

# Journal of Research in Education

Fall 2004

An International Journal Published  
by the Eastern Educational Research Association

## CONTENTS

Editor's Comments  
Darla Twale

*Conflict Management Strategies of Academic Chairs*  
Jay R. Dee, Alan B. Henkin, Carole A. Singleton, & Kelly M. Mack

*The Comprehensive of Expository Text: A Comparative Study*  
Deena K. Bernstein, Pearl L. Seidenberg, & Sandra Levey

*The Impact of Online Communication Design on Student Motivation and Performance*  
Leping Liu & Paul Jones

*A Study of Faculty Beliefs for the Teaching of Diversity in Colleges of Education: An Explanatory Mixed-method Approach*  
Xin Liang, Susan Colville-Hall, Suzanne MacDonald & Lynn Atkinson Smolen

*The University Volunteer Tutoring Experience: The Who, What, and Why of Working with Urban Elementary School Tutees*  
Gary W. Ritter, Kristin K. Higgins, & Rebecca A. Newgent

Methodological Issues  
David M. Shannon Guest Editor

*Expressing Logistic Regression Coefficients as De/ta-p Values: Presenting Information to Decision Makers*  
John W. Fraas & J. Michael Drushal

*Learning Styles and Mindstyles: An Examination of Validity*  
Maria Martinez Witte

*Data Analysis Techniques in Qualitative Research*  
Barbara B. Kawulich

## CONFLICT MANAGEMENT STRATEGIES OF ACADEMIC UNIT CHAIRS

**Jay R. Dee**  
**University of Massachusetts Boston**  
**Alan B. Henkin**  
**The University of Iowa**  
**Carole A. Singleton**  
**University of Northern Iowa**  
**Kelly M. Mack**  
**University of Maryland-Eastern Shore**

*Chairs lead academic units with changing student needs, high level faculty expectations for advocacy, major fiscal problems, and unstable and uncertain work environments. They are expected to manage effectively the day-to-day operations of units and, concurrently, deal with numerous other role expectations of internal constituencies and external organizational referents. Conflict, an organizational inevitability, is one of the only certainties under these conditions of work. Chairs in this study showed a preference for conflict management strategies that emphasize collaboration and compromise when dealing with faculty, and facilitation, integration, and compromise when dealing with upper-level administrators.*

Colleges and universities depend on academic unit chairs to provide the leadership needed to sustain and strengthen educational programs. Chairs are expected to provide curriculum leadership, support scholarship and service, hire and mentor new faculty, and assess student outcomes and faculty performance (Lucas, 1994; Sanders, Ward, & Bensimon, 2000; Stark, Briggs, & Rowland-Poplowski, 2002). They are called upon to communicate and interface effectively with line administrators, other academic units, and external organizational referents. Concurrently, chairs must manage the day-to-day operations of academic units in organizational environments that are increasingly turbulent, uncertain, and competitive (Denhardt, Denhardt, & Aristigueta, 2002; Ratcliff, Lubinescu, & Gaffney, 2001; Winston & Zimmerman, 2000).

Chairs endeavor to respond to environmental pressures, align unit goals with the mission of the institution, and coordinate interdependent activities related to the end purposes of the unit, its members, and the organization; increasingly under conditions of financial exigency. The only certainty in this scenario is conflict. Personal differences, information deficiencies, role incompatibilities, and environmental stress may be counted among the numerous sources of conflict in academic units (Whetton & Cameron, 1998). Academic units may be involved in continuous conflict, as they interpret and adapt to changing conditions of work (McCabe & Dutton, 1993). Where there are few slack resources, academic unit chairs and members may confront instances where the interests of people conflict with the interests of the larger organization. Institutional mandates for efficiency and coordination of effort, for example, may function as additional stressors that precipitate high levels of conflict (Collins, 1975; Folger & Poole, 1984).

Conflict management has been identified as a critical skill for academic unit leadership (Gmelch, 1995; Gmelch & Carroll, 1991; Lucas, 1994). The conflict management behaviors of chairs may define, in part, the extent to which academic units are successful in efforts to respond to organizational and environmental change, to resolve

operational problems, to acquire needed resources, and to develop and implement strategies in competitive markets. The challenges of conflict management, however, often exact a toll on organizational role incumbents in terms of increased stress and dissatisfaction with work (Gmelch & Burns, 1994).

Responsibilities for effective conflict management in the context of turbulent organizational and external environments may be redefining the role of academic unit leadership in ways that make the position increasingly unattractive. Chairs are called upon to address a range of non-routine, poorly-defined problems, and may expect to encounter incompatibilities in relationships that may become more pronounced under conditions of task interdependence and asymmetric power distribution (Cramer, 1991; Daft, 1998; Thomas & Pondy, 1977). Gmelch and Burns (1994) found that conflict mediation, among other roles of the chair, produced the highest level of stress. Chairs expressed high levels of role-related dissatisfaction, similarly, when questioned about responsibilities for dealing with problems of inter-faculty conflict (Gmelch, 1991).

Chairs depend on a spectrum of managerial alternatives as they endeavor to respond to the demands of conflict-laden work environments. In this study, we examine conflict management behaviors among social work unit chairs in the US. We focus on social work units for several reasons. First, academic units with “soft” disciplinary paradigms, such as social work, may be counted among those areas of study where internal and external conflicts become particularly acute. “Soft” paradigms are characterized by contested curricular goals, and lack of consensus regarding what “counts” as research and service (Anastas & Congress, 1999; Hearn & Anderson, 2002). Related conflicts may arise when academic units seek to define criteria for promotion and tenure, allocate economic rewards and symbolic recognition of faculty efforts, and determine performance expectations for students.

Second, conflicts may arise when academic unit autonomy is constrained by interdependent relationships with external professional associations and community organizations (Braskamp & Wergin, 1998; Pfeffer & Salancik, 1978). Social work programs are subjected to oversight from professional accreditation and regulatory bodies such as the Council of Social Work Education (CSWE), the National Association of Social Workers (NASW), and state boards of licensure. Social work chairs are expected to respond effectively to mandates, agendas, and priorities of these influential groups and, concurrently, to local referents including professional practitioner groups, social service agencies, agency-based student practicum supervisors, and organized advocacy associations representing populations served by social workers. Chairs are challenged to explain these external demands to faculty members, and translate environmental expectations in ways that mediate inevitable internal-external conflicts.

Finally, the culture of an academic field socializes faculty and chairs toward behavioral repertoires that are consistent with professional or discipline-based value systems (Robertson & Bean, 1998; Rosch & Reich, 1996; Smith & Stewart, 1999). Social work is a value-grounded profession (National Association of Social Workers, 1996) where professionals in the field operate within a highly moral context (Hasenfeld, 1992) regardless of practice venue. Chairs’ conflict management behaviors are likely to be consistent with norms and values prevalent in the profession.

In this study, we explore the roles of academic unit chairs, and offer a conceptual framework for understanding conflict management behaviors of chairs. Findings from a national survey of social work chairs are presented, and

reveal how chairs address conflict with faculty and with higher-level administrators. Findings are discussed in terms of the professional culture of social work, and recommendations for departmental leadership are offered.

### Roles of Academic Unit Chairs

Balderston (1995) suggested that academic unit chairs serve as “linking pins” in the organizational structure of higher education institutions. Chairs occupy roles at the confluence of faculty and administrative cultures (Birnbaum, 1988; Creswell, Wheeler, Seagren, Egly, & Beyer, 1990), and encounter a range of challenges associated with middle management. Faculty members, to cite one exemplar, have high-level expectations for support and advocacy (Garcia, 1997). Upper-level administrators, in contrast, seek systemic coherence, high levels of accountability, and justification for expenditures. Chairs, in linking roles, are challenged to allocate scarce resources among faculty members while simultaneously maintaining collegiality within the unit (Hubbell & Homer, 1997).

Chairs address manifest conflicts within and between two distinct authority systems; the professional authority system of faculty, which values autonomy, shared governance, and non-hierarchical leadership, and the administrative authority system, which adheres to principles of rational management (Birnbaum, 1988). A variety of skills and strategies may be necessary to address conflicts that emanate from different levels of the organization, especially where conflicting parties may have divergent views about the legitimate uses of power (Bess, 1988).

Academic unit chairs in social work, the focal population of this study, confront multiple, conflict-laden situations at the unit and organizational levels. They deal with the immediate problems of the unit, and concurrently represent the interests of programs that are differentiated from traditional disciplines in terms of emphases on applied models and professional practice. These chairs are called upon to mediate internal conflicts associated with manifest differences in faculty orientations toward theoretical, applied, and practical aspects of programs (Anastas & Congress, 1999).

The social work profession leaves a substantial imprint on the service technologies used by related organizations (Hasenfeld, 1992). Practitioners in the field expect experts in the academy to provide professional services, insights, guidance, and preparation that will enable successful interventions with a spectrum of client groups. Concurrently, colleges and universities often expect conformity with traditional patterns of instructional and service delivery (Bess, 1988). Social work chairs, under these circumstances, are likely to encounter conflicts involving internal pressures for coordination and coherence, and external expectations for responsiveness and localized adaptation to unique circumstances.

Social work chairs endeavor to support the standards stipulated in the profession’s (NASW) Code of Ethics that serve as guidance for professional conduct. Research has affirmed that the Code itself includes components with potential to induce conflict (Congress, 2000; Reamer, 1999). Academic unit chairs may become frontline participants in institutional conflicts involving policies that may contravene enactment of commitments to larger social needs associated with professional ethics. Conflicts may arise over issues of social justice, especially where institutional policies may be asynchronous with the profession’s core belief system and/or challenge the foundations

of an academic program and its curriculum (Reamer, 1999). Conflicts that occur within academic environments may involve internal disputes related to the broader world of professional practice over which unit chairs have minimal influence. As leaders of professional programs, moreover, chairs may confront personal contradictions where expectations of one role (chair) may be incompatible with expectations of another (social services professional). The high potential for conflict under conditions of work at the academic unit level suggests the importance of chairs' conflict management strategies.

### Theoretical Framework

Organizational conflict is frequently defined as an overt struggle between two or more parties in an organization, typically for scarce resources (Coser, 1956; Dahrendorf, 1958; Price, 1997). Katz and Kahn (1978) suggested that conflict is an interaction "marked by efforts at hindering, compelling, or injuring and by resistance or retaliation against those efforts" (p. 615). Conflict may also be subtle and couched in terms of values, cultural norms, and non-rational elements of organizational life (Kolb & Putnam, 1992). Despite the negative connotations of these definitions, conflict can have either positive or negative effects on organizations and academic units, depending upon how it is addressed (Robbins, 1974; Senge, 1990). Conflict can help academic units define and resolve problems, increase unit cohesion, keep faculty alert to others' interests, and enable the unit to function as a team (Eckel, 1998; Gmelch, 1995). Alternatively, poorly regulated conflict can yield interpersonal animosity, vengeful behaviors, sabotage, and other dysfunctional outcomes.

This study suggests that conflict management behaviors can be categorized using one of two models: contingency or enactment. The contingency model suggests that particular types of conflict management behaviors will be more effective under certain conditions than others (Blake & Mouton, 1964; Coser, 1967; Deutsch, 1973; Thomas, 1976). Leaders respond to a "given" conflict scenario, and seek to select strategies that are best suited to deal with a set of specific circumstances.

Compromise, for example, can be viewed as a contingency approach to conflict management. The positions of conflicting parties are taken as "givens," and leaders seek to split the difference between divergent positions (Thomas, 1976). Similarly, strategies of control and bureaucratic forcing (Blake & Mouton, 1964) are consistent with the contingency model. Here, leaders use formal rules to resolve disagreements; leaders identify or create rules that fit a given set of conflict circumstances.

Contingency approaches have the advantages of efficiency and uniformity. Formalized conflict management practices can serve as legitimate devices for quickly resolving disagreements that emerge with frequency and regularity. They may be less useful, however, in addressing conflicts induced by complex forces in the external environment (Sambamurthy & Poole, 1992). Conflicts that are external in origin are often ambiguous, and are not likely to be resolved through reliance on rules or by "splitting the difference" between opposing viewpoints. Where academic unit chairs take the conflict episode as a "given," they may forestall opportunities to reconcile conflicting information and positions. The benefits of broader perspective may not be realized when contingency models prevail.

The contingency model may be appropriate for addressing overt conflict, where the conflicting parties can be identified easily, and where issues and positions can be specified clearly. When conflict is subtle and based in differences over values, however, it may be difficult to identify positions and specify alternatives (Kolb & Putnam, 1992). Under such conditions, academic unit chairs may find that the enactment model yields more useful insights.

Enactment theory (Weick, 1969, 1995) suggests that conflict is a social construction created by organizational members. The conflict situation is not taken as a “given.” Instead, organizational members shape the parameters of conflict through processes of sensemaking, where conflicting parties and positions are continually redefined (Seedorf, 1992). Collaboration, for example, may be conceptualized as a conflict management strategy that is consistent with the enactment model. Organizational members seek to develop a synthesis of conflicting arguments. Elements from divergent perspectives are merged, re-formed, and re-developed; something new is created (Senge, 1990).

Enactment strategies enable the emergence of synergy. Participants develop responses to conflict that would not have emerged through compromise or rule-based resolution (Bazerman & Neale, 1992). Enactment approaches, however, are more time-consuming than contingency approaches, and may require extensive commitments of time and personnel resources (Faerman, 1996). Social constructions of conflict, moreover, may lack the legitimacy of an “objective,” rule-based response. The credibility of the conflict management process is, consequently, critical.

#### Method

In this study, we address two research questions in an effort to improve our understanding of how social work chairs utilized contingency and enactment approaches in conflicts with faculty and higher-level administrators.

1. What conflict management strategies do academic unit chairs use when dealing with faculty and administrators?
2. Do these conflict management strategies differ by demographic and institutional characteristics?

The study universe included all chairs of social work academic units in the US. Analyses are based on data derived from 190 instruments that were completed and returned. The 40% response rate falls within an acceptable range for mailed social science surveys in this category of research (Hager, Wilson, Pollack, & Rooney, 2002). Respondents completed the Organizational Communication Conflict Instrument (OCCI; Putnam & Wilson, 1982). The 29-item OCCI provides self-reports of behaviors used to define, participate in, and manage conflict. The Likert-type response continuum ranges from 1=never to 7=always.

The OCCI conceptualizes conflict as a communicative process. Putnam and Wilson (1982) focused on communication behaviors rather than conflict management “styles,” which refer to individual dispositions or habits (Hall, 1969). Measures of conflict management style do not take into account contingent, situational factors in the work environment, and provide only limited insight into enacted strategies. In contrast, the OCCI provides a context-specific measure that includes items that pertain to both contingency and enactment models of conflict management.

The OCCI reflects assumptions from the contingency model when it suggests that “the decision to use a particular conflict strategy is largely governed by situational rather than personality constraints” (Putnam & Wilson,

1982, p. 633). This assumption suggests that there is no “one best way” to manage conflict. Instead, effective conflict management strategies are contingent on the situation. The OCCI also incorporates assumptions from the emergent model. Conflicts are conceptualized as disagreements that “entail deep-seated latent differences rather than surface misinterpretations or semantic quibbles” (Putnam & Wilson, 1982, p. 633). Disagreements are socially constructed by organizational members who vary in their access to information and power, and who may interpret the social world in radically different ways.

A common limitation of conflict management instruments is linked to influences of social desirability. Specifically, respondents may self-report frequent use of strategies that they think will be received favorably, or will be viewed as socially acceptable. Putnam and Wilson (1982) found, however, only moderate correlations between the OCCI items and indices of social desirability. They asked respondents in their study to indicate the degree of desirability of each behavior, and the frequency with which they engage in that behavior. The item-desirability correlation coefficients ranged from .11 to .48; much lower than those found in previous questionnaires (Thomas & Kilmann, 1975).

Content validity refers to whether an instrument’s items adequately sample the behavioral strategies used to manage conflict. Assessments of content validity showed that the OCCI was superior to related instruments in that it provided more information about communication strategies than any other available instrument (Wilson & Waltman, 1988). In terms of reliability, previous tests for internal consistency reported alpha coefficients that ranged from .82 to .93 (Putnam & Wilson, 1982).

Survey instructions requested that respondents think of disagreements they have encountered with faculty and upper-level administrators. Then, they were asked to indicate how frequently they engaged in each of the behaviors described in the OCCI. The order of the faculty and administrator items was alternated to minimize potential for response bias; half of the surveys were printed with the faculty items first, and half were printed with the administrator items first. The questionnaire included additional items to obtain individual and institutional data.

#### *Sample*

The majority (61.0%) of respondents was female. Age data showed that respondents were in their 30s (3.3%), 40s (21.4%), 50s (51.1%), and 60s+ (24.2%). Respondents reported the number of years that they have served as a chair: 3 years or less (22.0%), 4 to 6 years (23.7%), 7 to 9 years (15.6%), and 10 years or more (38.7%).

Respondents represented institutions with a range of enrollment sizes: less than 1,000 students (7.8%), 1,000 to 2,000 (11.7%), 2,001 to 5,000 (21.7%), 5,001 to 10,000 (20.0%), 10,001 to 20,000 (24.4%), and more than 20,000 (14.4%). One-third of respondents worked at institutions with a faculty union.

#### *Procedures*

Principal components factor analysis (varimax rotation) was used to identify behavioral dimensions of conflict management reported by chairs in response to disagreements emanating from faculty and higher-level administrators. Factor analysis identifies the smallest number of common factors that “best explain or account for the correlations among the indicators” (Sharma, 1996, p. 99). This procedure revealed two sets of conflict management factors that parsimoniously explained chairs’ conflict management behaviors toward faculty and

higher-level administrators. Items that did not load .50 or higher on any of the factors were eliminated from subsequent analyses.

Tests and one-way analyses of variance were computed to determine whether conflict management strategies varied by gender, age, years of experience as chair, institutional size, or faculty unionization status.

The risk of committing a Type I error (claiming that an effect is significant when, in fact, it is not) increases with each hypothesis tested among the same set of subjects (Keppel, 1991). In this study, 10 hypotheses (five faculty conflict factors and five administrator conflict factors) were tested for each independent variable. To minimize vulnerability to Type I error, an alpha of .01 was used for all tests of statistical significance.

