or not they are conducted or supported under a program which is considered research for other purposes. For example, some demonstration and service programs may include research activities (Protection of Human Subjects, §46.102).

In other words, some researchers need to be educated in order to know that they need to be educated.

Grant funded projects generally pass through so many hands prior to study activation, that there are many opportunities for researchers to be made aware of the need for education and many points within the process for verification of education to take place. Non-funded projects are more likely to slip by without education requirements being met. Some institutions have begun requiring verification of education prior to approval of research projects by the IRB, as previously mentioned in the VA policy. This is effective for enforcing the education requirement for those researchers who understand that research protocols must pass through the IRB and comply, but does not address those who are unaware that their project must be reviewed by an IRB in the first place.

Implications for Research

Assuming that researchers are aware of the new push for education and for whom, there are several possible implications for research and researchers, some positive and some negative.

Research May Be Carried Out More Ethically

The primary impetus for the education requirement was to ensure that human subjects who participate in research

would be treated ethically. Hopefully this is and will continue to be the case and will only be improved by having better educated human subjects researchers.

Public Trust May Be Enhanced

A second potential positive outcome of increasingly well-trained researchers is that as research participants have contact with key personnel working within research projects, they may be better informed of their rights as research participants. Also, not only should participants be less likely to be harmed in the research process, but also may see it as a more positive experience than if the research personnel had not been trained. This would contribute to the public trust of research mentioned previously.

Less Research May Be Performed

However, there are several more negative possibilities created by new education policies. There is the possibility that potential researchers will choose not to undertake research they might have otherwise carried out because of the inconvenience of participating in training or ensuring that key personnel have completed training. The same could be said for traversing the IRB approval process as well. There are probably potential research projects that have never been initiated because of the inconvenience of obtaining IRB approval, but the new education requirements add one more hurdle, although a noble one, to the pre-research process. The primary implications of this possibility are that knowledge that would have been gained from that research is not and researchers forfeit the gratification and personal and professional growth that comes from conducting the research.

Research Designs May Be Modified

A second negative possibility is that research designs will be modified to reduce or eliminate the use of human subjects. Similarly, in the past, with restrictions on conducting research with children being made more stringent, there has been concern that children might be excluded from virtually all research (Robenznieks, 2001). While this removes risks of participating in a particular research study, it also reduces the knowledge gained about children, which would be beneficial. Again, eliminating the use of human subjects would reduce the amount of information gained from a study, either making it less informative, or even useless, whereas a study employing human subjects might be vastly more useful.

Similarly, because the education requirements extend to all key personnel at all research sites, researchers may tend to limit the number of key personnel and/or sites involved in a study. This could have two implications. One would be that smaller scale studies may be performed, which tend to have less external validity, and the second would be that fewer individuals have the opportunity to be involved in the conduct of research. While modifying research to avoid the use of human subjects, limiting study sites, limiting study personnel, and the abandonment of research projects altogether are not unethical, they do have negative implications. All may result in the reduction of new knowledge that would have been gained from research.

Researchers May Fail to Seek IRB Approval

Another possibility, and one that borders on unethical, is that some researchers will not seek IRB approval for their studies because of the added burden of having to document training for themselves and all key personnel. This could actually result in more potential risk for research participants than existed prior to the mandate for education. Of course, if this were to be discovered, the institution could be found "out of compliance" and be forced to suspend research until human subjects research violations were rectified.

Researchers May Falsify Training Documentation

A clearly unethical possibility is that researchers may lie about their training in order to receive study approval. Because institutions are providing their own education, testing, and certificating, the likelihood of this would vary among institutions. Individuals transferring between institutions and those conducting projects that utilize multiple sites would have more of an opportunity for such fabrication.

Requirements May Be Over-interpreted or Misinterpreted

A final possibility is that the requirement for education may be over-interpreted, and thereby too broadly applied, resulting in many of the problems outlined above, or misinterpreted. Education is clearly supported by federal policies at this point (see e.g., 65 Fed. Reg. 37137, June 13, 2000; "DHHS/NIH announce new initiatives", 2000; NIH: Office of Extramural Research, 2000, June 5). However, policies are so new, that few have been tested. It is therefore unclear exactly what penalties will be for failure to comply. The OHRP website currently indicates that comments are being received about the FWA and the *current* version [March, 2002] is available (Federalwide Assurance of Protection for Human Subjects, 2002, March 2). This implies that changes should be expected, but to ignore this emphasis on education would likely be a choice to have to do planning and implementation of an education plan quickly at a later date.

Reasons for Compliance

It is assumed that anyone conducting research will comply with the new education requirements, although reasons for compliance will vary. Ideally, researchers will see the benefit to the research subjects and the field of human research in general, and will comply willingly. However, many,

if not most, will comply because of the threat of consequences for themselves or their institution. Hopefully, research will only be improved and not curtailed.

Implications for Educational Researchers

A question of particular interest to educational researchers is whether there are implications that are specific to educational research? Educational researchers within higher education settings typically have access to an IRB and training opportunities and will be held accountable for meeting federal mandates, just as any other researchers within higher education because of the risk to the institution for noncompliance. However, educational researchers who conduct studies at other sites (e.g., K-12 schools) often have no IRB and have less access to training materials and current regulations and may be more likely to violate requirements for the protection of human subjects. Additionally, the determination of whether a project should be classified as research, and therefore governed as research, is often difficult within educational settings. What might be classified as research by 45 CFR 46 might appear to a classroom teacher to be a part of regular classroom practice. For example, if a teacher were to teach two Algebra I classes using two different textbooks, compare the knowledge of the two classes at the end of the school year, and present the findings at a national mathematics teaching conference, research has been done. New knowledge was generated and communicated.