### Results

Factor analysis (see Table 1) revealed five strategies for dealing with faculty conflict:

- Avoidance (7 items loading .50 or higher)--a contingency strategy where the person withdraws or postpones dealing with a given scenario (explained variance 21.9%, alpha reliability .89)
- Control (4 items loading .50 or higher)--a contingency strategy which forces a given conflict into a bureaucratic protocol (explained variance 5.8%, alpha reliability .70)
- Compromise (3 items loading .50 or higher)--a contingency strategy where the person "splits the difference" between conflicting positions (explained variance 4.4%, alpha reliability .71)
- Collaboration (4 items loading .50 or higher)--an enactment strategy where the person synthesizes divergent viewpoints and develops a new understanding of the problem (explained variance 13.2%, alpha reliability .72)
- Minimize (2 items loading .50 or higher)--a contingency strategy where the person attempts to downplay the significance of a given conflict (explained variance 8.8%, alpha reliability .71)

The five factors accounted for a total of 54.1% of the variance in chairs' conflict management responses to faculty. Nine of the original OCCI items did not load .50 or higher on any of the five faculty conflict factors; these items were eliminated from the analysis.

Chairs reported engaging in collaboration (mean 5.35, s.d. 0.55) most frequently, followed by compromise (mean 4.81, s.d. 0.63), minimize (mean 3.04, s.d. 0.79), control (mean 2.94, s.d. 0.81), and avoidance (mean 2.79, s.d. 0.77).



Table 1

*Faculty Conflict Factors*

Items	Factor Loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
<b>Avoidance</b>					
3. I shy away from topics that are sources of dispute	.796	-.069	.156	.073	.108
5. I steer clear of disagreeable situations.	.756	-.012	.044	.032	.197
7. I avoid a person I suspect of wanting to discuss a disagreement.	.537	-.092	.174	.292	-.079
12. I keep quiet about my views in order to avoid disagreements.	.707	-.079	.199	-.119	-.036
23. I withdraw when someone confronts me about a controversial issue.	.676	-.098	.264	.161	-.112
24. I sidestep disagreements when they arise.	.756	-.073	.360	.120	.064
25. I try to smooth over disagreements by making them appear unimportant.	.572	-.028	.560	.226	.081
<b>Collaboration</b>					
4. I suggest solutions that combine a variety of viewpoints.	-.074	.510	.238	-.051	.255
11. I offer creative solutions in discussions of disagreements.	-.193	.577	.030	.015	.018
20. I try to use everyone's ideas to generate solutions to disagreements.	.102	.709	-.235	-.181	.105
15. I reduce disagreements by saying that they are insignificant.	.362	-.098	.587	.034	.038
29. I ease conflict by claiming that our differences are trivial.	.444	-.048	.623	.046	.105
<b>Control</b>					
3. I insist that my position be accepted during a conflict.	.144	-.059	-.056	.516	-.106
17. I assert my opinion forcefully.	-.108	.117	-.002	.613	-.087
18. I dominate arguments until the other person understands my position.	.221	-.023	.104	.748	-.007
22. I argue insistently for my stance.	-.023	-.040	-.017	.617	.134
<b>Compromise</b>					
6. I give in a little on my ideas when the other person also gives in.	-.024	.081	.026	-.033	.599
9. I will go fifty-fifty to reach a settlement.	.167	.204	.249	-.108	.587
13. I frequently give in a little if the other person will meet me halfway.	.097	.212	.103	.049	.773

Factor analysis (see Table 2) also revealed a five-factor structure for conflict with higher-level administrators. Two factors were common to chairs' responses for both faculty and administrator conflict: compromise (4 items) and control (3 items). Responses related to collaboration diverged into two factors: facilitation (2 items) and integration (2 items). Facilitation involves bringing parties together to solve problems. Integration refers to the

synthesis of divergent viewpoints to develop a new understanding of the problem. Factor analysis indicated a “passive behavior” factor (10 items) that included both avoidance and minimization strategies.

The “passive behavior” factor explained the most variance (21.5%), and had acceptable reliability (alpha .90). The “compromise” factor explained 13.9% of the variance, and the alpha reliability coefficient was .74. The “control” factor accounted for 8.7% of the variance, and had an alpha reliability coefficient of .69. The “facilitation” factor explained 5.3% of the variance, and its corresponding alpha reliability coefficient was .74. Finally, “integration” accounted for 4.8% of the variance, and its alpha reliability coefficient was .64.

The five-factor structure explained a total of 54.2% of the variance in chairs’ conflict management behaviors with higher-level administrators. Eight of the original OCCI items did not load .50 or higher on any of the five administrator conflict factors, and these items were eliminated from the analysis.

Chairs reported engaging in facilitation (mean 5.59, s.d. 0.74) most often, followed by integration (mean 5.19, s.d. 0.71), compromise (mean 4.74, s.d. 0.67), passive behavior (mean 2.86, s.d. 0.80), and control (mean 2.79, s.d. 0.80).

Table 2

*Administrator Conflict Factors*

Items	Factor Loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
<b>Passive (Avoid/Minimize)</b>					
2. I shy away from topics that are sources of disputes.	.779	-.006	-.061	.030	-.043
5. I steer clear of disagreeable situations.	.763	-.007	.054	.017	.019
7. I avoid a person I suspect of wanting to discuss a disagreement.	.505	.085	.168	.002	-.293
12. I keep quiet about my views in order to avoid disagreements.	.719	-.022	-.121	.040	.030
15. I reduce disagreements by saying that they are insignificant.	.566	.202	.133	-.306	-.100
23. I withdraw when someone confronts me about a controversial issue.	.694	-.071	.076	-.120	-.174
24. I sidestep disagreements when they arise.	.779	.076	.101	-.132	-.108
25. I try to smooth over disagreements by making them appear unimportant.	.723	.160	.142	-.324	-.054
28. I hold my tongue rather than argue.	.594	.138	-.267	.040	-.095
29. I ease conflict by claiming that our differences are trivial.	.577	.280	.025	-.268	-.035
<b>Compromise</b>					
21. I offer tradeoffs to reach solutions in a disagreement.	.004	.506	.026	.323	.158
<b>Control</b>					
17. I assert my opinion forcefully.	-.167	.044	.647	.059	.001
18. I dominate arguments until the other person understands my position.	.195	-.130	.668	-.145	.063
22. I argue insistently for my stance.	-.079	.108	.667	.082	-.109
<b>Facilitation (collaboration)</b>					
19. I suggest that we work together to create solutions to disagreements.	-.240	.200	-.020	.662	.132
20. I try to use everyone's ideas to generate solutions to disagreements.	-.067	.175	-.019	.686	.250
<b>Integration (collaboration)</b>					
1. I blend my ideas with others to create new alternatives for resolving a conflict.	-.074	.089	.027	.086	.609
4. I suggest solutions that combine a variety of viewpoints.	-.129	.204	-.122	.079	.654

In terms of the second research question, conflict management behaviors did not differ significantly on the basis of respondent characteristics or institutional variables. T tests and one-way ANOVAs (see Table 3) for the five faculty conflict factors and the five administrator conflict factors showed no significant differences by gender, age, or years of experience as chair. Similarly, conflict scores did not differ by institutional size or faculty unionization status.

Table 3

*Significance Tests with Conflict Factors and Individual/Institutional Variables*

Conflict Factor	Significance Test				
	Gender	Age	Years of Experience	Institutional Size	Union Status
Faculty: Avoidance	t=0.38 (p=.708)	F=0.97 (p=.409)	F=0.63 (p=.600)	F=0.67 (p=.649)	t=-0.01 (p=.997)
Faculty: Collaboration	t=1.65 (p=.102)	F=0.57 (p=.638)	F=0.82 (p=.482)	F=1.47 (p=.203)	t=0.82 (p=.415)
Faculty: Minimize	t=0.01 (p=.993)	F=0.49 (p=.689)	F=1.21 (p=.309)	F=0.45 (p=.811)	t=0.65 (p=.517)
Faculty: Control	t=0.77 (p=.445)	F=0.71 (p=.550)	F=1.07 (p=.365)	F=1.47 (p=.203)	t=-0.79 (p=.432)
Faculty: Compromise	t=0.67 (p=.505)	F=0.23 (p=.878)	F=1.25 (p=.294)	F=2.42 (p=.038)	t=0.21 (p=.835)
Administrators: Passive Behavior	t=1.57 (p=.118)	F=1.69 (p=.171)	F=0.65 (p=.582)	F=0.83 (p=.533)	t=0.19 (p=.853)
Administrators: Compromise	t=0.87 (p=.387)	F=0.41 (p=.748)	F=0.81 (p=.490)	F=1.74 (p=.128)	t=0.87 (p=.386)
Administrators: Control	t=0.46 (p=.646)	F=0.13 (p=.945)	F=0.52 (p=.669)	F=1.10 (p=.363)	t=-0.65 (p=.518)
Administrators: Facilitation	t=1.39 (p=.167)	F=0.34 (p=.795)	F=0.43 (p=.732)	F=1.30 (p=.265)	t=-0.01 (p=.995)
Administrators: Integration	t=0.83 (p=.407)	F=0.95 (p=.416)	F=1.79 (p=.151)	F=0.93 (p=.465)	t=0.76 (p=.447)

## Discussion

Academic unit chairs in this study relied, primarily, on two strategies to address conflict with faculty: collaboration and compromise. Chairs utilized both enactment and contingency approaches in dealing with faculty conflict. While chairs sought to synthesize divergent ideas and create new interpretations of conflict scenarios, they also reported “splitting the difference” among opposing positions.

The profile of conflict with upper-level administrators revealed three preferred strategies: facilitation, integration, and compromise. Through facilitation, chairs sought to bring parties together to work out solutions. Additionally, they utilized an enactment strategy (integration) that encouraged administrators to seek new interpretations of disagreements. Though enactment strategies were prominent, a preference was shown for compromise where chairs would work to get opposing parties to meet half way.

Chairs used passive conflict management strategies (avoidance and minimization) less often than enactment approaches. They seldom used control strategies with either faculty or administrators.

Why may certain conflict management strategies be preferred by respondents in this study? The choices and preferred order of conflict management strategies used by study respondents may be linked, in part, to prior socialization to this profession (Robertson & Bean, 1998; Rosch & Reich, 1996). Preferences for collaborative approaches may be associated with preferred alternatives in social work professional practice. Such preference

accorded to collaborative approaches with both faculty and administrators--where leaders enact, shape, and reframe conflict scenarios, for example--may find associations and parallelisms in social work practice and therapeutic literatures (Egan, 2003; Minuchin, 1974; Minuchin & Fishman, 1981; Minuchin & Nichols, 1993; Neukrug, 2000; Watzlawick, Weakland, & Fisch, 1974) that provide models for professional practitioners to use as they endeavor to help clients reassess, reformulate, and create perspectives for problem solving.

Additional linkages to professional socialization may be suggested. Successful professional practice in helping relationships in social work requires certain "facilitative capacities" (Carkhuff, 1969; Ivey & Authier, 1978; Marshall, Charming, & Bell, 1979; Rogers, 1957). Emphases on the development of facilitative capacities involving empathy, respect, and authenticity are common in both social work education and practice. Similar elements appear to be prominent in collaborative approaches to conflict management, and may be adapted to preferred managerial repertoires of academic unit chairs in social work. Collaborative approaches to conflict management with administrators may be tempered, for example, by professionally linked dispositions toward facilitation (bringing parties together to solve problems) and integration (synthesizing divergent viewpoints into a new comprehension of the problem).

Academic unit chairs' preferences for compromise, another frequently elected management alternative, may be associated, in part, with professional preparation for work in communities, for work with families and individuals, and with socialization to the profession. In community and group work, for example, leaders and participants/clients commonly engage in give and take negotiations in terms of goal setting, contractual agreements, and in the implementation of change strategies (Egan, 2003; Neukrug, 2000). Chairs' choices of compromise as a conflict regulation strategy appear consistent, once again, with professional training, practice experience, and dispositions acquired in the course of academic preparation.

## Conclusion

Academic unit chairs may be viewed as leaders who manage conversational inquiry in order to engage colleagues in efforts that assure a sustainable future for the educational program. Skills required for effective performance include well-honed communication skills, cultural management skills, problem-solving skills, and conflict management skills, among others (Bowman, 2002; Hecht, Higgerson, Gmelch, & Tucker, 1999; Lippitt, 1982). Skillful execution of conflict management strategies can contribute to the development and maintenance of academic units distinguished by productive, healthy tensions; venues where faculty members can address professional goals and, at the same time, manage complex institutional and contextual dependencies.

Effective academic unit leaders are challenged to manage conflict at the interface of numerous organizational and environmental pressures. The vagueness and generality of most external demands may provide academic chairs with some level of latitude to enact responses that serve the best interests of their unit (Garcia, 1997). While external referents make numerous requests on academic units, the details of implementation are often left to the discretion of the unit. Chairs may have considerable leeway in terms of interpreting and shaping expectations, and may find advantage in enactment strategies that allow for the construction of conflict situations that facilitate consensus-building and functional outcomes (Helgesen, 1995; Senge, 1990).

Enactment approaches to conflict management may work best when chairs share information, promote a sense of trust, and emphasize solution-finding processes. Gmelch and Carroll (1991) described a conflict management process in which the chair focuses on interests rather than positions, and generates a variety of alternative solutions. Similarly, Lucas (1994) suggested that chairs assume facilitative roles to manage conflict. Such practices can reinforce beliefs in the fairness of outcomes and, concurrently, enable conflicting parties to feel that they have some control over the process.

Enactment may be a useful strategy for addressing complex conflicts of purpose, role, and identity. Conflict management behaviors, reported by respondents in this study, suggest that these chairs value collaborative approaches to leadership that appear compatible with characteristics of postsecondary institutions; specifically, shared governance, decentralization of power and authority, fluid participation in decision making, and an absence of clear performance indicators (Bess, 1988; Birnbaum, 1988; Cohen & March, 1974). Academic chairs may effectively advance the mutual interests of their institutions and departments by enacting new definitions of recurring problems, and by diminishing the impact of traditional dichotomies, such as autonomy and accountability, that tend to divide people into factions and rival constituencies.

Chairs work in fluid environments at the confluence of demands for organizational coherence, control, and accountability, on one hand, and professional autonomy, on the other. They may expect to confront a range of tensions and conflicts emanating, in part, from asymmetries in interests and values. Conflict, under such circumstances, may be regulated and managed by the chair; however, it is never resolved. Organizational conflict is both inevitable and enduring (Bolman & Deal, 1997).

The literature on conflict management in academic units remains limited. Previous empirical studies of conflict and academic unit chairs (Gmelch & Carroll, 1991; Roach, 1991; Wolverson, Gmelch, Wolverson, & Sarros, 1999) have not differentiated between the behaviors chairs use to manage conflict with faculty, and those employed to address disagreements with higher-level administrators. The extant literature, moreover, has not explored relationships between professional and discipline-based value systems and the conflict management behaviors utilized by academic unit chairs. This study revealed similar patterns and chair preferences for dealing with conflicts emanating from faculty and conflicts with higher-level administrators. Conflict management preferences for collaboration and compromise among these respondents appear to be informed by norms and values prevalent in the social work profession.

Research on academic chairs and conflict management should be extended in studies that focus on other disciplines and institutional types. Findings in this study revealed consistency between chairs' conflict management behaviors and the professional norms of social work. The study suggests that socialization experiences in the professional field or discipline may inform chairs' approaches to managing conflict. These findings suggest a need to examine how chairs form their perceptions of academic leadership and conflict management, and whether associated behaviors can be explained by professional socialization experiences. Future studies should also explore the effects of institutional socialization, and how the norms and values of an entire institution shape academic chair leadership and responses to conflict. This line of research may reveal strategies and alternatives for the

reconciliation of professional values and institutional expectations, and enable higher levels of commitment to goals and efforts that are personally and professionally meaningful.

### References

- Anastas, J., & Congress, E. (1999). Philosophical issues in doctoral education in social work: A survey of doctoral program directors. *Journal of Social Work Education, 35*(1), 143-153.
- Balderston, F. (1995). *Managing today's university: Strategies for viability, change, and excellence*. San Francisco: Jossey-Bass.
- Bazerman, M., & Neale, M. (1992). *Negotiating rationally*. New York: Free Press.
- Bess, J. (1988). *Collegiality and bureaucracy in the modern university*. New York: Teachers College Press, Columbia University.
- Birnbaum, R. (1988). *How colleges work: The cybernetics of academic organization and leadership*. San Francisco: Jossey-Bass.
- Blake, R., & Mouton, J. (1964). *The managerial grid*. Houston: Gulf Publishing.
- Bolman, L., & Deal, T. (1997). *Reframing organizations: Artistry, choice, and leadership* (2<sup>nd</sup> ed.). San Francisco: Jossey-Bass.
- Bowman, R. (2002). The real work of department chairs. *Clearing House, 75*(3), 158-162.
- Braskamp, L., & Wergin, J. (1998). Forming new social partnerships. In W. Tierney (Ed.), *The responsive university: Restructuring for high performance* (pp. 62-91). Baltimore: Johns Hopkins University Press.
- Carkhuff, R. (1969). *Helping and human relations*. New York: Holt, Rinehart and Winston.
- Cohen, M., & March, J. (1974). *Leadership and ambiguity: The American college president*. New York: McGraw-Hill.
- Collins, R. (1975). *Conflict sociology*. New York: Academic Press.
- Congress, E. (2000). What social workers should know about ethics: Understanding and resolving practice dilemmas. *Advances in Social Work, 1*(1), 1-42.
- Coser, L. (1956). *The functions of social conflict*. Glencoe, IL: Free Press.
- Coser, L. (1967). *Continuities in the study of social conflict*. New York: Free Press.
- Cramer, R. (1991). Intergroup relations and organizational dilemmas. In L. Cummings & B. Staw (Eds.), *Research in organizational behavior* (pp. 191-228). New York: JAI Press.
- Creswell, J., Wheeler, D., Seagren, A., Egly, N., & Beyer, K. (1990). *The academic chairperson's handbook*. Lincoln, NE: University of Nebraska Press.
- Daft, R. (1998). *Organization theory and design*. Cincinnati, OH: South-Western Publishing.
- Dahrendorf, R. (1958). *Class and class conflict in industrial society*. Stanford, CA: Stanford University Press.

- Denhardt, R., Denhardt, J., & Aristigueta, M. (2002). *Managing human behavior in public and nonprofit organizations*. Thousand Oaks, CA: Sage
- Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven, CT: Yale University Press.
- Eckel, P. (1998). Thinking differently about academic departments: The academic department as a team. *New Directions for Institutional Research*, 25(4), 27-38.
- Egan, G. (2003). *The skilled helper: A problem-management and opportunity-development approach to learning*. Pacific Grove, CA: Brooks/Cole.
- Faerman, S. (1996). Managing conflicts creatively. In J. Perry (Ed.), *Handbook of public administration* (pp. 632-646). San Francisco: Jossey-Bass.
- Folger, J., & Poole, M. (1984). *Working through conflict*. Dallas, TX: Scott, Foresman.
- Garcia, E. (1997). The chair as mediator: Balancing the new and the old while the winds roar. *ADFL Bulletin*, 28(3), 18-20.
- Gmelch, W. (1991, October). *The creation of constructive conflict within educational administration departments*. Paper presented at the annual meeting of the University Council for Educational Administration, Baltimore, MD.
- Gmelch, W. (1995). Department chairs under siege: Resolving the web of conflict. *New Directions for Higher Education*, 92, 35-42.
- Gmelch, W., & Burns, J. (1994). Sources of stress for academic department chairpersons. *Journal of Educational Administration*, 32(2), 79-94.
- Gmelch, W., & Carroll, J. (1991). The three Rs of conflict management for department chairs and faculty. *Innovative Higher Education*, 16(2), 107-123
- Hager, M., Wilson, S., Pollack, T., & Rooney, P. (2002, August). *Response rates for mail surveys on nonprofit organizations*. Paper presented at the annual meeting of the Academy of Management, Denver, CO.
- Hall, J. (1969). *Conflict management survey: A survey on one's characteristic reaction to and handling of conflict between himself and others*. Conroe, TX: Teleometrics.
- Hasenfeld, Y. (Ed.). (1992). *Human services as complex organizations*. Newbury Park, CA: Sage.
- Hearn, J., & Anderson, M. (2002). Conflict in academic departments: An analysis of disputes over faculty promotion and tenure. *Research in Higher Education*, 43(5), 503-529.
- Hecht, I., Higgerson, M., Gmelch, W., & Tucker, A. (1999). *The department chair as academic leader*. Phoenix, AZ: Oryx Press/American Council on Education.
- Helgesen, S. (1995). *The web of inclusion: A new architecture for building great organizations*. New York: Currency/Doubleday.
- Hubbell, L., & Homer, F. (1997). The academic department chair: The logic of appeasement. *PS: Political Science and Politics*, 30(2), 209-213.
- Ivey, A., & Authier, J. (1978). *Microcounseling: Innovations in interviewing, counseling, psychotherapy, and psychoeducation*. Springfield, IL: Thomas Press.



- Katz, D., & Kahn, R. (1978). *The social psychology of organizations* (2<sup>nd</sup> ed.). New York: John Wiley & Sons.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.
- Kolb, D., & Putnam, L. (1992). The dialectics of disputing. In D. Kolb & J. Bartunek (Eds.), *Hidden conflict in organizations: Uncovering behind-the-scenes disputes* (pp. 1-31). Newbury Park, CA: Sage.
- Lippitt, G. (1982). *Organizational renewal*. Englewood Cliffs, NJ: Prentice-Hall.
- Lucas, A. (1994). *Strengthening departmental leadership: A team-building guide for chairs in colleges and universities*. San Francisco: Jossey-Bass.
- Marshall, E., Charping, J., & Bell, W. (1979). Interpersonal skills training: A review of research. *Social Work Research and Abstracts*, 15, 10-16.
- McCabe, D., & Dutton, J. (1993). Making sense of the environment: The role of perceived effectiveness. *Human Relations*, 46(5), 623-643.
- Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.
- Minuchin, S., & Fishman, H. (1981). *Family therapy techniques*. Cambridge, MA: Harvard University Press.
- Minuchin, S., & Nichols, M. (1993). *Family healing*. New York: Free Press.
- National Association of Social Workers. (1996). *Code of ethics of the National Association of Social Workers*. Washington, DC: Author.
- Neukrug, E. (2000). *Theory, practice, and trends in human services*. Pacific Grove, CA: Brooks/Cole.
- Pfeffer, J., & Salancik, G. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper and Row.
- Price, J. (1997). Handbook of organizational measurement. *International Journal of Manpower*, 18, 303-558.
- Putnam, L., & Wilson, C. (1982). Communicative strategies in organizational conflict: Reliability and validity of a measurement scale. In M. Burgoon (Ed.), *Communication yearbook* (Vol. 6, pp. 629-652). Newbury Park, CA: Sage.
- Ratcliff, J., Lubinescu, E., & Gaffney, M. (2001). *How accreditation influences assessment: New directions for higher education*. San Francisco: Jossey-Bass.
- Reamer, F. (1999). *Social work values and ethics*. New York: Columbia University Press.
- Roach, K. (1991). University department chairs' use of compliance-gaining strategies. *Communication Quarterly*, 39(1), 75-90.
- Robbins, S. (1974). *Managing organizational conflict: A nontraditional approach*. Englewood Cliffs, NJ: Prentice-Hall.
- Robertson, L., & Bean, J. (1998). Women faculty in family and consumer sciences: Influences on job satisfaction. *Family and Consumer Sciences Research Journal*, 27(2), 167-193.
- Rogers, C. (1957). The necessary and sufficient conditions of psychotherapeutic personality change. *Journal of Consulting Psychology*, 21, 95-103.

- Rosch, T., & Reich, J. (1996). The enculturation of new faculty in higher education: A comparative investigation of three academic departments. *Research in Higher Education*, 37(1), 115-131.
- Sambamurthy, V., & Poole, M. (1992). The effects of variations in capabilities of GDSS designs on management of cognitive conflict in groups. *Information Systems Research*, 3, 224-251.
- Sanders, K., Ward, K., & Bensimon, E. (2000). *The department chair's role in developing new faculty into teachers and scholars*. Bolton, MA: Anker Publishing.
- Seedorf, R. (1992, April). *The problem solving role of the university department chair*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Senge, P. (1990). *The fifth discipline*. New York: Doubleday.
- Sharma, S. (1996). *Applied multivariate techniques*. New York: John Wiley & Sons.
- Smith, A., & Stewart, G. (1999). A statewide survey of new department chairs: Their experiences and needs in learning their roles. *New Directions for Community Colleges*, 27(105), 29-36.
- Stark, J., Briggs, C., & Rowland-Poplowski, J. (2002). Curriculum leadership roles of chairpersons in continuously planning departments. *Research in Higher Education*, 43(3), 329-356.
- Thomas, K. (1976). Conflict and conflict management. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 889-936). Chicago: Rand McNally.
- Thomas, K., & Kilmann, R. (1975). The social desirability variable in organizational research. *Academy of Management Journal*, 18, 741-752.
- Thomas, K., & Pondy, L. (1977). Toward an intent model of conflict management among principal parties. *Human Relations*, 30, 1089-1102.
- Watzlawick, P., Weakland, J., & Fisch, R. (1974). *Change: Principles of problem formation and problem resolution*. New York: W.W. Norton.
- Weick, K. (1969). *The social psychology of organizing*. Reading, MA: Addison-Wesley.
- Weick, K. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.
- Whetton, D., & Cameron, K. (1998). *Developing management skills*. Reading, MA: Addison-Wesley.
- Wilson, S., & Waltman, M. (1988). Assessing the Putnam-Wilson organizational communication conflict instrument. *Management Communication Quarterly*, 1, 367-388.
- Winston, G., & Zimmerman, D. (2000). Where is aggressive price competition taking higher education? *Change*, 32(4), 10-18.
- Wolverton, M., Gmelch, W., Wolverton, M. L., & Sarros, J. (1999). Stress in academic leadership: U.S. and Australian department chairs/heads. *Review of Higher Education*, 22(2), 165-185.
- Jay R. Dee is an Assistant Professor in the Graduate College of Education at the University of Massachusetts Boston, Alan B. Henkin is a Professor in the College of Education at the University of Iowa, Carol A. Singleton is a Professor in the Department of Social Work at the University of Northern Iowa, and Kelly M. Mack is an Associate Professor in the Department of Natural Sciences at the University of Maryland-Eastern Shore.*

*Correspondence concerning this article should be addressed to Dr. Alan B. Henkin, Professor, College of Education, University of Iowa, Iowa City, IA 52242.*

**Do not quote without permission of the authors.**

## THE COMPREHENSION OF EXPOSITORY TEXT: A COMPARATIVE STUDY

Deena K. Bernstein  
Sandra Levey  
Lehman College the City University of New York  
Pearl L. Seidenberg  
Long Island University

*Two hundred and seventy students in the 7<sup>th</sup>, 9<sup>th</sup>, and 11<sup>th</sup> grades participated in this study of skilled and less skilled readers' awareness and use of content and text structure knowledge to identify important text information. Their abilities were examined through presentation of an intact passage (allowing access to content and text structure schemata) and two modified passages (allowing access to only one schema). Developmental and passage differences were shown, with good readers demonstrating superiority over less skilled readers for all conditions. A strategy awareness task suggested that less skilled readers were not as sensitive to both schemata and were unable to coordinate the two types of knowledge when processing well-structured text.*

Research on text processing has increased our understanding of reading comprehension (Blank, 2004; Ehrlich, 1996; Feldt, Feldt, & Kilburg, 2002; Guthrie & Davis, 2003; Hammadou, 2000; Sanders & Noordman, 2000; Snyder & Downey, 1997). A body of evidence, based on schema theory (Anderson, 1994), reveals that prior knowledge about a topic, as well as awareness and knowledge about the characteristics of text, results in a reader's ability to identify important textual information and improves his or her comprehension of it (Duchan, 2004; Gersten, Fuchs, Williams, & Baker, 2001; Nelson, 1998; Richgels, McGee, Lomax, & Sheard, 1987; Roller, 1990; Slater & Graves, 1989; Westby, 1999). Researchers have also concluded that typically developing children (as well as mature readers) are more aware of the strategies they use to identify important textual information than their younger counterparts (Calisir & Gurel; 2003; Halldorson & Singer, 2002; Olhausen & Roller, 1988; Roller, 1990).

Another line of research has examined the processing strategies used by poor readers and readers with learning disabilities (Bashir & Scavuzzo, 1992; Kamhi & Catts, 2004; Lenski, 2002; Nation, Clarke, Marshall, & Durand, 2004; Nelson, 1998; Westby, 1999). These researchers concluded that poor readers and students with learning disabilities were constrained by their limited background knowledge and their decreased sensitivity to the organization of text (Cornoldi & Oakhill, 1996; Moss, 2004; Wallach, 1990; Westby, 1991, 1999).

The current study attempted to compare the text processing strategies of good readers (GR) and unskilled readers: poor readers (PR) and children with learning disabilities (LD). Our goal was to identify how content and text structure schemata influence their processing of expository text. Four general questions were addressed:

1. Do the three groups of children differ in their ability to identify important information under three different schema accessibility conditions?
2. Do the three groups differ in their responses when two schemata were accessible compared with when only one schema was accessible?
3. Do the three groups differ in the developmental patterns that emerge

for three different schemata accessibility conditions?

4. Do the three groups differ in the extent of the use of an accessible schema when the other schemata were also available?

## Method

### *Participants*

Two hundred and seventy students in the seventh, ninth, and 11th grades in the same middle class suburban school district in a northeastern state of the US participated, with 90 students in each grade level: 30 GRs, 30 PRs, and 30 classified as students with LDs. At each grade level, subjects in each of the three ability groups were randomly assigned to one of the passages representing the three schema accessibility conditions. Data from 8 LD students were not included in the final analysis; therefore, analysis was based on the scores of 262 subjects.

The 90 GRs were selected at random from a larger sample of students with scores above the 60th percentile on the reading comprehension section of the Metropolitan Reading Achievement Test. This standardized measure was used by the school district so that the classroom teachers could effectively group children for reading instruction. The GRs had an average reading comprehension grade level score of 11.8 (SD = 1.5; range = 4.9) and average readers had an average reading comprehension grade level score of 6.6 (SD = 1.9; range = 7.10). The PRs were selected at random from students with scores below the 40th percentile on the same test. The 90 students with LDs came from a pool consisting of most of the identified LD students in the three grade levels. All had received psychological and education evaluation and had been classified as LD by the school district. Objective criteria for placements in the LD program included a full scale IQ score around 80 and a discrepancy of at least 1.5 standard deviation between full scale IQ and a standardized achievement test. The school district special education program prohibited access to IQ scores. All of these LD students scored below the 40<sup>th</sup> percentile on the Metropolitan Achievement Test with an average reading comprehension grade level score of 6.3 (SD = 2.2; range = 9.1).

### *Materials*

Passages were about an obscure country named Melanesia, drawn from those used in a previous study (Olhausen & Roller, 1988). The intact version, the CS passage, allowed use of both content and structure schemata and consisted of nine paragraphs (see Appendix). The first, second, and sixth paragraphs provided structural information about the organization of the passage. Each of the remaining six paragraphs began with a sentence that signaled that the main idea, or target sentence, would appear next. The remaining sentences in each of the six paragraphs provided supporting details for the main idea. The second passage, the C passage, was designed to block the use of the text structure schema while permitting the reader to use a content ("nation") schema to perform the criterion test (e.g., information about geography, culture, history, etc.). Although similar to the CS version, signal sentences and structural paragraphs were deleted. Remaining sentences were scrambled and then arranged in six paragraphs. Each paragraph contained one target sentence, which appeared as either the second or the third sentence in the paragraph. The target sentence contained important information relevant to the "nation" schema such as location, landform, and other relevant data. The third passage, the S passage, blocked the use of a content schema (the "nation" schema) while allowing the use of a text structure schema. The structural features of the text remained

intact but vocabulary items that permitted access and use of the “nation” schema were replaced with nonsense words, forcing the reader to rely on only the structure schema to identify important information.

### *Procedures*

All students received a packet consisting of three pages. The first page provided specific instructions of completing the task. The second pages contained one of the three passages representing one of the three schema accessibility conditions described above, the: CS passage, C passage, and S passage. The third page directed the students to explain, in writing, how they decided which sentences were most important.

All students were given the same instructions. They were told that the purpose of the task was to help identify how readers figure out the most important information in their textbooks. Next, the test passage was read aloud by the examiner, with each student instructed to read along. The listening-while-reading activity was included to diminish the effects of any decoding or word recognition difficulty on text comprehension. After hearing the passage, they were instructed to reread the passage at least three times and then, when familiar with the passage, to underline the six most important sentences. The students who received the C and S passages were told that the passages might seem strange, but if they read them carefully, that they should be able to figure out what was most important. After completing the underlining task, subjects described how they decided which sentences were most important.

### *Scoring*

Responses to the underlining task were scored either right or wrong, depending on whether or not the underlined sentence was a target sentence. Scores ranged from 0-6 on the underlining task. In order to estimate the reliability (internal consistency) of the six target sentences (6 items) for subjects in this study, coefficient alpha (Cronbach, 1984) was computed. The reliability coefficients were as follows: CS passage = .809; C passage = .703; and S passage = .790. The strategic awareness scores could range from a low of 1 to a high of 4. The content scoring was based on the following criteria: (a) 1 = no strategy; (b) 2 = general mention of topics or details; (c) 3 = 1, 2, or 3 topics of the 6 “nation” schema topics (such as location, landforms, climate, history, culture, or economy) specifically mentioned; (d) 4 = 4 or more topics specifically mentioned: awareness of the 6 major categories and a search for the sentence that best covered each topic.

The text structure score was based on the following criteria: (a) 1 = no obvious strategy (e.g., “I guessed”); (b) 2 = indication of searching for general structural features (e.g., “I looked for the main idea”); (c) 3 = indication of a strategy related to the specific structure of the passages (e.g., mention of signaling or topical paragraphs); and (d) 4 = specific structure strategy related to the text materials (e.g., mention of signaling in the first sentence of the paragraph and then a search for the main idea sentence that followed).

Two raters independently scored the responses of two randomly selected subjects at each ability level and schema accessibility condition for each grade level (a total of 54 subjects and 21 percent of written reports). Agreement reached 92% for the GRs, 89% for the PRs, and 86% for students with LDs. Power analyses were conducted to estimate the sensitivity of the analyses. Power was computed for all interactions ( $p = .05$ ), using the Statistica (2002) software package.

## Results

### *Underlining Task*

A Multiple Analysis of Variance was conducted with scores on the sentence, underlining task as the dependent variables and grade level (GL), reading group (RG), and schema accessibility (SA) as the independent variables. Statistically significant main effects for all three independent variables, as well as a significant interaction between grade level and schema accessibility, were found (see Tables 1 and 2). Since RGs did not interact statistically significantly with either of the other independent variables, the means of the LD (1.66), PR (1.67), and GR (2.81) groups could be compared across GLs and SA conditions. Post hoc Tukey HSD tests indicated that the GR mean was significantly higher than both the LD and PR means, which were almost identical.