At this point, because the recommendation for human subjects education is so new, colleges, universities, and medical schools have been the focus of energy with regard to developing and overseeing compliance. However, as most of those institutions have their programs in order, the focus may broaden to include other organizations. At that point, individuals within schools and school systems will most likely be expected to meet and verify that they have met the human subjects protection education guidelines.

Implications for Professional Organizations

Finally, what are the implications for professional organizations that sponsor conferences or publish journals in which research findings are shared? Is there any responsibility to verify whether the research was done in compliance with the institution's IRB? While many scholarly journals do ask for authors to sign a statement indicating that research was conducted ethically, most professional presentations do not require such certification. Journals and professional organizations have not historically sought to verify this, nor would it be feasible to do so. Should an organization or journal be concerned about whether an author or presenter has fulfilled the education recommendation, or would that fall under any statement of ethical practice that is already requested? We believe the latter is more practical.

Verification of ethical practices or education probably falls beyond what is likely, or even possible for most journals or organizations, however, professional organizations could provide some of the required education in the ethical treatment of human research subjects as a service to participants. This could serve to make it more likely that researchers were aware of ethical research practices, not only through the content, but also just the publicity of having the topic on the program at a conference.

Conclusions

While many researchers probably see this new educational recommendation (or requirement, depending on the type of institution) as a burden that will impede their research, it looks as though it is here to stay. It is not clear yet was the results will be for failure to comply. Various types of noncompliance with Federal guidelines regarding human subjects research have resulted in suspension of one investigator's personal research, to the suspension of an entire institution's research. For decades most researchers have followed non-legislatively enacted codes such as the Nuremburg Code, The Declaration of Helsinki, and The Belmont Report for ethical guidance. However, due to the continuation of and publicizing of unethical research practices, the public and certain lawmakers have acted on the belief that oversight needs to be *ratcheted up a notch*. Thus, *strongly recommended* education; and is education not what educational researchers are in the business of promoting?

References

Council on Governmental Relations (2000). *COGR June 2000 meeting report.*

Curry, D. (2001). U.S. restricts research at Johns Hopkins after a volunteer's death. *The Chronicle of Higher Education, 47*(47), p.A25(1)

DHHS/NIH announce new initiatives for human subject research. (2000, July/August). *Federal Grants News for Colleges and University*, 1-3.

Federalwide Assurance of Protection for Human Subjects (March, 2, 2002). U.S. Department of Health and Human Services (DHHS) Office for Human Research Protections (OHRP). Retrieved from the World Wide Web on September 23, 2002 http://ohrp.osophs.dhhs.gov/humansubjects/assurance/filasurt.htm

Hansen, J., & Spencer, T. (2000, January 21). University of Alabama, Birmingham research suspended. *Birmingham News*. Retrieved June 28, 2002, from http://www.naim.org/wwwboard/messages/202.htm

Levine, S. (2001, July 21). Agency to hasten Hopkins review: University allowed to continue studies. *Washington Post*. Retrieved June 28, 2002 from http://www.washingtonpost.com/ac2/wp-dyn/A11691-2001Jul17?language=printer

Milgram, S. (1994a). Obedience in Vietnam. In Behrens, L., & Rosen, L. J. (eds.), *Writing and Reading Across the Curriculum (5th ed.), New York: HarperCollins.*

Milgram, S. (1994b) The perils of obedience. In Behrens, L., & Rosen, L. J. (eds.). *Writing and Reading Across the Curriculum* (5th ed.), New York: HarperCollins.

National Institutes of Health: Office of Extramural Research. (2002, May 8). *Frequently asked questions for the requirement for education on the protection of human subjects.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/policy/he_educ_faq.htm

National Institutes of Health: Office of Extramural Research. (2001, September 5). *Required education in the protection of human research participants.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/guide/notice-files/NOT-OD-01-061.html

National Institutes of Health: Office of Extramural Research. (2000, June 5). *Required education in the protection of human research participants.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/guide/notice-files/NOT-OD-00-039.html

Office of Human Research Protection. (2001, June 21). *IRB guidebook.* Retrieved July 26, 2002 from http://ohrp.osophs.dhhs.gov/irb/irb_guidebook.htm

Office of Human Subjects Research: National Institutes of Health. (2000). *Answers to questions frequently asked of NIH'S Office of Human Subjects Research (OHSR).* Retrieved June 6, 2002 from http://ohsr.od.nih.gov/info/hinfo_8.php3

Office of Research Integrity. (2001, February). *Responsible conduct of research (RCR) education.* Retrieved July 23, 2002 from http://ori.dhhs.gov/html/programs/congressionalconcernsinquiry.asp

Protection of Human Subjects, 45 C.F.R. 46 (Revised November 13, 2001) (Effective December 13, 2001)

Riccarrdi, N., & Monmaney, T. (2000, April 27). King/Drew medical research suspended: Voluntary action comes after regulators find two dozen violations of rules designed to protect human subjects and ensure credibility at south L.A. facility. *Los Angeles Times.* Retrieved June 28, 2002 from http://aegis.com/news/lt/2000/LT000404.html

Robiznieks, A. (2001a, September 24). University of Michigan suspends top researcher. *American Medical News.* p. 14.