**Table 1**

*Means (and Standard Deviations) on Sentence Underlining Task by Grade Level (GL), Reading Group (RG), and Schema Accessibility (SA)*

	GRADE 7				GRADE 9				GRADE 11				Total 11 Totals
	LD	PR	GR	Total 7	LD	PR	GR	Total 9	LD	PR	GR	Total 11 Totals	
CS	1.70 (1.89) n=10	1.80 (1.99) n=10	3.10 (1.91) n=10	2.20 (1.97) n=30	2.50 (1.51) n=10	1.90 (1.37) n=10	3.40 (1.58) n=10	2.60 (1.57) n=30	2.50 (1.78) n=10	2.80 (1.48) n=10	4.00 (1.33) n=10	3.10 (1.63) n=30	2.63 (1.75) n=90
C	1.00 (1.05) n=10	1.10 (.74) n=10	1.20 (1.92) n=10	1.10 (.88) n=30	1.00 (.93) n=8	1.30 (.82) n=10	2.00 (1.05) n=10	1.46 (1.00) n=28	1.63 (.92) n=8	2.20 (1.48) n=10	3.20 (.79) n=10	2.39 (1.26) n=28	1.64 (1.18) n=86
S	.60 (.84) n=10	.50 (.71) n=10	2.40 (1.07) n=10	1.17 (1.23) n=30	2.40 (1.71) n=10	2.00 (1.63) n=10	3.10 (1.60) n=10	2.50 (1.66) n=30	1.33 (1.51) n=6	1.40 (1.07) n=10	2.90 (2.18) n=10	1.96 (1.78) n=26	1.87 (1.64) n=86

Note. Good reader (GR), poor reader (PR), and learning disabled (LD) students' performance was compared.



Table 2

*Analysis of Variance of Sentence Underlining Task by Schema Accessibility (SA), Reading Group (RG), and Grade (GR)*

Source of Variation	Sum of squares	DF square	Mean	F
Condition (SA)	48.057	2	24.028	12.291**
Reading Group (RG)	77.503	2	38.752	19.823**
Grade	44.411	2	22.205	11.359**
SA x grade	6.882	4	1.721	.880
SA x grade	19.254	4	4.721	2.462*
RG x grade	3.817	4	.954	.488
SA x RG x grade	6.847	8	.856	.438
Residual	459.408	235	1.955	
Total	668.141	261	2.560	

\*\*p < .001, \*p < .05.

Tests of simple main effects conducted for grade level within schema accessibility indicated that the grades differed statistically significantly within all three accessibility conditions. For the CS passage,  $F(2,235) = 3.12$ ,  $p < .05$ ; for the C passage,  $F(2,235) = 6.35$ ,  $p < .01$ ; and for the S passage,  $F(2,235) = 6.86$ ,  $p < .001$ . Post hoc Tukey HSD comparisons determined that 11th grade students had a statistically significant higher mean than seventh grade students in the CS condition and a statistically significant higher mean than both seventh and ninth grade students in the C condition. In the S condition, both ninth and 11th grade students scored statistically significantly higher than did seventh grade students.

Overall, the data for the underlining task indicated that the GR's ability to identify important information was superior to both the PRs and children with LDs, whose performance was almost identical. For the most part, the important information on the C and S passages was less accessible than that in the intact text (CS). In addition, subjects' performances for the three conditions increased as GL increased. The fact that there was no difference in performance on the C and S passages suggests that they were equally effective in relating important information. An analysis of the strategy awareness scores, which is discussed below, provides additional insight into how content and structure schemata interact and influence text processing for students at the three different ability levels.

*Strategic Awareness (SA)*

The goal was to examine a critical metacognitive skill: how aware readers of different abilities were in their use of the different schemata on the underlining task.

#### *Content Strategy Awareness*

A Multiple Analysis of Variance was conducted with scores on the content strategy awareness, task as the dependent variables, and GL, RG, and SA as the independent variables. All three main effects were significant, but because of the significant interaction between SA and RG, only the means for GL were compared across the other two independent variables (see Tables 3 and 4). Post hoc analyses indicated that these were all statistically significantly different from one another so that the content SA means of 1.23, 1.52, and 1.88 for Grade 7, Grade 9, and Grade 11 respectively, exhibited a clear developmental pattern of increasing content SA as GL increased.

Table 3

*Means (and Standard Deviations) on Strategy Awareness (Content) by Grade Level (GL), Reading Group (RG), and Schema Accessibility (SA)*

	GRADE 7				GRADE 9				GRADE 11				Total 11 Totals
	LD	PR	GR	Total 7	LD	PR	GR	Total 9	LD	PR	GR	Total 11	
CS	1.20 (.42) n=10	1.20 (.42) n=10	1.60 (1.07) n=10	1.33 (.71) n=30	1.30 (.67) n=10	1.90 (.74) n=10	1.90 (1.10) n=10	1.70 (.88) n=30	1.50 (.53) n=10	1.90 (.74) n=10	3.00 (1.05) n=10	2.13 (1.01) n=30	1.72 (.92) n=90
C	1.00 (.00) n=10	1.30 (.67) n=10	1.40 (.70) n=10	1.23 (.57) n=30	1.63 (.74) n=8	1.20 (.63) n=10	1.80 (.79) n=10	1.54 (.74) n=28	2.13 (.83) n=8	2.00 (.82) n=10	2.00 (.94) n=10	2.04 (.84) n=28	1.59 (.79) n=86
S	1.20 (.42) n=10	1.10 (.32) n=10	1.10 (.32) n=10	1.13 (.35) n=30	1.10 (.57) n=10	1.30 (.67) n=10	1.60 (.70) n=10	1.33 (.66) n=30	1.67 (1.21) n=6	1.30 (.48) n=10	1.40 (.52) n=10	1.42 (.70) n=26	1.29 (.59) n=8

Table 4

*Analysis of Variance of Strategy Awareness (Content) by Schema Accessibility (SA), Reading Group (RG), and Grade Level (GL)*

Source of Variation	Sum of squares	DF square	Mean	F
Condition (SA)	8.220	2	4.110	8.187***
Reading Group (RG)	6.447	2	3.223	6.420**
Grade	17.060	2	8.530	16.990***
SA x RG	4.989	4	1.247	2.484*
SA x grade	2.642	4	.661	1.316
RG x grade	.594	4	.149	.296
SA x RG x grade	7.944	8	.993	1.978
Residual	117.983	235	.502	
Total	167.118	261	.640	

\*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$ .

Tests of simple main effects of RG within accessibility condition yielded significant results only for the CS passage where the GR mean was found to be significantly higher than both the PR and LD means,  $F(2,235) = 10.51$ ,  $p < .001$ .

Although there was an increase in the reported use of a content strategy as GL increased, the RG by SA interaction reflected the fact that the mean content scores for the C passage were higher than the scores for the S passage only for the GR group. This finding suggests that, at each GL, only subjects in the GR group were able to report the use of the appropriate strategy for the appropriate passage. In addition, the GR subjects reported increased use of the appropriate strategies as they grew older. These findings show that, for the GR group, the "nation" schema is available from seventh grade on and appears to validate the content schema as a strategy that GRs use when identifying important information in text.

In contrast, the PR and LD groups did not report an increased use of the content strategy for the C passage compared with the S passage and, therefore, the PRs appear to be less sensitive to the "nation" schema's influence in the processing of text than the GRs. Also of interest is the finding that content scores for the CS passage were higher than those for the C passage for the GRs. It appears that the extra structural support in the form of textual cues in the CS passage consistently helped the GRs focus on the content of the passage and thus report more use of the content strategy.

*Structure Strategy Awareness*

A Multiple Analysis of Variance was conducted with scores on the structure strategy awareness task as the dependent variables and grade level (GL), reading group (RG), and schema accessibility (SA) as the independent variables. Analysis revealed statistically significant main effects for all three of the independent variables (see Tables 5 and Table 6). In addition, GL was found to interact significantly with both SA and RG. Tests of simple main effects of GL within SA indicated statistically significant differences among the grades for the CS passage,  $F(2,235) = 3.87, p < .05$ , and for the C passage,  $F(2,235) = 5.21, p < .001$ , but the GLs did not exhibit statistically significantly different structure SA scores in the case of the S passage. Post hoc analyses indicated that for both the CS and C passages, seventh graders had significantly higher structure SA scores than did ninth and 11th graders.

Table 5

Means (and Standard Deviations) on Strategy Awareness (Structure) by Grade Level (GL), Reading Group (RG), and Schema Accessibility (SA)

	GRADE 7				GRADE 9				GRADE 11				Total 11 Totals
	LD	PR	GR	Total 7	LD	PR	GR	Total 9	LD	PR	GR	Total 11	
CS	1.10 (.32) n=10	2.10 (.57) n=10	2.30 (.67) n=10	1.83 (.75) n=30	1.30 (.67) n=10	1.30 (.48) n=10	2.10 (.74) n=10	1.57 (.73) n=30	1.10 (.32) n=10	1.40 (.84) n=10	1.80 (.79) n=10	1.43 (.73) n=30	1.61 (.75) n=90
C	1.20 (.42) n=10	1.60 (.52) n=10	2.20 (.42) n=10	1.67 (.61) n=30	1.00 (.00) n=8	1.30 (.48) n=10	1.40 (.52) n=10	1.25 (.44) n=28	1.13 (.35) n=8	1.30 (.48) n=10	1.40 (.52) n=10	1.29 (.46) n=28	1.41 (.54) n=86
S	1.00 (.31) n=10	1.80 (.59) n=10	1.40 (.67) n=10	1.40 (.56) n=30	1.30 (.67) n=10	1.40 (.52) n=10	2.00 (.94) n=10	1.57 (.77) n=30	1.83 (.75) n=6	1.50 (.53) n=10	1.60 (.52) n=10	1.62 (.57) n=26	1.52 (.65) n=86

Table 6

*Analysis of Variance of Strategy Awareness (Structure) by Schema Accessibility (SA), Reading Group (RG), and Grade Level (GL)*

Source of Variation	Sum of squares	DF square	Mean	F
Condition (SA)	2.166	2	1.083	3.367*
RG	16.348	2	8.174	25.406***
GR	2.367	2	1.183	3.678*
SA x RG	2.558	4	.640	1.988
SA x GR	4.468	4	1.117	3.471**
RG x GR	4.200	8	.525	1.632**
Residual	75.608	235	.322	
Total	111.439	261	.427	

\*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$ .

The interaction between GL and RG was also explored. Tests for simple main effects indicated no statistically significant difference on structure SA scores among GLs within the LD reading group. However, the GL means did differ statistically significantly within the PR group,  $F(2,235) = 6.87$ ,  $p < .001$ , and the GR group,  $F(2,235) = 3.21$ ,  $p < .05$ . In the PR group, the seventh grade mean (1.83) was statistically significantly higher than both the ninth (1.33) and 11th grade means (1.40). In the GR group, the only statistically significant difference was between the seventh grade mean of 1.97 and the 11th grade mean of 1.60.

In general, the structure strategy awareness scores were not as straightforward as the content scores. While the reporting of the increased use of a content strategy was evident as GL increased, the findings for the structure strategy awareness scores exhibited a disruption of the developmental trend. For example, the mean structure strategy awareness scores for the S passage did not differ across grades, and seventh graders showed higher structure strategy awareness scores for both the CS passage and the C passage (where it was not possible to use it) than did the ninth and 11th graders.

The behavior of the seventh graders may represent an over-reliance on the use of a partially learned schema that is still vague, ill-defined, and not well understood (in this case, text structure organization), which is introduced to students in the seventh grade. Seventh graders even attempted to use the schema when it was not supported by the

content of the passage. It appears that, when faced with a difficult passage, both good and less able younger readers may attempt to activate a structure schema in order to process the text even when the strategy is not appropriate.

There was a clear pattern of increasing reported use of the content strategy as the grade increased, which was not the case for the structure strategy. One fact that may have made the structure schema less accessible than expected is the structure of the discourse type used in this study. According to Meyer and Freedle (1984), the description text structure type, of which our MelanAsia passage is an example, is the least organized and least well-defined of the various text structure types. Therefore, it provides fewer text signals or cues.

#### Discussion

The findings in the current study mirror those reported by Olhausen and Roller (1988) for the underlining task, which measured the ability to locate important information in both intact and manipulated text. While our subjects' performances improved with age, their performances were influenced by passage differences. GRs at each GL also demonstrated a clear superiority over less able readers (PR and LD) on the underlining task for the three passage conditions, with improved scores as GL increased. These findings were consistent with those of Olhausen and Roller (1988), which showed a correlation between age, reading ability, and the ability to identify important text information. The performances of the PR and LD groups on the underlining task were almost identical, suggesting that the overall text processing abilities of these two ability groups may be similar. GRs demonstrated a consistent superiority over less able readers in reported use of the content strategy for both the appropriate manipulated passage (C) and the intact passage (CS). Only GRs reported a higher level of use of the content strategy for the appropriate passage (C) than for the inappropriate passage (S), and they were more likely to use the appropriate strategy as their GL increased. At the same time, GRs' content strategy scores were consistently higher for the intact text (CS passage) than on the appropriate manipulated passage (C passage). It appears that, when it is possible also to use a structure schema (as in the CS passage), good readers were able to report increased use of their available content schemata.

In general, the findings for the content strategy awareness scores were more straightforward than those for the structure strategy awareness scores, suggesting that the structure schema was less well-defined and, therefore, less accessible than the content schema for all the subjects in the study. Even GRs were not able to identify their use of a structure strategy in the same way that they were able to identify their use of the content strategy. This finding is not inconsistent with results reported in prior studies for these two lines of research. Although a reader's background knowledge may play a role in their ability to construct meaning from text, the studies of text structure knowledge have produced mixed findings. The results of the underlining task, coupled with the data generated by the content and structure strategy awareness task, provided a consistent pattern, with GRs better than the less able readers to regulate or coordinate the use of both content and structure schemata, resulting in a superior ability to identify important information in a well-structured passage.

As reported earlier, GRs at all GLs reported a higher level of use of the content strategy for the intact text (CS passage) than for the appropriate manipulated passage (C passage), and they appeared able to make fuller use of their content schema when they were able to coordinate the content schema with the structure schema when presented with the intact text. It is also of interest to note that even the oldest readers (the 11th graders) in both the



PR and LD groups failed to report a high level of use for either strategy when processing the intact text. Both the PR and LD groups, even at Grade 11, still appear to have difficulty in coordinating both content and text structure schemata when processing the intact text. GRs appear to make fuller use of both content and structure schemata in processing intact text than less fluent readers, who appear to be less sensitive to both content as well as text structure cues. The difference between GRs and less able readers appears to be in their coordination of the two strategies available to them rather than their knowledge about these strategies. Even when the older PR and LD subjects reported the use of both strategies; it was not reflected in their ability to identify important information in well-structured text. For the PR and LD subjects, strategy awareness alone does not appear to lead to automatic coordination of the two schemata and improved text processing. One may speculate that the less able readers' difficulty coordinating their strategic knowledge may be a result of over reliance on only one schema (Holmes, 1987; Kletzien, 1991). The significance of the current findings on the comprehension of expository text is clear: an integration and coordination of two types of knowledge (i.e., content knowledge that characterizes domain-specific knowledge and text structure knowledge that characterizes text organization) is necessary for identifying important information in text for all readers of different abilities (Gallagher & Pearson, 1989; McCormick, 1992; McKeown, Beck, Sinatra, & Loxterman, 1992; Roller, 1990).

#### Implications and Future Research

Both PRs and students with LDs experience serious problems in text comprehension and need training in the use and integration of content and structure schemata. Although there were some qualitative differences between the PR and LD groups in strategy awareness and use, our findings suggest that students in both groups would benefit equally from more explicit systematic instruction. Specific intervention approaches for students with text processing deficits should prove beneficial for typical children as well as for students who are poor readers and students with learning disabilities (Scruggs, 1990). Systematic, explicit, and guided instruction should provide all children, including less able readers, with differentiated awareness of the content and text structure knowledge and their interdependence. This might include: (a) summarizing, outlining, or concept mapping; (b) inducing text structure by differentiating among superordinate and/or subordinate ideas; (c) providing visual guides that highlight conceptual relationships (Duchan, 2004; Gallagher & Pearson, 1989; Richgels, McGee, Lomax, & Sheard, 1987). Future research should examine the processing strategies used by children when they encounter other than descriptive text structure and identify the most effective instructional approaches that will enable them to comprehend a greater variety of texts.

## References

- Anderson, R. (1994). Role of the reader's schemata in comprehension, learning and memory. In R. B. Ruddell, M. Rapp, & H. Singer (Eds.), *Theoretical models and processes of reading* (4<sup>th</sup> ed., pp. 469-482). Newark, DE: International Reading Association.
- Bashir, A., & Scavuzzo, A. (1992). Children with language disorders: Natural history and academic success. *Journal of Learning Disabilities, 21*, 53-65.
- Blank, M. (2004). Classroom discourse: A key to literacy. In K. G. Butler & E. R. Silliman (Eds.), *Speaking, reading, and writing in children with language learning disabilities: New paradigms in research and practice* (pp. 151-173). Mahwah, NJ: Erlbaum.
- Calisir, F., & Gurel, Z. (2003). The influence of text structure and prior knowledge of the learner on reading comprehension, browsing, and perceived control. *Computers in Human Behavior, 19*, 135-146.
- Cornoldi, C., & Oakhill, J. (Eds.). (1996). *Reading comprehension difficulties: Processes and intervention*. Mahwah, NJ: Erlbaum.
- Cronbach, L. J. (1984). *Essentials of psychological testing*. New York: Harper & Harper.
- Duchan, J. F. (2004). *Framework in language and literacy: How theory informs practice: How theory informs practice*. New York: The Guilford Press.
- Ehrlich, M. F. (1996). Metacognitive monitoring in the processing of anaphoric devices in skilled less skilled comprehenders. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 221-249). Mahwah, NJ: Erlbaum.
- Feldt, R. C., Feldt, R. A., & Kilburg, K. (2002). Acquisition, maintenance, and transfer of a questioning strategy in second- and third-grade students to learn from science textbooks. *Reading Psychology, 23*, 181-198.
- Gallagher, M., & Pearson, P. D. (1989). *Discussion, comprehension and knowledge acquisition in content area classrooms* (Technical Report 480). Urbana: University of Illinois, Center for the Study of Reading.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review in Educational Research, 71*, 279-320.
- Guthrie, J. T., & Davis, M. H. (2003). Motivating struggling readers in middle school through an engagement model of classroom practice. *Reading and Writing Quarterly, 19*, 59-85.
- Halldorson, M., & Singer, M. (2002). Inference processes: Integrating relevant knowledge and text information. *Discourse Processes, 34*, 145-161.
- Hammadou, J. (2000). The impact of analogy and content knowledge on reading comprehension: What helps, what hurts. *The Modern Language Journal, 84*, 38-50.
- Holmes, B. C. (1987). Children's inferences with print and pictures. *Journal of Educational Psychology, 79*, 14-18.
- Kamhi, A. G., & Catts, H. W. (2004). The language basis of reading: Implications for classification and treatment of children with reading disabilities. In K.G. Butler & E. R. Silliman (Eds.), *Speaking, reading, and writing in children with language learning disabilities: New paradigms in research and practice* (pp. 45-72). Mahwah, NJ: Erlbaum.
- Kletzien, S. B. (1991). Strategy use by good and poor comprehenders reading expository text of differing levels. *Reading Research Quarterly, 26*, 67-86.