Robiznieks, A. (2001b, October 8). Ruling bans research where children are at "any risk." *American Medical News.* p. 20.

U.S. Department of Health & Human Services: Office of Research Integrity (ORI) (2002, February 12). *Programs: Responsible conduct of research (RCR) education.* Retrieved July 23, 2002 from http://ori.dhhs.gov/html/programs/congressionalconcerns.asp

Author Note

Please direct all correspondence to Dr. Andrea D. Clements, Human Development and Learning Department, Box 70548, East Tennessee State University, Johnson City, Tennessee, 37614 or clements@etsu.edu.

Andrea D. Clements, Associate Professor and Coordinator of the Foundations of Human Development and Learning Program at East Tennessee State University, is the Director of Ethics for the Eastern Educational Research Association. She also is a member of the Institutional Review Board (IRB) and Expedited Review Subcommittee of the IRB at East Tennessee State University.

Pat W. Myrick is the Compliance Manager for the Institutional Review Boards at East Tennessee State University and James H. Quillen College of Medicine. She is a certified Institutional Review Board Professional and a Certified Clinical Research Professional. She developed the Responsibility In Human Subject Research Compliance Education Program at East Tennessee State University.

Research Practices that Produce Type VI Errors

Isadore Newman
The University of Akron

John W. Fraas
Ashland University

Carole Newman
The University of Akron

Russell Brown
Summa Health Care System

A Type VI error occurs when an inconsistency exists between the study's research question and the analytic technique and/or research design used in the study. This article presents various research practices that lead to Type VI errors. Specifically discussed are the practices that (a) fail to distinguish between statistical analysis and research design issues, (b) do not match the model used in structural equation modeling with the research question, (c) analyze a research question that involves practical significance with an analytical technique that fails to do so, (d) use methods designed to control for inflated Type I error rates that do not match the nature of the research question, and (e) employ multivariate data

Milgram, S. (1994a). Obedience in Vietnam. In Behrens, L., & Rosen, L. J. (eds.), *Writing and Reading Across the Curriculum (5th ed.), New York: HarperCollins.*

Milgram, S. (1994b) The perils of obedience. In Behrens, L., & Rosen, L. J. (eds.). *Writing and Reading Across the Curriculum* (5th ed.), New York: HarperCollins.

National Institutes of Health: Office of Extramural Research. (2002, May 8). *Frequently asked questions for the requirement for education on the protection of human subjects.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/policy/he_educ_faq.htm

National Institutes of Health: Office of Extramural Research. (2001, September 5). *Required education in the protection of human research participants.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/guide/notice-files/NOT-OD-01-061.html

National Institutes of Health: Office of Extramural Research. (2000, June 5). *Required education in the protection of human research participants.* Retrieved June 6, 2002 from http://grants.nih.gov/grants/guide/notice-files/NOT-OD-00-039.html

Office of Human Research Protection. (2001, June 21). *IRB guidebook.* Retrieved July 26, 2002 from http://ohrp.osophs.dhhs.gov/irb/irb_guidebook.htm

Office of Human Subjects Research: National Institutes of Health. (2000). *Answers to questions frequently asked of NIH'S Office of Human Subjects Research (OHSR).* Retrieved June 6, 2002 from http://ohsr.od.nih.gov/info/hinfo_8.php3

Office of Research Integrity. (2001, February). *Responsible conduct of research (RCR) education.* Retrieved July 23, 2002 from http://ori.dhhs.gov/html/programs/congressionalconcernsinquiry.asp

Protection of Human Subjects, 45 C.F.R. 46 (Revised November 13, 2001) (Effective December 13, 2001)

Riccardi, N., & Monmaney, T. (2000, April 27). King/Drew medical research suspended: Voluntary action comes after regulators find two dozen violations of rules designed to protect human subjects and ensure credibility at south L.A. facility. *Los Angeles Times.* Retrieved June 28, 2002 from http://aegis.com/news/lt/2000/LT000404.html

Robiznieks, A. (2001a, September 24). University of Michigan suspends top researcher. *American Medical News.* p. 14.

Robiznieks, A. (2001b, October 8). Ruling bans research where children are at "any risk." *American Medical News.* p. 20.

U.S. Department of Health & Human Services: Office of Research Integrity (ORI) (2002, February 12). *Programs: Responsible conduct of research (RCR) education.* Retrieved July 23, 2002 from http://ori.dhhs.gov/html/programs/congressionalconcerns.asp

Author Note

Please direct all correspondence to Dr. Andrea D. Clements, Human Development and Learning Department, Box 70548, East Tennessee State University, Johnson City, Tennessee, 37614 or clements@etsu.edu.

Andrea D. Clements, Associate Professor and Coordinator of the Foundations of Human Development and Learning Program at East Tennessee State University, is the Director of Ethics for the Eastern Educational Research Association. She also is a member of the Institutional Review Board (IRB) and Expedited Review Subcommittee of the IRB at East Tennessee State University.

Pat W. Myrick is the Compliance Manager for the Institutional Review Boards at East Tennessee State University and James H. Quillen College of Medicine. She is a certified Institutional Review Board Professional and a Certified Clinical Research Professional. She developed the Responsibility In Human Subject Research Compliance Education Program at East Tennessee State University.