- Lenski, S. D. (2002). Strategy instruction for a sociocognitive perspective. *Reading Psychology, 23*, 27-143.
- McCormick, S. (1992). Disabled readers' erroneous responses to inferential comprehension questions: Description and analysis. *Reading Research Quarterly, 27*, 55-77.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly, 27*, 79-93.
- Meyer, B. F. J., & Freedle, R. O. (1984). Effects of discourse type on recall. *American Education Research Journal, 21*, 121-143.
- Moss, B. (2004). Teaching expository text structures through information trade book retellings. *The Reading Teacher, 57*, 710-718.
- Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment. *Journal of Speech, Language, and Hearing Research, 47*, 199-211.
- Nelson, N. (1998). *Childhood language disorders in context*. Boston: Allyn & Bacon.
- Olhausen, M. M., & Roller, C. M. (1988). The operation of text structure and content schemata in isolation and interaction. *Reading Research Quarterly, 23*, 70-88.
- Richgels, D. J., McGee, L. M., Lomax, R. G., & Sheard, C. (1987). Awareness of four text structures: Effects on recall of expository text. *Reading Research Quarterly, 22*, 177-196.
- Roller, C. M. (1990). Commentary: The interaction of knowledge and structure variables in the processing of expository prose. *Reading Research Quarterly, 25*, 79-89.
- Sanders, T. J. M., & Noordman, L. G. M. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes, 29*, 37-60.
- Scruggs, T. E. (1990). Commentary: Foundations of intervention research. In T. E. Scruggs & B. Y. L. Wong (Eds.), *Intervention research in learning disabilities* (pp. 66-77). New York: Spring-Verlag.
- Slater, W., & Graves, M. F. (1989). Research on expository text: Implications for teachers. In K. D. Muth (Ed.), *Children's comprehension of text* (pp. 140-166). Newark, DE: International Reading Associates.
- Snyder, L., & Downey, D. (1997). Developmental differences in the relationship between oral language deficits and reading. In K. Butler (Series Ed.) & P. Menyuk & M. Chesnick (Vol. Eds.), *Topics in language disorders: Perspectives on relations between oral knowledge and reading* (3<sup>rd</sup> ed., pp. 27-40). Rockville, MD: Aspen.
- Statistica* [Computer software]. (2002). Tulsa, OK: StatSoft, Inc.
- Wallach, G. P. (1990). Magic buries Celts: Looking for broader interpretations of language learning and literacy. *Topics in Language Disorders, 10*, 63-80.
- Westby, C. E. (1991, October). *Understanding classroom texts* [Audio conference]. Steps to Developing and Achieving Language Based Curriculum in the Classroom. Parkville, MD: American Speech-Language Hearing Association.
- Westby, C. E. (1999). Assessing and facilitating text comprehension problems. In H. W. Catts & A. G. Kamhi (Eds.), *Language and reading disabilities* (pp. 154-223). Boston: Allyn & Bacon.

## Appendix

Melanesia is a relatively unknown country. In order to learn more about Melanesia, we will describe both the physical geography and the culture of the country. We will first describe the physical geography of Melanesia. Specifically, we will focus on its location, then the landforms, and finally the climate. First, we look at location as an aspect of physical geography. Melanesia is an island nation located in the Pacific Ocean, northeast of Australia. This curving chain of islands stretches from New Guinea to the Fiji Islands. A second aspect of physical geography is landforms. Some of the islands in Melanesia are large, single islands, and some are made up of island groupings. The large islands of Melanesia are called continental islands, which means they are the tops of the underwater mountains that have risen from the ocean depth. There are an abundance of mineral resources and dark, rich soil in many of the islands. The final topic included in the physical description of the physical geography is climate. The climate of Melanesia is uncomfortably hot and humid, and Melanesians most often face severe tropical storms called typhoons. The climate is also partly responsible for the persistence of tropical diseases such as malaria. An average of 300 typhoons a year occurs in the Pacific island region, and often the winds reach 150 miles per hour, destroying wildlife and damaging houses and vegetation. We will next describe the culture of Melanesia. More specifically, we will examine the history, society, and economy of the country. First, we will examine the historical aspects of the culture. The history of Melanesia shows an ancient people primarily engaged in agriculture. Historians think that people came to the islands of Melanesia perhaps as long as 25,000 years ago. Those historians believe that people from Southeast Asia immigrated to the Pacific islands in groups, or waves. Each group settled in a different area of the region and developed its own unique society. The early Melanesians were gardening people. They cleared garden plots with stone and shell tools and by burning jungle growth. Each tribe exchanged goods such as food, animals, clay and wooden bowls, woven mats, weapons, and even canoes with other tribes. Religious ceremonies were often used to protect crops from harm or to increase their yields. Next, we will focus on the society of the country. Today, most people in Melanesia live in small villages ruled by a leader called a headman who was chosen by the others in the village. To qualify to be chosen as a headman, a villager has to be wealthy. Each person in a village is related by birth or marriage to everyone in the village. A final important topic is economy. The island economy of Melanesia consists of mining of many valuable minerals and plantation farming. Important minerals include coal, petroleum, gold, nickel, chrome, and aluminum ore. Villagers depend on a few root crops for food; taro, a starchy root, is the most important. Melanesians produce sweet potatoes and yams, breadfruits, bananas, and coconuts. They also rely on hunting and fishing. Today a growing population is creating economic difficulties for some of these islands.

*Deena K. Bernstein is Professor and Chair and Sandra Levey is an Assistant Professor in the Department of Speech-Language-Hearing Sciences at Lehman College of the City of New York. Pearl L. Seidenberg is Professor Emeritus at the CW Post Center of Long Island University. Correspondence concerning this article should be addressed to Dr. Sandra Levey, Department of Speech-Language-Hearing Sciences, Lehman College The City University of New York, 250 Bedford Boulevard, W., Bronx, NY 10468-1589.*

# THE IMPACT OF ONLINE COMMUNICATION DESIGN ON STUDENT MOTIVATION AND PERFORMANCE

Leping Liu  
University of Nevada, Reno  
Paul Jones  
Towson University

*In this study, we proposed two types of design, Type I (static) design and Type II (dynamic) design. We examined (a) the impact of two approaches of online communication design (single-layer, and multi-layer design) on student motivation toward online learning, and on student learning in terms of the completion of communication tasks, (b) the relationship between student motivation and their online communication performance, and (c) student satisfaction level to the use of each communication format. The participants of this study were 224 university students who enrolled in online courses. Findings suggest that (a) motivation scores, as well as online communication performance scores, were different between students who performed single-layer communication activities and those who performed multi-layer activities; and (b) a linear relationship existed between student motivation and online communication performances.*

Over the past decade, online education has been promoted from the stage of using Web-based resources to the stage of designing Web-based instruction (Fishman, 1997; Riel, 1992; Trentin, 2001). While the emphasis of online education is switching from “what to use” to “how to use,” online course design becomes one of the most discussed topics in the literature (Boer & Collis, 2001; Liu, 2003; Liu & Maddux, 2003a; Schweizer, Whipp, & Hayslett, 2002). Among the major components of online course design, communication design is a critical issue that ensures the success of an online course (Aviv & Golan, 1998; Berge, 1997).

Based on a review of motivation theories, and our experiences with communication design in Web-based courses, this study examined the impact of online-communication design—in terms of its format, frequency, and range—on student motivation toward online learning and on student communication performances. The purpose of this study was to develop a sub model of online communication design by examining the impact of different approaches of online communication design on student motivation toward online learning, in terms of the effectiveness of different approaches to online communication design on student learning in terms of the completion of communication tasks, the relationship between student motivation and online communication performance, and student satisfaction level to the use of each communication format.

## Motivation and Learning Achievement

Generally, the term *motivation* can be defined as a willingness to try or attempt to perform an act (Kellenberger, 1996; Kelsey & D'souza, 2004). Existing theories and models of motivation are attempts to explain future behavior (Harlen & Crick, 2003; Keller, 1993; Wishart, 2001). Achievement motivation and goal setting theories are the most widely discussed and applied theories in this field.

### *Achievement Motivation Theory*

*Achievement motivation theory* (Atkinson & Feather, 1966; Kelsey & D'souza, 2004; Lewin, Dembo, Festinger & Sears, 1944; McClelland, 1951; Raynor, 1969; Wishart, 2001) is based primarily on the findings of

studies in which individuals are classified as high or low in their need for achievement. These different levels of need for achievement and different levels of aspiration toward achievement are related to performance in one of two ways: *the tendency to achieve success* and/or *the tendency to avoid failure* (Atkinson, 1974). The *tendency to achieve success* is a multiplicative function of three components: (a) motive to achieve success, a personality component representing an individual's willingness for success; (b) expectancy to achieve success, an environmental component referring to an individual's expectations of the probability of success; and (c) the incentive value of success, another environmental component concerning the attractiveness of success. The general implication of this function is that individuals will perform better on a task when (a) they are highly motivated, or are willing to achieve success; (b) they believe their performance will be most probably followed by success; and (c) they recognize the value of the consequences of their performance. Typically, "the differences in behavior are due to differences in strength of tendency to achieve in individuals who differ in strength of achievement motive" (Atkinson, 1974, p. 15). Many motivational models developed in education have been based on this "motive—expectancy—value theory" (Crookes & Schmidt, 1991; Harlen & Crick, 2003; Keller, 1993; Pintrich, 1990).

The *tendency to avoid failure* (Ames & Ames, 1989; Atkinson, 1974; Crookes & Schmidt, 1991; Kelsey & D'souza, 2004) can be considered a function separate from the tendency to achieve success. Achievement motivation theories assume that the tendency to avoid failure is also a multiplicative function of expectancy to achieve success and the incentive value of success. In contrast to the tendency to achieve success, the motive to avoid failure can be thought of as a capacity for reacting with humiliation and shame when one fails. The expectancy to avoid failure refers to one's expectations on the probability of failure, and the incentive value of failure indicates a negative tendency such as the repulsiveness of failure. The tendency to avoid failure is aroused and expressed when one has the fear of failure, or realizes that some act may lead to failure, or is aware of the negative value of the consequence of failure. Obviously, the three components in the function of the tendency to avoid failure all appear to be negative. Therefore, their combined motivational effect might be thought of as, if not a negative effect, a low-level motivational effect toward achievement.

### *Goal Setting Theory*

Another motivation theory that addresses the relationship between motivation and performance is *goal setting* theory (Locke & Latham, 1990). Goal setting theory is based on the premise that much human action is purposeful (Binswanger, 1991; Dembo, 2000; Kelsey & D'souza, 2004; Shih & Gamon, 2001). In other words, why do some people perform better on work tasks than others? When ability and knowledge are controlled, the question must be answered within the realm of motivation. According to goal setting theory, the reason some people perform better on work tasks than others is because they have different performance goals.

According to Locke and Latham (1990), goal setting consists of two attributes: *content* and *intensity* (Kelsey & D'souza, 2004; Locke & Latham, 1990). The *content* attribute includes two characteristics: specificity and difficulty. At different levels of specificity, the goal content can be vague ("do it") or specific ("solve 50 math problems in Chapter Two correctly within the next two hours"). At different levels of difficulty, the goal content can be easy ("solve two problems in two hours") or difficult ("solve 100 problems in two hours"). Greater specificity and difficulty result in relatively higher goal levels. Consistent findings from over 400 studies (Locke & Latham, 1994) suggest that there is a linear relationship between performance and goal level, and higher level goals (specific

and difficult) lead to higher performance than lower level goals (such as vague but challenging, vague but unchallenging, or no goals). The harder and the more specific the task, the better the performance.

The other attribute of this goal setting model is *goal intensity*. Goal intensity refers to the degree of the individual's intensity in thinking about the goal. Gollwitzer, Heckhausen, and Ratajczak (1990) found, for example, that subjects who were most likely to become committed to and subsequently solve a problem were those who thought most intensely about how to solve it. Also, many goal-intensity studies have examined the degree of commitment of an individual to the goal. In goal-setting theory, the degree of commitment to a goal is similar to the incentive value of success in the theory of achievement motivation. It was found (Locke & Latham, 1994) that high commitment is more likely to lead to goal attainment than low commitment.

#### *Motivation and Learning Achievement*

The motivation theories are based on the relationship between motivation and performance. A "good performance" usually refers to the relatively complete process or quality of performance. In education, the measurement of student performance in their learning is their learning achievement. Theoretically, motivation influences performance and, hence, influences achievement. Many studies have investigated the relationship between motivation and achievement. Uguroglu and Walberg (1979), for example, examined the results of several studies of motivational factors on achievement. Using linear multiple regression models, they found that motivation contributed significantly to variance in achievement. According to Parkerson, Lomax, Schiller, and Walberg (1984), in a meta-analysis over 40 studies, motivation was found to be one of the consistent causes or correlates of learning achievement. Means, Jonassen, and Dwyer (1997) reviewed other investigators' findings that motivation accounted from 16% to 38% of the variance in student achievement. In their own study, they found 42.9% of the variance in learning achievement could be accounted for by motivation.

In summary we found that, (a) the tendency (motive, expectancy, and incentive value) to achieve success produces better performance; (b) the tendency to avoid failure produces decrements in performance; (c) high level goals (harder and more specific tasks) are more likely to result in better performance; (d) high commitment is more likely to lead to goal attainment than low commitment; and (e) motivation influences achievement and is a function of achievement. Correspondingly, it is reasonable to assume that these motivational factors would influence students' performances in an online learning environment. Therefore, we want to consider these attributes of motivation into the design of online communication, from overall strategies to particular activities.

#### Online Communication Design and Motivation

##### *An Online Communication Design Model -- ADIE*

An online course can be viewed as an interactive online information system. From a designer's view, its development should follow the major phases of information systems development: analysis (A), design (D), implementation (I) and evaluation (E) (Burch, 1992). As an important portion of an online course, online communication design includes the same ADIE phases: (a) *analysis of online communication* – to determine the purposes and needs of communication, analyze the characteristics of students involved in the communication, and locate available communication tools; (b) *design of online communication* – to determine the format, schedule, and range of communication, and decide the means or methods to assess the effectiveness of the communication; (c) *implementation of the communication design* – to set up the communication templates in different formats and at

different levels, create training materials for using the communication tools for students, and develop alternative tools as backups for each format; and (d) *evaluation of the communication design* – by examining the completion of tasks assigned to each communication format, the smoothness of communication tools, and student satisfaction level to the use of each communication format. The four phases formulate an ADIE online communication design model. This study will focus on the second phase – *design of online communication*.

### *Online Communication Design and Motivation*

In the second phase, *design of online communication*, we determine the format, frequency, and range of communication. First, online communication provides both asynchronous and synchronous environments for collaboration (Harlen & Crick, 2003; Hewitt, 2003; Ohlend, Yu, Jannasch-Pennell, & Digangi, 2000; Swan, Shea, Fredericksen, Pickett, Pelz & Maher, 2000; Vonderwell, 2003). In WebCT [the online course management system used in the current study], the format of asynchronous communication can be the use of a built-in class mailing list or the bulletin board; and the format of synchronous communication can be the use of class chat rooms. In previous studies, students had higher commitment and displayed more of the achievement-success motive in their communication tasks when using the class mailing list and bulletin board, but showed lower commitment and more tendency to avoid failure when they used class chat rooms (Liu, 2003; Liu & Maddux, 2003a).

Second, frequency of online communications can be daily-based, weekly-based, and need-based. According to our observations and communications with students, we found that if daily communication activities such as daily discussion and responses are required, students felt more pressure, less enjoyment, and had low commitment to the tasks; when the communication activities were weekly-based, students tended to set higher level of goals for their performances, and hence had higher commitment to their work; and when communication opportunities were provided by the instructor whenever students needed help, students showed higher motive and commitment to their work (Liu, 2003; Liu & Maddux, 2003a).

Third, the range of communication could be an important factor that influences student motivation. Communication activities can be distributed class-wide, group-wide, or privately. We also found that in class-wide communication activities, some students showed greater tendency to avoid failure, and some showed more motivation to achieve success. Furthermore, all students set a higher level of goals and had high commitment toward their work in group-wide activities and in private one-to-one communications between the instructor and the student.

Major online communication activities in our classes were (a) online discussions that included reading chapters, posting reading journals, raising discussion questions, and responding to other journals and questions; (b) online team building that included getting to know team members, choosing roles, electing team leader, formulating team work plan, and completing team project; and (c) exchange of information that included obtaining instructions, submitting course work, getting help from the instructor or among students, and reporting progresses to the instructor or within group.

Obviously, certain format, frequency, and range of communication are appropriate for certain online communication activities or tasks. But, what interests us more is whether there is a pattern of using proper format, frequency, and range that would motivate students in participating in online communication activities and better their performances. Exploration of the “pattern” will be based on two types of design.



### *Type I and Type II Design in Online Communications*

Type I and Type II design have been proposed to strategically determine the rules, principles, structures, and macro processes of educational designs (Liu & Johnson, 2004; Liu & Maddux, in press). The Type I design, *Static Design*, characterizes design principles and strategies in four ways. First, it is a linear design; all design units proceed one after another (Specific units and their micro processes can vary when dealing with different applications). Second, it is a single dimensional design, emphasizing certain major perspective of the application. Third, it focuses on the state of each individual unit of the design, interested in the weight of each unit. Fourth, it is a close-ended design without any room for further change.

The Type II design, *Dynamic Design*, distinguishes design principles and strategies in four other ways. First, it is a nonlinear design; all design units can be in a “tree” or “net” structure and do not have to be performed one after another. Second, it is a multiple dimensional design, emphasizing interactions among dimensions. Third, it focuses on the continuous process of all units of the design, especially the connections among them. Fourth, it is an open-ended design with developmental potentials. A design can be considered Type I or Type II if it meets any one feature or a combination of certain features under that type of design (see Table 1).

*Table 1*

### *Type I and Type II Design in Online Communication*

Features	Types of Design	
	Type I: Static Design	Type II: Dynamic Design
1	Linear	Non-linear
2	Single-dimensional	Multiple-dimensional
3	State-focused	Process-focused
4	Close-ended	Open-ended
<b>Online Communication Design</b>	Single-layer	Multiple-layer

In our online teaching experiences, two types of communication design were applied: single-layer, and multi-layer design. In a single-layer/single dimensional design, communication activities conducted with one format, at one frequency, and in one range, can be considered a Type I design. In a multi-layer/multiple dimensional design, communication activities performed with more than one format, at several frequencies, and in various ranges, can be viewed as Type II design (see Table 1).

The above review reveals some considerable components in online communication design, and offers a theoretical framework that student motivation may influence their online communication performances, and the use of either Type I or Type II methods to design the format, frequency, and range of communication in an online course could influence student motivation, and hence their performances. To carefully address the importance of the current study, we want to emphasize its contribution to the literature from three perspectives: (a) It explores the practical value of this theoretical framework, (b) it examines the effectiveness of the new proposed Type I and Type II design methods in the field of online learning, and (c) it presents some useful strategies and methods for online

communication design, which together will add a sub model into the *design* phase of the ADIE online communication design model.

Based on the purposes of this study, the following research questions were examined: (a) Is there any difference in the scores of motivation toward online learning between students who complete online activities via a single-layer communication design and those via a multi-layer communication design? (b) Is there any difference in the communication task scores between students who complete online activities via a single-layer communication design and those via a multi-layer communication design? (c) Can students' online communication performance be predicted by their motivation toward online learning? and (d) How are students satisfied with different formats, frequency, and range of online communication?

## Methodology

### *Participants*

The participants of the study were 224 education majors enrolled in two online education courses over three semesters at a Western state university from 2002 to 2003. One online course was an introductory technology course required for all the graduate students in the college; and the other one is a design course that was offered as a dual-level course to both graduates and undergraduates. The 224 participants consisted of 192 graduates and 32 undergraduates, 168 females and 56 males ranging in age from 19 to 42 with an average age of 26.6 (SD = 3.12). Approximately 84% of the participants had not had online learning experiences, 9% had taken one online course, and 7% had taken two or more online courses. They all had convenient Internet access at home or work, and had experiences on Web browsing, but only 22% of them rated themselves as "comfortable" with using technologies.

One author of this article was the instructor of both online courses. Sampling the participants from one instructor, we believed, would strengthen the study by controlling possible extraneous variables that might influence student motivation and performance. Such variables could be the instructor's philosophy of teaching, teaching style, personality, organization or communication skills, and attitudes toward students. "Types of design principles" is controlled as well. However, it might also be a limitation for not considering different impacts from the instructor on student learning which addresses another scope of the research.

### *The Online Courses*

The graduate required course, "Information Technology in Education," posed two main learning objectives such that students should be able to demonstrate an understanding of the background knowledge of technology integration in the field of education; and use and integrate current technology tools into teaching and learning.

According to the objectives, the course work consisted of non-technology work where students were expected to read assigned materials, search relevant information, write position journals, and discuss with the class items such as history, issues and trends, and theoretical background of technology integration. Course work in technology included learning basic technology skills where students were expected to complete 10 labs (such as advanced functions of Word, desktop publishing, digital photo and graphing, presentation software, spreadsheets, databases, web page design and publishing), and create an electronic portfolio. To learn the design and implementation of technology integrating, students were expected to develop a group project of technology-based teaching; and for their non-technology work were assigned online communication tasks, and technology work as individual or group offline lab work.

The design course, "Methods of Using Information Technology in Teaching and Learning," aimed at two main learning objectives such that students should be able to understand the fundamental theories of computer-based-instruction (CBI) design, and apply them into the development of CBI; and also use an authoring tool to create multimedia CBI lesson segments.

According to the objectives, online non-technology course work included completing assigned reading, writing position journals, planning and analyzing the CBI design, and discussing with the class ways to learn the related design theories. Offline technology course work allowed students to create a 5-minute run time multimedia program on a selected teaching/learning topic using an authoring tool such as *ToolBook*, *Director*, or *HyperStudio*.

Both courses were offered as hybrid online course, and all instructions and materials were delivered through the WebCT system. The course activities and communications were completed online. As a hybrid Web course, there were two required face-to-face class meetings on the first day and last day of the class and two optional meetings during the semester. The instructor was also available when students needed special help.

#### *Instruments and Measurements*

Three measurements were used for the classes. First, an attitude survey derived from Liu and Johnson's (1998) instrument was used. In the survey, six statements are categorized to measure student motivation (See Appendix, Part I). Among the six statements, items 1, 3 and 5 are positive statements, scoring from 1 (*strongly disagree*) to 5 (*strongly agree*); and items 2, 4 and 6 are negative statements scoring from 1 (*strongly agree*) to 5 (*strongly disagree*). The sum score of the six statements are the motivation score with a maximum of 30, where a higher score indicates more motivated attitudes. The reliability coefficients alpha for this instrument was .835 from the previous study (Liu & Johnson, 1998), and .822 for the current study.

Second, the *Survey of Student Preference* (See Appendix, Part II) comprised of 18 statements was adopted to obtain student preference to the communication formats (Webmail, discussion board, and chat room), communication frequency levels (daily-based, weekly-based, and need-based), and communication range (class-wide, within-group, and "private"). In this survey, for each format, frequency level, and range, a positive statement and a negative statement were provided. The positive statements, items 1 to 9, are scored from 1 (*strongly disagree*) to 5 (*strongly agree*); and the negative statements, items 10 to 18, are scored from 1 (*strongly agree*) to 5 (*strongly disagree*). The sum score of the two statements are the preference scores for each format, frequency and range of online communication with a maximum score of 10, where a higher score indicates a higher preference. Based on the data from the current study, the alpha reliability coefficient for this instrument was .819, suggesting that the scores were reasonably reliable for participants like those in the study.

Third, students' online communication performance grade was a sum of scores of their online theme discussions and online team-building projects. By the time the data were collected, students had completed communication tasks of a total point of 100: 30 points for completing the reading materials and posting five journals, 30 points for responding to 10 other journals, 20 points for team building (contacting their team members, introducing themselves to the team, determining their role in the team) using WebCT mail and chat room, and 20 points for completing the plan for team project. These tasks required a large amount of online activities.

The attitude survey and online course survey were regular anonymous measurements upon which the instructor could adjust the course pace and materials. For the purpose of the current study, only those volunteers put their names on the survey so that the instructor could match their performance scores.

*Research Design and Procedures*

This study applied an experimental design to explore differences and relationships. Data were collected during three semesters from six online classes. In each class, students were randomly assigned to two groups. One group performed online communication activities with the *single-layer* communication design, that is, they performed with one format (discussion board), at one frequency level (weekly-based), and in one range (class-wide only). The other group performed with the *multiple-layer* design, that is, they completed all the communication activities at all frequency levels (daily-based, weekly-based, and need-based), in all ranges (class-wide, within-group, and “private”), and used different communication tools (Webmail, personal email, discussion board, and chat room).

At the mid-term of the semester, the communication performance scores were collected. By then, students had completed online tasks of 100 points as described previously. Motivation scores were collected at the same time, which were expected to relate to their online communication experiences.

After the mid-term, the multiple-layer group stayed the same, but the single-layer group switched to perform under the multiple-layer communication environment so that they could experience all tools and means of online communication. Scores of student preference to each communication format, frequency, and range were collected at the end of each semester from both groups (see Table 2). The data collected at the mid-term were used in the current study to develop a sub model of the ADIE online communication design model.

*Table 2*

*Research Design*

	Communication Design Method			
	1 <sup>st</sup> half of the Semester		2 <sup>nd</sup> half of the Semester	
Group A	Single Layer	Collect Motivation & Performance data	Multiple Layer	Collect Preference data
Group B	Multiple Layer		Multiple Layer	

**Data Analysis and Results**

To explore different effects of the two communication design methods, the dependent variables were *motivation* and *performance*, and the independent variable was *design-method* at two levels – single-layer and multiple-layer. To explore the predicting relationship between student motivation and online communication performance, *motivation* was used as predictor variable, and *performance* as response variable. Data exploration showed that there were no significant outliers, and the equal variance assumption and the normality assumption were not violated. To answer the four research questions, four procedures of data analysis were completed.

### Data Analysis One

First, an independent sample *t* test was conducted to compare the means of motivation scores of students from the two groups (see Table 3). Significant difference was found ( $t_{(222)} = 31.601, p < 0.0001$ ) that students in the multiple-layer group (mean = 27.33,  $SD = 1.783$ ) had higher motivation scores than those in the single-layer group (mean = 19.73,  $SD = 1.816$ ). The measure of nonoverlap *U* (Cohen, 1988, p. 19) associated with effect size (Hinkle, Wiersma, & Jurs, 2003, p. 297; Sprinthall, 2000, p. 242) of this test was 97%, indicating that 97% of the scores in the two groups were not overlapped, thus, were different.

Table 3

### Mean Scores of the Groups

Variables	Group	Mean N=112	SD
Motivation	Single Layer	19.73	1.816
	Multiple Layer	27.33	1.786
Performance	Single Layer	79.53	3.351
	Multiple Layer	95.56	3.218

### Data Analysis Two

Second, another independent sample *t* test was conducted to compare the means of online-communication performance scores of students from the two groups (see Table 3). Significant difference was found ( $t_{(222)} = 35.552, p < 0.0001$ ) that students in the multiple-layer group (mean = 95.56,  $SD = 3.218$ ) had higher performance scores than those in the single-layer group (mean = 79.53,  $SD = 3.531$ ). The measure of nonoverlap *U* (Cohen, 1988, p. 19) associated with effect size (Hinkle, Wiersma, & Jurs, 2003, p. 297; Sprinthall, 2000, p. 242) of this test was larger than 97.7%, indicating that more than 97.7% of the scores in the two groups were not overlapped, thus were different.

### Data Analysis Three—A

Third, a simple regression analysis was performed, in which *motivation* was the predictor variable, and *performance* was the response variable regressed by *motivation*. To examine the predicting relationship in general, we used the combined data from both groups in regression analysis. In the results, the *F* ratio for the linear model was significant ( $F_{(1, 223)} = 5612.558, p < 0.0001$ ), indicating that a linear relationship existed between the response variable *performance* and the predictor variable *motivation*. Since this was a simple regression including only one predictor variable, the *F* value is equal to the square of the *t* statistics; the *t* statistics showed the same significant predictive relationship ( $t = 74.917, p < 0.0001$ ). The regression analysis generated a set of coefficients to formulate the *Regression Equation 1*:

$$(\text{PERFORMANCE}) = 39.77 + 2.03 (\text{MOTIVATION})$$

According to this equation, a one-unit increase in one's score of *motivation* will increase 2.03 units on his/her *performance* score.

R-Square of the model ( $R^2 = 0.862$ ) shows the effect size of the test that a major portion (approximately 86%) of the variation in the response variable *performance* was explained by this model, or by the variation in the

predictor variable *motivation*. Since a significant  $F$  value determines a significant linear model, according to Cohen (1988, p. 414), the power analysis for a regression analysis is usually performed on the  $F$  test. The linear interpolation finds the power of the  $F_{(1, 223)}$  test to be 0.99, indicating a very high probability of finding a significant linear relationship in this set of data.

#### *Data Analysis Three—B*

After the general analysis on the combined data of the two groups, we wanted to see whether the same relationship existed in each group, and whether the portion of variation in *performance* predicted by *motivation* was different by group. Therefore, we conducted two separate regression analyses using the data from each group. In each analysis, *motivation* was the predictor variable, and *performance* was the response variable. The results from the regression analysis on the data from the multiple-layer group indicated that a linear relationship existed between the response variable *performance* and the predictor variable *motivation* ( $F_{(1, 111)} = 236.707, p < 0.0001$ ). The  $t$  statistics showed the same significant predictive relationship ( $t = 15.358, p < 0.0001$ ). The regression analysis generated a set of coefficients to formulate the *Regression Equation 2*:

$$(\text{PERFORMANCE}) = 54.798 + 1.492 (\text{MOTIVATION})$$

According to the effect size of this model ( $R^2 = 0.683$ ), *motivation* predicted 68% of the variation of the response variable *performance*. The linear interpolation found the power of the  $F_{(1, 111)}$  test to be 0.97, indicating a relatively high probability of finding a significant linear relationship in the data from the multiple-layer group.

Next, the results from the regression analysis on the data from the single-layer group indicated that a linear relationship existed between the response variable *performance* and the predictor variable *motivation* ( $F_{(1, 111)} = 952.098, p < 0.0001$ ). The  $t$  statistics showed the same significant predictive relationship ( $t = 30.856, p < 0.0001$ ). The regression analysis generated a set of coefficients to formulate the *Regression Equation 3*:

$$(\text{PERFORMANCE}) = 43.196 + 1.841 (\text{MOTIVATION})$$

According to the effect size of this model ( $R^2 = 0.896$ ), *motivation* predicted approximately 90% of the variation of the response variable *performance*. The linear interpolation found the power of the  $F_{(1, 111)}$  test to be 0.98, indicating a relatively high probability of finding a significant linear relationship in the data from the multiple-layer group.

In summary, the linear relationship between *motivation* and *performance* existed in general, as well as in each group; and in the single-layer group, the portion of variation (90%) in *performance* predicted by *motivation* was relatively higher than that (68%) in the multiple-layer group. Mathematically, these results can be explained by the *Regression Equations 2 and 3*: A one-unit increase in the *motivation* score of a student in the single-layer group will increase more units on his/her *performance* score (1.841 units) than that of a student in the multiple-layer group will (1.492 units).

#### *Data Analysis Four*

Last, data from the *Survey of Student Preference* (Appendix, Part II) were analyzed, and the mean score for each of the nine items was calculated. A mean score larger than 8.00 was considered a high preference. As shown in Table 4, students expressed high preferences to one communication format, the use of discussion board (mean = 8.24); two frequency levels, weekly-based (mean = 8.66), and need-based (mean = 8.55); and two ranges, group communication (mean = 8.84), and individual communication (mean = 9.48).

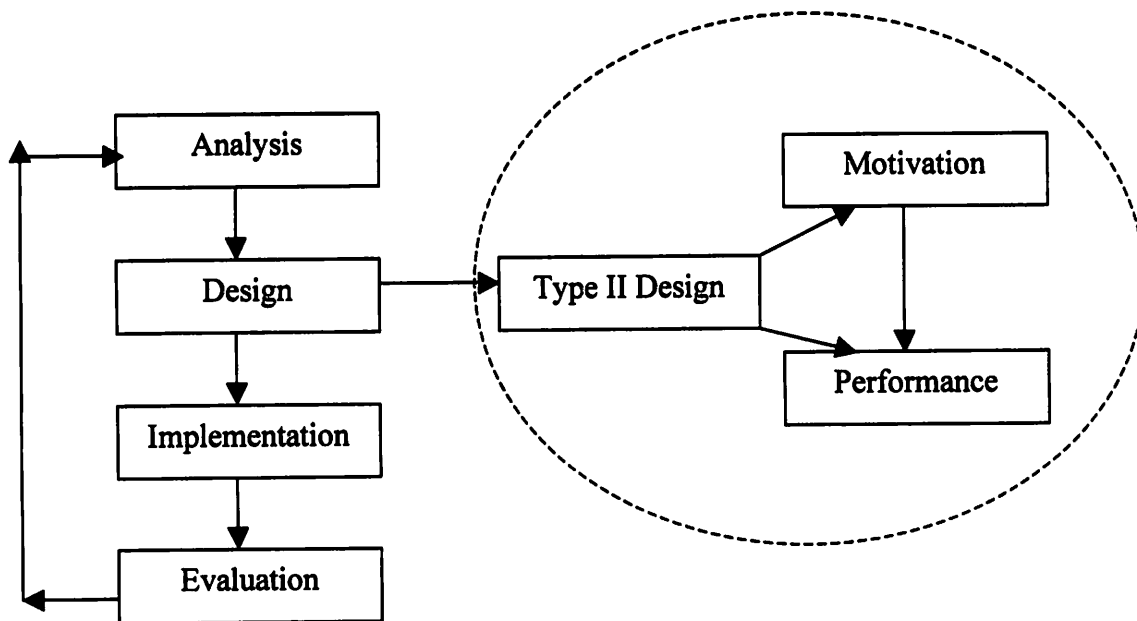
Table 4

*Student Preference to Online Communication*

		Mean N=224	SD
Format	Discussion Board	8.24	0.112
	WebCT Mail	7.02	0.524
	Chat	5.43	0.426
Frequency	Weekly	8.66	0.087
	Need-Based	8.55	0.242
	Daily	6.11	0.308
Range	Individual	9.48	0.093
	Within Group	8.84	0.165
	Class-Wide	7.86	0.243

*A Sub Model of ADIE Online Communication Design Model*

The above results added a new portion into the second phase (*design of online communication*) of the four-phase ADIE online communication design model. As shown in Figure 1, the left part of the model is the main ADIE design model, and the right part in the oval shape is the sub model – motivation based Type II online communication design model. The sub model presents the relationships in Type II design, student motivation, and performance.



*Figure 1. Online communication design model*

## Conclusions

According to the results of data analysis and the new developed model, we can draw the conclusions that (a) multiple-layer online communication design motivates students more than single-layer design; (b) students perform better in online communication activities in a multiple-layer online communication environment than in a single-layer environment; (c) students' online communication performance is a function of student motivation and can be predicted by student motivation, which is also supported in the motivation-achievement theory in a new context of an online learning environment; and (d) students prefer to use discussion board, weekly and need-based communication, group communication, and individual contact, rather than other formats, frequencies, or ranges of communication. The results are consistent with the theoretical framework summarized from the literature review and our experiences.