Research Practices that Produce Type VI Errors

Isadore Newman
The University of Akron

John W. Fraas
Ashland University

Carole Newman
The University of Akron

Russell Brown
Summa Health Care System

A Type VI error occurs when an inconsistency exists between the study's research question and the analytic technique and/or research design used in the study. This article presents various research practices that lead to Type VI errors. Specifically discussed are the practices that (a) fail to distinguish between statistical analysis and research design issues, (b) do not match the model used in structural equation modeling with the research question, (c) analyze a research question that involves practical significance with an analytical technique that fails to do so, (d) use methods designed to control for inflated Type I error rates that do not match the nature of the research question, and (e) employ multivariate data

analysis techniques for research questions that require the application of univariate techniques. An examination of these types of mismatches between the analytical technique and/or the research design and the research question should assist researchers in avoiding Type VI errors.

Current Research Practices that Produce Type VI Errors

The research and analytical techniques used by the researcher should not lead to Type VI errors, that is, a condition in which an inconsistency exists between the study's research question and the analytic technique and/or research design (Newman, Deitchman, Burkholder, Sanders, & Ervin, 1976). Type VI errors are committed when researchers (a) fail to distinguish between statistical analysis and research design issues, (b) do not match the model used in structural equation modeling with the research question, (c) analyze a research question that involves practical significance with an analytical technique that fails to do so, (d) use methods designed to control for inflated Type I error rates that do not match the nature of the research question, and (e) employ multivariate data analysis techniques for research questions that require the application of univariate techniques.

The current state of computer hardware and software allows researchers to apply sophisticated statistical techniques with relative ease. It must be noted, however, that the point-and-click environment of today's computer statistical software allows researchers to use these sophisticated techniques in a rather mechanical fashion. Thus, while such ease of use leads to efficiency, it may also foster the inappropriate match between the analytical technique and the research question, which will lead to a Type VI error. An examination of these types of mismatches between the study's research question and the analytic technique and/or research design should assist researchers in avoiding Type VI errors.

Differences Between Issues Addressed by Statistical Analysis and Research Design

A Type VI error is likely to occur when a researcher does not distinguish between the issues addressed by the analytical techniques employed in a study and those issues dealt with by the study's research design. Although this distinction may seem obvious, it is vastly overlooked by too many researchers. To illustrate, in 1980 the United States Supreme Court based decisions in racial and sex discrimination suits on statistical analyses that reported statistically significant relationships (i.e., correlation values). However, these studies were void of the types of research design controls discussed by Fisher (1971), Campbell and Stanley (1963), and Newman and Newman (1994). These decisions were based on the misunderstanding that the use of sophisticated statistical analyses allows one to assume causation. If the research question was designed to determine a cause for the discrepancies between opportunities for advancement or pay, the use of correlations would be one form of a Type VI error. If, however, the research question attempted to establish the existence of relationships, the applications of correlation analytical techniques would not produce a Type VI error unless the researcher attempted to make casual inferences from those correlation results.

To further understand the appropriate role of all statistical tests, it is important to note that all tests of significance are tests of relationships. Therefore, each statistical test employed by researchers can be mathematically translated into the degree of relationship between the variables. Conversely, if one knows the sample size, all mathematical relationships can be converted into tests of significance. For example, as noted by McNeil, Newman, and Kelley (1996), a *t* test is equivalent to a point-biserial correlation; an *F* test is equivalent to an *eta* value or an *R* value; and a 2X2 chi-square value is equivalent to a *phi* coefficient. The point biserial correlation, *phi*, *eta*, *t*, *R*, *F* and chi-square values are all related. Thus, any test of significance merely gives a probability estimate that a relationship is not due to random variation. It does not yield any information about the causal nature of that relationship (i.e., the degree of internal validity). Only to the extent that a research design has a high degree of internal validity can one assume causality.

Type VI Error Concerns with Structural Equation Modeling

Structural Equation Modeling (SEM) has been conceptualized as the "marriage of path analysis to factor analysis." Thus, SEM is purported to have many advantages including the ability to (a) represent latent variables (or factors) that are theoretically presumed to be the cause of observed variables, (b) model theoretical relationships, (c) estimate measurement error, and (d) estimate the overall "fit" of the model (Quintana & Maxwell, 1999). Of these advantages, perhaps, the most important is that the models used in conjunction with SEM tend to be theoretically based. It is important to note, however, that researchers who mechanically make certain assumptions and interpretations with SEM increase the chances of committing Type VI errors. Five of these assumptions and interpretations are (a) the assumption of causation, (b) the practice of not including interaction terms, (c) the interpretation of the goodness-of-fit indices, (d) assumptions regarding factor fit, and (e) the assumption regarding measurement error due to factor instability.

The assumption of causation. Some authors who use SEM have the tendency to interpret their results in a manner that implies, if not directly states, that a causal relationship exists. For example, in a study by Aquino, Russell, Cutrona,

and Altmaier (1996), which examined the relationship between social support, employment and life satisfaction in the elderly, the authors provided a causal interpretation of an ex post facto design. With respect to the appropriateness of this causal interpretation, it is important to note that SEM is an "extension of the collection of statistics belonging to the general linear model (GLM)" (Quintana & Maxwell, 1999, p. 486), and is, therefore, an analytic technique and not a research design. Yet, SEM has been explicitly described as causal modeling (Fassinger, 1987). The components of SEM (factor analysis and path analysis) are both based on the analysis of the covariance/correlation matrix. As such, SEM is still based on the degree of relationship (correlation) among variables. Thus SEM does not inherently control for threats to internal validity. Correlation and causation are not the same. Since causal inferences depend on appropriate research design and a statistical method does not allow for causal inference, researchers must be keenly aware of what researchers can conclude based on the statistical analysis and what limitations the study's research design imposes on those conclusions, even with SEM.

The practice of not including interaction terms. SEM is infrequently used to test interaction effects (Newman & Marchant, 1993; Ridenour, Marchant, & Newman, 1994). Therefore, if the theory assumes interaction, which should be reflected in the research question, an inconsistency between the analysis and the research question would exist. It should be noted, however, that researchers who use SEM have dealt with interaction effects by analyzing separate models for variables that interact. This is a simple and elegant solution when (a) there are very few variables that interact, (b) it is a first order interaction, and (c) the variables are dichotomous (e.g., gender) in nature rather than ordinal or interval.

The interaction problem is relatively easy to resolve for a dichotomous variable as was modeled by Robitschek and Kashubeck (1999). They developed separate models for males and females in order to model the relationship of family functioning and alcoholism to a person's well-being and level of distress. The inclusion of interaction terms in a SEM analysis becomes considerably more complicated when there are many interactions, the variables are ordinal or interval in nature (e.g., intelligence), and higher order interaction exists (e.g., intelligence by age, and intelligence by age by gender). It is not unlikely for researchers to encounter these types of interactions and variables when testing social science theories. Since there is no mathematical model, of which we are aware, that can adequately assess such research questions when using SEM (Newman & Marth, 1995), the use of SEM to test research questions that include these types of interactions would produce Type VI errors. Equally important is the possibility that researchers who are committed to using SEM may be less inclined to propose research questions involving these types of interaction effects. Thus, an inconsistency between the analytic technique used and the "appropriate" research question would exist.

Interpretation of the goodness-of-fit indices. An inconsistency may exist between the method used with SEM to assess the model's goodness of fit and the intent of the research question. Hair, Anderson, Tatham, and Black (1998) present a plethora of goodness-of-fit measures that are used in conjunction with SEM. They divided these goodness-of-fit measures into three types — measures of absolute fit, measures of incremental fit, and measures of parsimonious fit. The measures of absolute fit include: (a) the likelihood-ratio chi-square statistic, (b) the noncentrality parameter, (c) the goodness-of-fit-index, (d) the root mean square residual, (e) the root mean square error of approximation, (f) the expected cross-validation index, and (g) the cross-validation index. Some of the measures of incremental fit are: (a) the adjusted goodness-of-fit index, (b) the Tucker-Lewis index, (c) the normed fit index, (d) the relative fit index, (e) the incremental fit index, and (f) the comparative fit index. Parsimonious fit measures include: (a) the parsimonious normed fit index, (b) the parsimonious goodness-of-fit index, (c) the normed chi-square, and (d) the Akaike information criterion.

These goodness-of-fit measures are based on a similar concept (i.e., the model's ability to reproduce the correlation matrix). This method of assessing model fit can lead to an incorrect assessment under two possible scenarios. First, it is possible for a model that contains numerous paths to produce a goodness-of-fit value that indicates a good level of fit in spite of the fact that none of the path values are statistically significant. Second, it is possible for the model to effectively reproduce the correlation matrix even when some of the relationships are in the opposite direction from the direction predicted by the theory. While each of these methods assesses the ability of the model to reproduce the correlation matrix, it does not provide a direct means to assess the number of relationships that are statistically significant and in the direction predicted by the theory. In fact, a number of the assessment measurements could be high and all of the relationships could be in the opposite direction of the theory! Interpretation of fit indices, alone, as support or lack of support for a theoretical model could produce a Type VI error.

Consideration of the number of path values that are statistically significant in the predicted direction would decrease the likelihood of committing a Type VI error when using SEM. This method of assessing the model's goodness of fit is not based on the model's ability to reproduce the correlation matrix, but rather on the ability of the model to reflect the path values as suggested by theory. The use of this approach, which has been referred to as the Binomial Index of Model Fit (Fraas & Newman, 1994; Newman, Fraas & Norfolk, 1995), would reduce the chance of an inconsistency between the assessment of the model's goodness of fit and the research question.

Assumptions regarding factor fit. A Type VI error can be committed when SEM is used to determine if the factors are good estimates of the theoretical factor structure (i.e., SEM is used for confirmatory factor analysis). In a

measurement model with three factors, it is very possible that two factors fit the structure well and one does not. For such a case, the traditional goodness-of-fit techniques may simply indicate a poor fit without revealing which factors do not fit the structure. If the research question requires the analysis to reveal which factors do and do not fit the structure and the assessment technique is not capable of providing such evidence, a Type VI error will be committed. To avoid this error, researchers could employ the Kaiser Factor Matching technique (Newman, Dimitrov, & Waechter, 2000). This technique would indicate which of the two factors fit well. Newman, et al. suggested that Kaiser Factor Matching be used along with the more traditional estimates to get a better picture of the relationships of the factors to the underlying theoretical constructs and thus avoid a Type VI error.