We want to emphasize another important finding from the current study: strategically, using Type II design (dynamic design) principles in online communication design produces more positive outcomes in student motivation and performances than using Type I design (static design) principles. We see it as important because it broadens the way people view or think about *design*; the four features of Type II design reflect a dynamic pattern that helps us to formulate a combination of appropriate decisions that best fit the needs of our students (Liu & Maddux, in press). For example, multiple-layer design can emphasize different combinations of options regarding format, frequency, and range of online communication, or different combinations with other variables related to online learning such as technology skills, accessibility, or cultural consideration.

In general, Type II design can be used in any educational application to promote learning. For example, if an instructor uses Type II design principles to design courses, lessons, assignments, projects, instructional materials, learning activities, or assessment instruments, he/she will be able to provide dynamic learning opportunities to students, and explore their potential to the greatest extent (Liu & Johnson, 2004; Liu & Maddux, in press). If designers use the four dynamic Type II features to design educational software, online learning applications, multimedia programs, or any educational information systems, they will be able to develop products that best fit the needs of learners (Liu, 2000; Liu & Johnson, 2003). We hope that the Type II design principle can be used differently in as many ways as possible by instructors and designers, and tested in as many learning environments as possible.

This study contributes to the literature of educational design theoretically and practically. Theoretically, it exhibits a new approach to design (Type II dynamic design), and it extends the traditional ADIE design model by adding the motivation based, sub model proposed Type II design model (as in Figure 1). This theoretical framework reveals many unsolved issues and problems, or new questions in the field of educational design, and provides an open-ended and multiple-dimensional research agenda. Practically, it demonstrates examples of using Type II design; other educators can easily duplicate the experiences, or apply Type II design in their own practice; and researchers can simply replicate the study, or continue examining the effectiveness of Type II design in a variety of educational applications.

## Limitations and Suggestions

We also realized the limitations of this study. First, this study used a convenient sample. The participants only include education major students and they were from the classes we taught. The results of the study reflected



the situations in these education classes, can be used as references by other educators, but may not be generalized to a broader population of online learners. Second, the sample was from classes taught by one instructor. Although this may control the possible influence from the instructor and course design, it might also be a limitation for not considering different impacts that different instructors may have on students' motivation and learning. Further studies need to be conducted focusing on the impacts of different factors from the perspective of the instructor on student online learning and communications. Third, this study only detected the impact of student motivation on online learning in general. Notice that student motivation may vary when they each perform different learning tasks. All the courses in this study had technology work and non-technology work. Further studies may investigate their motivation toward different tasks or the design of those tasks.

Fourth, this study only included one attitude variable – student motivation. Note that, in the single-layer group (Type I communication design), the portion of variation in *performance* predicted by *motivation* is relatively higher than that in the multiple-layer group (Type II communication design). Although this can be explained by the regression equations, the source of influence that causes this difference is yet to be explored. Motivation may influence student learning associated with many other variables, such as anxiety to the use of technology, interests in subject, access to the Internet, convenient help and technical support, stability of the online-course system, or instructor factors (as described above). *Motivation*, with different combinations of such variables, may contribute to a different portion of the variation in *performance*. The exploration of these related variables could produce an agenda of research for other educators and researchers to conduct further studies.

We have presented two types of online communication design, and the impact of the designs. We also recognize the limitations of the designs, and proposed problems for further research. We hope that other educators and researchers can use this first step in carrying on further studies to validate or revise our design model.

#### References

- Ames, C., & Ames, R. (1989). *Research in motivation in education*. San Diego, CA: Academic Press.
- Atkinson, J. W. (1974). The mainsprings of achievement oriented activity. In J. W. Atkinson & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 13-41). New York: Halsted Press.
- Atkinson, J. W., & Feather, N. T. (Eds.). (1996). *A theory of achievement motivation*. New York: Wiley.
- Aviv, R., & Golan, G. (1998). Pedagogical communication patterns in collaborative telelearning. *Journal of Education Technology Systems*, 26, 201-208.
- Binswanger, H. (1991). Volition as cognitive self-regulation. *Organizational Behavior and Human Decision Processes*, 50, 154-178.
- Berge, Z. (1997). Computer conferencing and the on-line classroom. *International Journal of Educational Telecommunications*, 3(1), 3-21.
- Boer, W. D., & Collis, B. (2001). Implementation and adaptation experiences with a WWW-based course management system. *Computer in the Schools*, 17(3/4), 127-146.
- Burch, J. G. (1992). *Systems analysis, design, and implementation*. Boston: Boyd and Fraser.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum.

- Crookes, G., & Schmidt, R. W. (1991). Motivation: Reopening the research agenda. *Language Learning, 41*, 469-512.
- Dembo, M. H. (2000). *Motivation and learning strategies for college success: A self-management approach*. Mahwah, NJ: Lawrence Erlbaum.
- Fishman, B. J. (1997, March). *Student traits and the use of computer-mediated communication tools: What matters and why?* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Gollwitzer, P. M., Heckhausen, H., & Ratajczak, K. (1990). From weighing to willing: Approaching a change decision through pre- or postdecisional mentation. *Organizational Behavior and Human Decision Processes, 45*(1), 41-46.
- Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169-209.
- Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research, 28*(1), 31-45.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (5<sup>th</sup> ed.). Boston: Houghton Mifflin Co.
- Kellenberger, D. W. (1996). Preservice teachers' perceived computer self-efficacy based on achievement and value beliefs within a motivational framework. *Journal of Research on Computing in Education, 29*(2), 124-140.
- Keller, J. M. (1993). Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional design theories and models: An overview of their current status* (pp. 384-434). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kelsey, K. D., & D'souza, A. (2004). Student motivation for learning at a distance: Does interaction matter? *Online Journal of Distance Learning Administration, 7*(2) Retrieved August 24, 2004, from <http://www.westga.edu/~o7Edistance/ojdlasummer72/kelsey72.html>
- Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, (Vol. I, pp. 333-378). New York: Ronald Press.
- Liu, L. (2000). Different storyboarding methods in multimedia courseware design. In D. A. Willis, J. D. Price, & J. Willis (Eds.), *Technology & teacher education annual 2000* (pp. 784-790). Charlottesville, VA: AACE.
- Liu, L. (2003). Communication design for online courses: Effects of a multi-layer approach. In G. Marks & A. Rodssett (Eds.), *E-Learn in corporate, government, healthcare and higher education annual 2003* (pp.1699-1702). Charlottesville, VA: AACE.
- Liu, L., & Johnson, L. (1998). A computer achievement model: Computer attitude and computer achievement. *Computers in the Schools, 14*(3-4), 33-54.
- Liu, L., & Johnson, L. (2003). Structure-based versus function-based method in multimedia courseware design. In C. Crawford, N. Davis, J. Price, R. Weber & D. A. Willis (Eds.), *Technology & teacher education annual 2003* (pp. 744-747). Charlottesville, VA: AACE.
- Liu, L., & Johnson, L. (2004). Static and dynamic design in online course development. In C. Crawford, N. Davis, J. Price, R. Weber, & D. A. Willis (Eds.), *Technology & teacher education annual 2004* (pp. 2946-2952). Charlottesville, VA: AACE.
- Liu, L., & Maddux, C. (2003a). Online course design and research. In G. Marks & A. Rodssett (Eds.), *E-Learn in corporate, government, healthcare and higher education annual 2003* (pp.1078-1081). Charlottesville, VA: AACE.

- Liu, L., & Maddux, C. (in press). Influences of course design on student evaluations: An initial logistic prediction model. *Journal of Excellence in College Teaching*.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.
- Locke, E. A., & Latham, G. P. (1994). Goal setting theory. In H. F. O'Neil & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 13-29). Hillsdale, NJ: Lawrence Erlbaum.
- McClelland, D. C. (1951). *Personality*. New York: Wm. Sloane.
- Means, T. B., Jonassen, D. H., & Dwyer, F. M. (1997). Enhancing relevance: Embedded ARCS strategies vs. purpose. *Educational Technology Research and Development*, 45(1), 5-17.
- Ohlend, B., Yu, C. H., Jannasch-Pennell, A., & Digangi, S. A. (2000). Impact of asynchronous and synchronous Internet-based communication on collaboration and performance among K-12 teachers. *Journal of Educational Computing Research*, 23, 405-420.
- Parkerson, J. A., Lomax, R. G., Schiller, D. P., & Walberg, H. J. (1984). Exploring causal models of educational achievement. *Journal of Educational Psychology*, 76, 638-646.
- Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching for teacher education. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 826-857). New York: Macmillan.
- Raynor, J. O. (1969). Future orientation and motivation immediate activity: An elaboration of the theory of achievement motivation. *Psychological Review*, 76(4), 606-610.
- Riel, M. (1992). A functional analysis of education telecomputing: A case study of learning circles. *Interactive Learning Environments*, 2(1), 15-29.
- Schweizer, H., Whipp, J., & Hayslett, C. (2002). Quality control in online courses: Using a social constructivist framework. *Computers in the Schools*, 19(3/4), 143-158.
- Shih, C. C., & Gamon, J. (2001). Web-based learning: Relationships among student motivation attitude, learning styles, and achievement. *Journal of Agricultural Education*, 42(4), 12-21.
- Sprinthall, R. C. (2000). *Basic statistical analysis*. Needham Heights, MA: Allyn and Bacon.
- Swan, K., Shea, P., Fredericksen, E., Pickett, A., Pelz, W., & Maher, G. (2000). Building knowledge building communities: Consistency, contact and communication in the virtual classroom. *Journal of Educational Computing Research*, 23, 359-383.
- Trentin, G. (2001). Designing online education courses. *Computers in the Schools*, 17(3/4), 47-66.
- Uguroglu, M. E., & Walberg, H. J. (1979). Motivation and achievement: A quantitative synthesis. *American Educational Research Journal*, 16, 375-389.
- Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and Higher Education*, 6(1), 77-90.
- Wishart, J. (2001). Motivation and learning styles in young children with Down syndrome. *Down Syndrome Research and Practice*, 7(2), 47-51.

*Leping Liu is an Associate Professor in the Department of Counseling and Educational Psychology at the University of Nevada, Reno; Paul Jones is Chairperson of the Department of Reading, Special Education and*

*Instructional Technology at Towson University. Correspondence concerning this article should be sent to Dr. Leping Liu, Department of Counseling and Educational Psychology, College of Education, University of Nevada, Reno, Reno, NV 89507.*

## Appendix

### Part I – Items of Motivation Measurement from the Survey:

1. I want to develop my skills in an online learning environment and study more through Web courses.
2. I don't want to take any more online courses than I have to.
3. I am interested in acquiring further knowledge through online learning.
4. I am not willing to take more than the required number of online courses.
5. I plan to take as many online courses as I can during my education.
6. I am not motivated to work very hard in an online learning environment.

### Part II – Survey of Student Preference

1. I prefer to use Webmail for our class communication with classmates and the instructor.
2. I prefer to use Discussion Board for our class communication with classmates and the instructor.
3. I prefer to use Chat Room for our class communication with classmates and the instructor.
4. I like the instructor to provide daily-based communication to the class.
5. I like the instructor to provide weekly-based communication to the class.
6. I hope we can communicate with the instructor whenever we need.
7. I prefer class-wide communication with classmates and the instructor.
8. I prefer group-wide communication with classmates and the instructor.
9. I prefer private communication with classmates and the instructor.
10. I don't feel comfortable to use Webmail for our class communication with classmates and the instructor.
11. I don't prefer to use Discussion Board for our class communication with classmates and the instructor.
12. I don't think that Chat Room is a good tool for our class communication.
13. I don't think daily-based communication with the instructor is necessary in this class.
14. I don't think weekly-based communication is appropriate for this class.
15. I don't think that communication for special need is necessary so long as I keep the regular communication with the instructor.
16. I think class-wide communication with classmates and the instructor is not efficient for our online learning.
17. I don't think group-wide communication is necessary.
18. I don't think I need private communication with classmates and the instructor.

## **A Study of Faculty Beliefs for the Teaching of Diversity in Colleges of Education: An Explanatory Mixed-method Approach**

**Xin Liang  
Susan Colville-Hall  
Suzanne MacDonald  
Lynn Atkinson Smolen  
University of Akron**

*There is a call for more inquiry to what teachers of teachers themselves need to know, and what institutional supports need to be in place to meet the complex demands of preparing teachers for the 21<sup>st</sup> century (Cochran-Smith, 2003). The purpose of this study is to explore the complexity of teacher educators' beliefs and commitment to the teaching of diversity in four universities with an explanatory mixed research design. The concept of mixed methods research is not new to teacher education research. However, simply mixing both quantitative and qualitative data in a single study is not necessarily a mixed method approach. Researchers should ask questions such as, how should the data resources be integrated and connected to unfold the research questions in the design, data collection, and data analysis procedures. The unique features of data collection and data analysis procedures for explanatory mixed method design were emphasized and discussed in this study.*

Educators are being held accountable for the social context of schooling as they plan for and deliver instruction to a more and more diverse student population (MacDonald, Colville-Hall, & Smolen, 2003). With official acknowledgement from accrediting and professional organizations such as the National Council for Accreditation of Teacher Education (NCATE) associating the education of teacher candidates with diverse students and families, and the culturally pluralistic society in which they will teach, how do universities best prepare candidates to be knowledgeable, effective, and culturally responsive to teach in more and more diverse schools? This is a common concern being posed to college of education (COE) faculty who are in charge of this task.

Researchers (Haberman, 1991; Ladson-Billings, 2001) have expressed skepticism about the commitment of teacher educators to a multicultural perspective in teacher education. Purposes and missions related to the role of teacher educators in meeting the needs of a culturally pluralistic society can be confounded by ambiguity and ambivalence within the institution and college (or department) of education itself (MacDonald, Colville-Hall, & Smolen, 2003). As many educators have begun to study the effect on the acquisition of an identified knowledge base related to the teaching of diversity, they recommend closely examining the beliefs, understanding, and commitment of those who teach future teachers (Zeichner, 1996).  
*Faculty Attitudes and Beliefs Related to Diversity*

Previous research has demonstrated a strong link between beliefs and behaviors. Pajares (1992, 2003) pointed out that attitudes, values, perceptions, understandings, and images are beliefs in disguise. The psychological frame of current theory suggests that learning is an active and meaning-making process

that is influenced by an individual's existing beliefs, attitudes, and preconceptions (Resnick, 1989). Faculty in teacher education programs come into teaching with strong theories of and beliefs about teaching and learning that have been acquired over the years from their own personal experiences of schooling. The personal experiences and beliefs direct how they perceive and react in their own classroom, working with their students – teacher candidates. Research suggests that teachers' beliefs lead to differential expectations and treatment of students. Smylie (1988) concluded in his study "teachers' perceptions and beliefs are the most significant predictors of individual change" (p. 23). Richardson (1996) concluded in his study that there is a strong relationship between beliefs and knowledge. They are strongly intertwined and form a filter through which ideas and experiences are interpreted.

The overwhelming presence of Whiteness in teacher education (Sleeter, 2001) indicates a lack of multicultural experience among teacher educators themselves. Teacher educators may sometimes be at the same level as their students in terms of diverse cultural experience (Allen & Herman-Wilmarth, 2004). Acquiring competence to teach in a culturally responsive manner requires work at the programmatic level. It also necessitates research on what takes place during teacher preparation when COE faculty bring in their own beliefs, knowledge, and dispositions. Johnson (2002) revealed in her study that teachers' perceptions of racial awareness were influenced by a perceived identity, living and working experience with individuals of other races, and personal beliefs. Teacher educators need to be connected with community-based cross-cultural immersion, provide multicultural course work with field experiences aligned with school-university partnerships, and engage in classroom dialogue within a social context that allows individuals to examine their belief systems (Allen & Herman-Wilmarth, 2004). If experience, knowledge, and dispositions are intertwined with COE faculty's belief systems to educate culturally responsive teacher candidates, we should identify COE faculty's own experiences and beliefs regarding cross-cultural immersion, multicultural coursework, comfort level with discussing beliefs, and racial and ethnic identity that connect them as teacher educators and agencies for change.

In order to understand COE faculty's commitment to multicultural education and to the development of effective strategies for preparing future teachers to teach in a diverse society, it is essential to investigate the beliefs and values the COE faculty hold and bring with them to the teacher education program, and how these perceptions and values are shaped and intertwined with their knowledge to guide and direct behavior as teacher educators. The purpose of this study was to answer such questions with both quantitative and qualitative data. Diversity in this study was defined as cultural and personal experience with people who are different. The questions that guided this study were: (a) How do COE faculty perceive the importance of diversity and what is their commitment to diversity? and (b) Which variables differentiated COE faculty's beliefs and commitment to diversity?

## Method

### *Research Design*

A mixed research method with an explanatory approach was designed to answer the research questions. The qualitative analysis, as an explanatory approach, describes and explains the reasons for the

quantitative results (Creswell, 2002). We conducted an integrative analysis of quantitative and qualitative data in which quantitative results were used to initiate the trend of the diversity experience of COE faculty, the values and beliefs they bring to teacher education programs, and the variables that differentiate their beliefs and commitment. The quantitative results were also used to provide a context for the qualitative analyses. The qualitative data provided a discussion about the themes emerging from the data and how they support or refute the statistical analysis.

#### *Participants*

The participants in the study were full-time faculty in the COE from four institutions that house teacher education programs. The student population in the four research sites (Site A, Site B, Site C, and Site D) ranged from 12,500 to 33,747 where 10% to 22% of the total student population was minority students. Among the 116 participants, 51 (44%) were males and 65 (56%) were females. The majority or 95 (81.9%) identified themselves as European Americans, while 21 participants (18.1%) identified themselves in the African American/Other category, which combined African Americans, Asians, Hispanics/Latinos, and Native Americans/Pacific Islanders. Most faculty were between 40 to 59 years old (91.2%), and represented a fairly even split among the ranks of assistant, associate, and full professor.