The assumption regarding measurement error due to factor instability. The use of SEM could lead to another form of Type VI error based on how it purports to control for measurement error. That is, the underlying factors generated by SEM may not adequately represent the variables contained in the research question. A somewhat over simplification of this procedure may assist in the explanation of this mismatch. Instead of using the indicator variables, SEM uses the underlying factors as predictor variables. The assumption is that the factors are stable and therefore control for measurement error. It is likely that this procedure does reduce measurement error but it does not eliminate it. Thus, the use of these factors may cause other types of errors due to the fact that the factors may be quite specific to the sample analyzed. This problem is likely compounded by the tendency of researchers to use only a small number of indicators, such as three or four per latent variable, a practice that has been advocated in the literature (Quintana Maxwell &, 1999). In some cases, single indicators are used to represent latent variables. For example, Tuason and Friedlander (2000), in a study of Bowen's theory of differentiation, used single indicators to represent three of six of the latent variables in their study. This was accomplished by supplying estimates of the reliability of the measurement of each of the indicator variables. Traditionally, it has been reported that five or more items are necessary to create a stable factor (Nunnally, 1972). Thus, a good but infrequently used practice would be to cross-validate the model to estimate the stability of the latent variables. The use of cross-validation would assist researchers in avoiding Type VI errors.

Importance of Statistical Significance and Practical Significance

Numerous researchers and journal editors are calling for the inclusion of practical significance (e.g., effect sizes) in research results and, thus, presumably in research questions (Cohen, 1977, Fraas & Newman, 2000; Huberty, 1993; Levin & Robinson, 2000; Newman & Fraas, 2001; Robinson & Levin, 1997; Shaver, 1993; Thompson, 1996, 1998, 1999a, 1999b, 1999c). If research questions contain a practical significance component and the analytic techniques used to test that research question do not incorporate a practical significance component, a Type VI error will be committed. Thus, in order to avoid a Type VI error for such research questions, it has been argued that the analytical techniques used must address both the issues of statistical significance and practical significance (Fraas & Newman, 2000; Newman & Fraas, 2001; Robinson & Levin, 2000).

Fraas and Newman (2000) and Newman and Fraas (2001) discussed the testing of non-nil null hypotheses (i.e., hypotheses that incorporate non-zero values) as a means of incorporating both the test of statistical and practical significance. To illustrate the use of a non-nil null hypothesis, consider a study designed to test the effectiveness of two treatments intended to produce weight losses for cardiac patients. An important issue for researchers to initially address is the establishment of how large the difference in the weight-loss means of the two groups must be in order for it to be considered important, that is, practically significant. Fraas and Newman expressed the view that this process is probably best undertaken by involving practitioners (i.e., people directly involved in the process being studied) and researchers in the field.

In this illustration, we are assuming that practitioners and researchers have identified that any difference between the group means greater than 10 pounds would be considered practically significant. For this practical significance level, the non-nil null hypothesis that the mean for Group 1 minus the mean for Group 2 equals 10 pounds is more meaningful clinically than the nil null hypothesis that the mean for Group 1 minus the mean for Group 2 equals zero pounds. That is, instead of saying we are 95% confident that the difference is greater than zero, we can say we are 95% confident that the difference is greater than 10. If experts determined that a difference of 10 pounds was minimally required for clinical benefit, the latter statement is far more meaningful and useful for the practitioners who will utilize the study's results.

It is important to note that the use of non-nil null hypotheses require the researchers and practitioners to identify what is clinically or pragmatically meaningful before initiating the research. The decisions determining the appropriate effect size need to be related to the purpose for doing the research. Newman, Ridenour, Newman, and DiMarco (2002) indicate it is important that researchers frame their research question so it is consistent with their purpose. This clear understanding of purpose is necessary to help the researcher identify the appropriate effect size. Without looking at the research question in the context of its purpose, one cannot identify an appropriate and useful clinical effect size. The incorporation of the effect size into the research question dictates that statistical significance is not considered independent of the practical significance level.

An additional point should be noted with respect to establishing the practical significance level. When dealing with the issue of how large should an effect size be in order for it to be considered practically significant, Newman and Newman (2000) presented the case that even small effect sizes, as measured by small R-square values, may be important, that is, practically significant. They concluded that small R-square values could be very valuable and useful, depending on the research question being asked and the size of the population. This position, which is supported by Rosenthal and Rosnow (1991) and Deming (1982), suggests that practical significance should be measured in a relative sense rather than an absolute manner. This position again suggests that the involvement of practitioners in establishing the practical significance level is an essential component in establishing a meaningful research question.

It is essential that the established practical significance level be incorporated into the analytic technique used to test the research question in order to avoid committing a Type VI error. Fraas and Newman (2000) suggested that one technique that could be used to test a non-nil null hypothesis is a randomization test. The use of a randomization test, which generates its own distributions, could be used to determine if the observed difference adjusted for the practical significance level is unlikely to be due to chance variation. It should be noted, however, that Newman and Fraas (2001) conducted a Monte Carlo study regarding the impact of including the practical significance level into the research question on the Type I error rates of independent *t* tests of two group means. The results revealed negligible impacts of such tests on the Type I error rates. Thus, non-nil null hypotheses may, under certain conditions, be tested with parametric tests.

Controlling Type I Error Rates

Researchers are familiar with the practice of controlling for inflated Type I error rates (i.e., the chance of rejecting a true null hypothesis) in studies that involve multiple statistical tests. If the adjustment technique the researchers use, however, does not match the research question with respect to the assumed error rate unit and directional and non-directional relationships, a Type VI error will be committed. Newman, Fraas and Laux (2000) suggested a non-mechanical approach that requires the researcher reflect on three elements of the adjustment process.

First, the researcher identifies the error rate unit or units (i.e., the various groupings of tests for which adjustments are to be made). Second, adjustments are not made for directional tests (i.e., the hypotheses reflect a direction based on theory and/or previous empirical results). Newman, Fraas, and Laux (2000) suggested that if a researcher adjusts the alpha levels for tests that are predicted from previous experience or theory, the researcher would probably overcorrect for possible inflated Type I errors. Such an outcome would increase the likelihood of Type II errors being committed. Third, the alpha levels for the non-directional tests in a given error rate unit are adjusted for the number of such tests. This approach requires researchers to reflect on the adjustment process at two stages of the research process (a) the point at which the research question is being formulated and (b) the point at which the analytic technique is being chosen and implemented. Reflecting on the adjustment process at both of these stages should decrease the likelihood that the adjustment technique used does not match the research question, that is, decrease the chance of a Type VI error.

Multivariate Tests versus Univariate Tests

Various statistics books recommend that if one has multiple dependent variables, multivariate analyses should be used. One reason for this recommendation relates to controlling for inflated Type I error rates when multiple statistical tests are conducted on a set of dependent variables (Stevens 1996). Unfortunately this recommendation is frequently misunderstood since the overall multivariate test of significance is only one aspect of the multivariate testing procedure. When a significant multivariate test is followed by univariate tests to determine where the significant differences lie, a control for multiple tests is still necessary to keep the Type I error rate from being inflated (Croom, 1987; Newman, Croom, Mugrage & Hoedt, 1983).

The most important consideration in determining whether one should use multivariate or univariate measures should be how well the statistical procedures reflect the question of interest (Newman & Benz, 1987). If the question of interest is not related to the underlying construct of a set of independent variables, but is instead related to each dependent variable, univariate analyses should be used. Multivariate analyses should only be used when one is interested in the underlying hypothetical construct of the set of dependent variables, which is frequently not the case (J. Finn, personal communication, October, 2001; Thompson, 1994). The use of multivariate analyses with a research question that reflects an interest in each dependent variable will produce a mismatch between the analytic technique and the research question, which is a Type VI error.

An additional point is important to consider regarding the use of multivariate analyses. All multivariate analysis of variance techniques are subsets of canonical analysis, just as traditional analysis of variance techniques are subsets of multiple regression analyses. A canonical correlation measures the degree of association between the underlying construct of a set of predictor variables with the underlying construct of a set of criterion variables. It does not answer the question: What is the degree of association among the individual predictor variables and the individual criterion variables simultaneously? If a researcher is interested in each dependent variable separate from the others and the researcher conducts a multivariate analysis, a Type VI error would be committed. Similarly, if a researcher is interested in

a linear combination of the dependent variables and the researcher conducts a univariate analysis, a Type VI error would also be committed. The research must match the multivariate and univariate analytical techniques with the appropriate research questions.

Implications for Researchers

The intent of this article is to alert researchers to possible research practices that may lead to inconsistencies among the research questions, the analytical techniques employed and/or the research designs (i.e., Type VI errors). To avoid Type VI errors, researchers must be careful not to approach the research process in a mechanical fashion, but rather they must evaluate whether their analytic techniques and research design are appropriate for their research questions.

Specifically, a researcher can decrease their chances of committing Type VI errors by addressing a number of key questions:

1. When stating the conclusions resulting from my study, did I clearly delineate and separate the issues addressed by my research design from those issues related to the analytic techniques employed?

2. If I used SEM, did my model and analysis reflect my research question, including the need to: (a) incorporate interaction variables, (b) use appropriate goodness-of-fit measures, (c) reveal which factors do and do not fit the structure, and (d) have the underlying factors generated by SEM adequately represent the variables.

3. If my research question contains a practical significance component, do my analytic techniques allow me to adequately test that research question? If my research question does not contain a practical significance component, should it?

4. If my study involves multiple statistical tests, does the method I used to control for inflated Type I error rates match the nature of my research question? That is, have I accurately identified my error rate units and appropriately considered the directional versus the nondirectional nature of my hypotheses?

5. Did the analytic technique that I employed accurately reflect whether the nature of my research question was univariate or multivariate?

We believe that Type VI errors would be less common in today's research if researchers would reflect on the issues addressed by these questions.

References

Aquino, J., Russell, D., Cutrona, C., & Altmaier, E. (1996). Employment status, social support, and life satisfaction among the elderly. *Journal of Counseling Psychology, 43*(4), 480-489.

Campbell, D.T., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press, Inc.

Croom, W. (1987). *Selected psychometric properties and measures of interpretability of two methods of estimating population parameters: Canonical correlation and multiple linear regression.* Unpublished doctoral dissertation, The University of Akron, Akron, OH.

Deming, E. (1982). *Out of crisis: Center for advanced engineering study.* Cambridge: Massachusetts Institute of Technology.

Fassinger (1987). Use of structural equation modeling in counseling psychology research. *Journal of Counseling Psychology, 34*(4), 425-436.

Fisher, R.A. (1971). *Collected papers of R.A. Fisher.* Edited by J.H. Bennett. Adelaide: The University of Adelaide.

Fraas, J.W., & Newman, I. (1994). A binomial test of model fit. *Structural Equation Modeling: A Multidisciplinary Journal, 1*(3), 268-273.

Fraas, J.W., & Newman, I. (2000, October). *Testing for statistical and practical significance: A suggested technique using a randomization test.* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Huberty, C. J. (1993). Historical origins of statistical testing practices: The treatment of Fisher versus Neyman-Pearson views in textbooks. *The Journal of Experimental Education, 61*(4), 317-333.

Levin, J.R., & Robinson, D.H. (2000) Rejoinder: Statistical hypothesis testing, effect size estimate, and the conclusion of coherence of primary research studies. *Educational Researcher, 29*(1), 34-36.

McNeil, K., Newman, I., & Kelly, F. (1996). *Testing research hypotheses with the general linear model.* Carbondale, IL: Southern Illinois University Press.

Newman I. & Benz, C. (1987). *Comparing canonical correlation strategies to model comparison multiple regression on one data set* (technical report). Akron, OH: The University of Akron, Office of Educational Research and Evaluation.

Newman, I., Croom, W., Mudgrage, B., & Hoedt, K. (1983). Note: A suggested method for correcting alpha error build-up on orthogonal dependent variables. *Multiple Linear Regression Viewpoints, 12*(1), 56-60.

Newman, I., Deitchman, R., Burkholder, J., Sanders, R., & Ervin, L. (1976). Type VI Error: Inconsistency between the statistical procedure and the research question. *Multiple Linear Regression Viewpoints, 6*(4), 1-19.

Newman, I., Dimitrov, D., & Waechter, D. (2000). Factor structure of perceived individualization of instruction: Argument for multiple perspective. *Educational Research Quarterly, 24*(1), 20-29.

Newman, I., & Fraas, J.W. (2001, October). *Testing non-nil null hypotheses with randomization and t tests: A Monte Carlo study.* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Newman, I, Fraas, J.W., & Laux, J. (2000). A three-step adjustment procedure for Type I error rates. *Journal of Research in Education, 10*(1), 84-90.

Newman, I., Fraas, J.W., & Norfolk, T. (1995). Binomial index of model fit: An elaboration. *Structural Equation Modeling: A Multidisciplinary Journal, 2*(2), 155-162.

Newman, I. & Marchant, G. J. (1993, April) *Type VI error in path analysis: Testing for interactions.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, Ga.

Newman, I., & Marth, J. (1995). To path analyze or to not path analyze: Is there an alternative approach. *Multiple Linear Regression Viewpoints, 22*(1), 7-12.

Newman, I., & Newman, C. (1994). *Conceptual statistics for beginners.* (2nd ed.). New York: University Press of America.

Newman, I., & Newman, C. (2000). A discussion of low R^2: Concerns and use. *Educational Research Quarterly, 24*(2), 3-9.

Newman, I., Ridenour, C., Newman, C., & DiMarco, G. (2002). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori & C. Teddie (Eds.), *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: Sage Publications.

Nunnally, J. (1972). Educational measurement and evaluation. New York: McGraw-Hill.

Quintana, S., & Maxwell, S. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist, 27*(4), 485-527.

Ridenour, T. A., Marchant, G. J., & Newman, I. (1994, August). *Interaction effects in path analysis in psychology.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Robinson, D.H., & Levin, J.R. (1997). Reflection on statistical and substantive significance, with a slice of replication. *Educational Researcher, 26*(5), 21-26.

Robitchek, C., & Kashubeck, S. (1999). A structural model of parental alcoholism, family functioning, and psychological health: The mediating effects of hardiness and personal growth orientation. *Journal of Counseling Psychology, 46*(2), 159-172.

Rosenthal, R., & Rosnow, R. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education, 61*(4), 293-316.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Thompson, B. (1994, February). *Why multivariate methods are usually vital in research: Some basic concepts.* Paper presented as a Featured Speaker at the biennial meeting of the Southwestern Society for Research in Human Development (SWSRHD), Austin, TX.

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher, 26*(5), 29-32.

Thompson, B. (1998). Review of what if there were no significance tests? *Educational and Psychological Measurement, 58*(2), 334-346.

Thompson, B. (1999a). If statistical significance tests are broken/misused, what practices should supplement or replace them? *Theory and Psychology, 9*(2), 165-181.

Thompson, B. (1999b). Statistical significance tests, effect-size reporting and the vain pursuit of pseudo-objectivity. *Theory and Psychology, 9*(2), 191-196.

Thompson, B. (1999c). Journal editorial policies regarding statistical significance tests: Heat is to fire as p is to importance. *Educational and Psychological Review, 11*(2), 157-169.

Tuason, M., Friedlander, M. (2000). Do parents' differentiation levels predict those of their adult children? And other tests of Bowen theory in a Philippine sample. *Journal of Counseling Psychology, 47*(1), 27-35.

Dr. Isadore Newman is a Distinguished Professor at the University of Akron, and a Professor in the Department of Educational Foundations and Leadership where he teaches research methods, advanced statistics and Ph. D. research seminars. He has also held the Harrington Endowed chair. Dr. Newman has writen several books on qualitative and quantitative research methods and has over 100 published articles.

Dr. Carole Newman is an Associate Professor in the Department of Curricular and Instructional Studies at the University of Akron where she teaches graduate and undergraduate courses in teaching methods. She is an educational consultant for area school districts and educational service centers.

John W. Fraas is a professor of business administration and a Trustees' Professor at Ashland Universitiy. He was

the initial appointee to the position of Trustees' Professor, a position awarded by the University's Board of Trustees to faculty members who have excelled in the areas of teaching, scholarship, and University service.

Russell Brown is completing his doctorate in Counseling Psychology at The University of Akron. He is currently an adjunct faculty member at Cleveland State University and the University of Akron where he teaches Statistics, Research Methodology and Developmental Psychology.