#### *Data Collection Procedures*

*Quantitative data.* We administered a 44-item self-report survey to the COE faculty. The first 24 items asked for demographic and diversity experience information. The other 20 items measured faculty perceptions of the importance of diversity using a 7-point Likert-type scale where 1 = *strongly disagree* and 7 = *strongly agree*.

Prior to the final instrument, a pilot survey was conducted in one of the four institutions to examine the item structure and internal consistency. Eight content experts participated in the instrument development; all were actively engaged in teaching diversity courses or included diversity issues in their courses. As a result of this pilot survey, five sub content areas were selected to measure COE faculty beliefs and perceptions on diversity. The first subscale was the importance of diversity training of teachers (General), for which the internal reliability (4 items) was .84. The second was college support of diversity training (College Support), for which the internal reliability (3 items) was .72. The third was emphasis on diversity training in college coursework (Courses), for which the internal reliability (7 items) was .82. The fourth subscale was comfort level with discussing racial issues (Comfort), for which the internal reliability (3 items) was .86. The last subscale was how people perceive you according to your race/ethnicity (Racial Sensitivity), with an internal reliability (3 items) of .75. The overall reliability was .94.

All full-time faculty members at the four state institutions (N=205) received a request for a response to the survey. One hundred and sixteen COE faculty (56.6%) agreed to participate in the study.

*Qualitative data.* We included three qualitative data resources in the study: open-ended questions in the survey, semi-structured interviews with individual COE faculty, and focus group interviews. Both individual interviews and focus group discussions were conducted as follow-up data collection procedures to obtain a detailed description of the participants' perspectives and experiences regarding diversity issues



in their home college. Ten faculty members were interviewed individually. Five focus group interviews were held in the four institutions with approximately 4-5 faculty in each group. Both the interviews and focus groups were audio-recorded and transcribed by two independent assistants. We directed qualitative data collection with the following questions: (a) Why was there a discrepancy among the COE faculty between perceptions and commitment to diversity? (b) How had the identified four independent variables impacted the perceptions and determined the commitment? (c) To what extent did the qualitative data support and/or raise questions regarding the quantitative findings of the study? and (d) What explanatory material was offered to account for COE faculty's feelings, reactions, and experiences in the context? The open-ended questions included at the end of the survey were: (a) Do you consider diversity an important issue? Why? Why not? and (b) How does your college handle the teaching of diversity to teacher candidates?

The 10 individually interviewed participants were 5 female and 5 male faculty. Among them, 4 were African Americans and 6 were White. The COE faculty interviewed taught both undergraduate and graduate students. Some were engaged in diversity training in their teacher education program and some were not. Five focus group meetings were held in the four institutions. Participants in focus groups consisted of faculty members of different ranks (assistant, associate, and full professor), across different age groups. They represented discipline areas typical for a college of education in a midwestern university. Pseudonyms were used in this report.

Both individual interviews and focus groups were semi-structured. Initial responses to interview questions were probed further by the interviewers to invite increasingly detailed and thoughtful reports of faculty members' self-reflections, perceptions, and beliefs. The design of using the mixed method approach was to connect and communicate multiple data resources to triangulate the research perspectives and increase the understanding of the participants' perspectives (Stake, 1995).

#### Data Analysis

Quantitative data were first analyzed to grasp a general trend of faculty diversity experience, beliefs, perceptions of diversity, and factors that differentiated COE faculty's beliefs and perceptions regarding diversity. We compared 24 independent variables (IV) in the first part of the survey with the five subscales as dependent variables. The 24 IVs included site, age, gender, race, state, language, second language, specialty, education, family, rank, appointment status, teacher preparation experience, public school experience, teaching environment, community location, experience in pre K-12, school location, time spent in a diverse community, types of teaching settings, requirement of diversity course, contact with diverse students, self-identity, and travel outside the US. Multivariate analysis of variance (MANOVA) was conducted to compare the differences when the analyses involved two or more dependent variables (Mertler & Vannatta, 2001). The major advantage of conducting MANOVA was that the influence of several dependent variables (the five subscales) was considered in combination with one another, thus creating a more powerful test. One of the assumptions of conducting MANOVA is that the multiple subscales together consist of different measures essentially measuring the same thing--view of diversity. At

a minimum, the dependent variables should have some degree of linearity and share a common conceptual meaning (Stevens, 1992). Based on this assumption, prior to conducting the MANOVA, correlation coefficient analysis was conducted to determine if some degree of linearity existed among the five subscales to measure a common conceptual meaning - diversity perception. The results of correlation coefficient analyses demonstrated that there existed a significant relationship among the five subscales, as was reported previously in this study. MANOVA analysis results indicated that four variables were statistically significant. These variables were site, amount of travel, self-identity, and the amount of contact with undergraduate students of diverse backgrounds. Based on the initial comparisons to determine the significance, further comparisons were conducted to obtain detailed differences for each significant item via Bonferroni post hoc tests.

Qualitative analysis was initiated by taking the quantitative outcomes as starting points in a defined context. We performed a content analysis in which all the data were partitioned into content domains for the comparison of the themes across individual responses. To ensure the accuracy and trustworthiness of the qualitative data analysis, three other members of the research team responded to and verified emergent themes initiated by the first author.

Data analysis included open coding of all transcripts by the first author. As themes emerged from the data, the research team constructed a report, analyzed the content of open-ended questions and interviews, and developed results by comparing the themes with quantitative data. The themes derived from the open coding by the first author were compared by the other researchers. The goal was to understand and explain the quantitative findings by taking the participants' reflections and voices into account. The two sets of findings are synthesized through a single discussion section.

## Results

### *Quantitative Findings*

#### *Discrepancy between perceived importance and importance of commitments.*

Descriptive analysis indicated that the COE faculty's educational, ethnic, and cultural backgrounds, along with their diversity experience, were not very diverse. Although ethnographically diverse, 81.9% were White. Most (68%) grew up in predominantly European-American communities. Over half (59.1%) rarely or never had any contact with ethnically and racially different individuals in their pre K-12 education. Over 98% of the faculty spoke only English. Less than one third (26.8%) traveled extensively outside the US. Almost half (43.2%) of the faculty had no pre K-12 classroom teaching experience in settings with diverse student populations.

The item means measuring COE faculty perceptions and beliefs indicated a wide range. The difference between the highest and the lowest item means was 3.45 on a 7-point Likert scale. It was notable that the higher scored items tended to be related to perceptions while the lower mean items were more about faculty's perceptions of their commitment to diversity. For example, the highest item, 'Teachers should be trained to effectively teach diverse students' had a mean of 6.19. As to what the COE faculty do

in their own classroom teaching, the item rate was much lower. For example the lowest item, 'Issues of diversity are the main focus of the classes,'  $M=2.74$ . Other similar items such as 'I emphasize diversity in classes I teach,' and 'I bring in outside speakers to address diversity in my class' were among the lowest rated items.

*Items that differentiate COE faculty's beliefs and perceptions.* MANOVA analysis suggested that four variables were statistically significant at the .05 level: institution sites (Wilks' Lambda > .001); travel outside the US (Wilks' Lambda = .014); self-identity (Wilks' Lambda = .003); and contact experience (Wilks' Lambda = .027). Further group comparisons were conducted to obtain detailed differences using the five subscale dependent variables, site, travel, self-identity, and contact experience as independent variables via Bonferroni post hoc tests.

There was a significant difference among COE faculty in the four institutions in responding to General, College Support, and Racial Sensitivity subscales. The mean differences ranged from -.948 to -1.37. There was a significant difference between the COE faculty who traveled extensively and those who traveled some in responding to the General subscale ( $MD = .623, p = .032$ ). There was a significant difference among the COE faculty who perceived their self-identity as connected greatly, moderately, little, or not at all in responding to General, Course, and Racial Sensitivity as mean differences ranged from -.869 to -1.96. There was a significant difference among the COE faculty who had most, some or no contact experience with diverse students in responding to General, Course, Racial Sensitivity, and Comfort. Mean differences ranged from -1.22 to -2.21

As expected, the other 20 variables did not turn out to be significant. The research team was interested in knowing more about the impact of these variables - sites, self-identity, contact experience, and the amount of travel outside the United States on COE faculty's beliefs and efforts regarding diversity in the defined five aspects. The research team also wanted to know why COE faculty rated some items more important than others.

#### *Qualitative Findings*

*General consensus on the importance of diversity.* The overwhelming majority of the COE faculty believed that diversity has become a survival issue for future teachers who will teach in an increasingly diverse society. For example, Marie responded to the impact of the school population shift with, "The majority of students are going to be minority. One-hundred percent of the time, diversity is linked to society's survival, literally."

However, some faculty did express indifference. Martin stressed, "The mistake made by African-Americans was not in choosing integration over segregation, but in choosing integration over education." For a faculty member like Martin, it seemed the current college culture neglected the biggest issue--student learning, which was overshadowed by the distraction of diversity. He tended to perceive diversity as an independent political issue different from a more "knowledge based" concept--effective teaching and learning.

*Institutional efforts-A mixed perception.* “Don’t know” was a very frequent response when the faculty were asked about how their home college viewed and prepared teacher candidates to teach in diverse settings. Mark’s response represented the lack of faculty involvement: “We have your diversity course. I thought you all were handling that.” Diversity issues were considered an important part of the college mission because of the pressure from NCATE. Judith commented, “People make verbal commitment to it. Helping make decisions tends to be a question. The reason is because many of us have not had any training in terms of diversity. We should not have to scuffle and scramble when NCATE comes around.” A lack of support and recognition in the home college was another concern. “However, the money just isn’t there in the same way with diversity as with technology. The services we provide certainly are not recognized for either tenure or promotion.” (Jeff).

*Fear of the unknown.* COE faculty shared their observations of the fear existing among teacher candidates to teach in urban schools where the student population is more diverse. “Most of my students expect to teach in communities where they grew up, (e.g., all White, semi-rural)” (Allen). The COE faculty believed that the fear was really a fear of the unknown related to the previous family and cultural experiences they brought with them to the teacher education program. Louise shared her experience with this fear. “I took the class down to XY High School, and I said this is a city school. Is there any graffiti on the walls? Nope. The students were just shocked to see how orderly it was down there, and I believe it is the unknown that scared them away.”

*Self-identity, modeling, and cultural immersion.* The role model of teachers and the school climate played an essential part in rethinking the meaning of self-identity in relation to one’s own perceptions, beliefs, and biases. “Our students are nothing other than a reflection of ourselves. How many of our faculty here would be comfortable going to “XYZ” (an inner city school)?” Sam shared, “My own personal background does have an effect on how I approach the topic and how I deal with it in the classroom.”

In our discussion with the participants, the COE faculty asserted that they could change beliefs; even deep-seated ones, by engaging students in meaningful educational experience. They were hopeful that what they were doing would ultimately contribute to the change process. Melissa expressed this hope: “I think we have more power than we sometimes think we do. I had a student in Georgia who was very prejudiced against Mexicans and talked about how lazy they were. We really challenged him in class about those beliefs. When he finally graduated, he wrote me a note ‘I did not always agree with you, but really you made me think.’ He is probably still thinking the same... At least we gave him the opportunity to think, to see a different perspective, and that’s worth something.”

#### Discussion

The study was undertaken with an explanatory mixed method research approach to answer research questions on the COE faculty’s perceptions and beliefs in regard to diversity issues in teacher education programs. The descriptive analysis entailed the COE faculty’s diversity experience. It also revealed that some issues were perceived to be more important. MANOVA analysis identified four factors that significantly influenced the COE faculty’s perceptions and beliefs regarding the importance of

diversity in the aspects of General Importance, Coursework, College Support, Comfort Level talking about diversity and Cultural Sensitivity. The comments from the open-ended questions in the survey described how the COE faculty in the four institutions perceived the importance of and commitment to diversity, and why they had such perceptions. The feelings and experiences that COE faculty shared with the researchers in the interview and focus group discussions also enriched our understanding of how these different factors were shaped. At a deeper level, qualitative data documented how and why the challenges and barriers were embedded individually and institutionally.

The data results revealed within and between group differences in faculty's perceptions regarding the importance of diversity. Most importantly, aligned with NCATE standards, the meaning and task of diversity based on the missions and goals of each individual college were left undefined. There was enormous inconsistency in faculty members' knowledge, information and depth of understanding about diversity issues (Kitano, Lewis, Lynch, & Graves, 1996). From this aspect, much needs to be addressed regarding the lack of professional development, financial support, systematic evaluation criteria, and institutional recognition concerning the faculty in the four institutions. There needs to be a clarified definition of what diversity training means within the college, and how the issues are related to each individual faculty, each class, field experience, and student teaching experience. These findings have implications for future research, and potentially, for establishing goals for teacher education programs to enhance multicultural competency of COE faculty and teacher candidates.

To explore the perceptions, beliefs, possibilities and constraints in regards to the roles of teacher educators as change agents requires systematic inquiry. One of the challenges for this endeavor is the lack of a knowledge base supported by rich data resources. Cochran-Smith (2003) summarized the various forms of practitioner inquiry such as action research (Dahlstrom, Swarts & Zeichner, 1999; Noffke, 1997), self-study (Hamilton, 1998), and autobiographical inquiry (Florio-Ruane, 2001).

The mixed research method, integrating both statistical analysis and participants' reflections and experiences to explain a social phenomenon, adds a scientific component to the research outcome. The journey of conducting a mixed method research in general, and an explanatory mixed method research design in particular, in this study presents the dynamics and complexity of data collection and data analysis to capture the trends and the details of the faculty's perceptions and beliefs regarding the multicultural agenda across the four institutions to integrate and connect multiple data resources in a meaningful way to enhance the power of the research outcome.

## References

- Allen, J., & Herman-Wilmarth, J. (2004). Cultural construction zones. *Journal of Teacher Education, 55*, 214-226.
- Cochran-Smith, M. (2003). Learning and unlearning: The education of teacher educators. *Teaching and Teacher Education, 19*, 5-28.
- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.
- Creswell, J. W., Fetters, M. D., & Ivankova, N. V. (2004). Designing a mixed methods study in primary care. *Annals of Family Medicine, 2*, 7-12. Retrieved August 18, 2004, from <http://www.annfammed.org>
- Dahlstrom, L., Swarts, P., & Zeichner, K. (1999). Reconstructive education and the road to social justice: The case of post-colonial teacher education in Namibia. *International Journal of Leadership in Education, 2*, 149-164.
- Florio-Ruane, S. (2001). *Teacher education and cultural imagination*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Haberman, M. (1991). Can cultural awareness be taught in teacher education programs? *Teaching Education, 44*, 25-31.
- Hamilton, M. L. (Ed.). (1998). *Reconceptualizing teaching practice: Self-study in teacher education*. London: Falmer Press.
- Johnson, L. (2002). "My eyes have been opened": White teachers and racial awareness. *Journal of Teacher Education, 53*, 153-167.
- Kitano, M., Lewis, R., Lynch, E., & Graves, A. (1996). Teaching in a multicultural classroom: Teacher educators' perspectives. *Equity and Excellence in Education, 29* (3), 70-77.
- Ladson-Billings, G. (2001). *Forum two: Programmatic issues in urban education*. Cleveland, OH: Cleveland State University
- MacDonald, S., Colville-Hall, S., & Smolen, L. (2003). Program ownership and the "transformative agenda" in colleges of education: Teachers of teachers working towards a more equitable society. *Multicultural Education, 10* (4), 7-14
- Mertler, C. A., & Vannatta, R.A. (2001). *Advanced and multivariate statistical methods: Practical application and interpretation*. Los Angeles, CA: Pyrczak Publishing.
- Noffke, S. (1997). Professional, personal, and political dimensions of action research. In M. Apple (Ed.), *Review of research in education* (Vol. 22, pp. 305-343). Washington, DC: American Educational Research Association.
- Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*, 307-332.
- Pajares, F. (2003). Self-efficacy, beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly, 19*, 139-159.

- Resnick, L. B. (1989). Introduction. In L. B. Resnick. (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 1-24). Hillsdale, N J: Lawrence Erlbaum Associates.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (pp. 102-119). New York: Simon & Schuster/Macmillan.
- Sleeter, C. (2001). Preparing teachers for culturally diverse schools: Research and the overwhelming presence of Whiteness. *Journal of Teacher Education*, 52, 94-106.
- Smylie, M. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. *American Educational Research Journal*, 25, 1-30.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2<sup>nd</sup> ed). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zeichner, K. (1996). Educating teachers for cultural diversity. In K. Zeichner, S. Melnick, & M. Gomez (Eds.), *Currents of reform in preservice teacher education* (pp. 133-175). New York: Teachers College Press, Columbia University.

*Xin Liang is an Assistant Professor and Suzanne MacDonald is an Associate Professor in the Department of Educational Foundations & Leadership at the University of Akron; Susan Colville-Hall is an Associate Professor and Lynn Atkinson Smolen is a Professor in the Department of Curriculum and Instructional Studies at the University of Akron. Correspondence concerning this article should be addressed to Dr. Xin Liang, Assistant Professor, Department of Educational Foundations & Leadership, University of Akron, Akron, OH 44325-4201.*

**THE UNIVERSITY VOLUNTEER TUTORING EXPERIENCE:  
THE WHO, WHAT, AND WHY OF WORKING WITH URBAN ELEMENTARY SCHOOL TUTEES**

**Gary W. Ritter, Kristin K. Higgins, and Rebecca A. Newgent  
University of Arkansas**

*This study explores the effects of a large-scale university-assisted partnership on the college-aged tutors by asking what types of students participate, why these students tutor, and if and how the tutors benefit. Data gathered from a Likert type survey and from interviews with the nearly 400 college-aged tutors reveal that the tutors participate in the program out of a desire to work with children, a sense of social responsibility, and an interest in exploring educational vocations. Overall, the data suggest that volunteer tutors' expectations were indeed met and that they accrued real benefits from participation in the program.*

There is abundant evidence that urban elementary schools are facing staggering challenges. For example, urban schools serve communities with a 67% higher rate of child poverty than for the nation as a whole (Jerald & Curran, 1998). Children in urban schools are three times more likely to live in impoverished neighborhoods. Further, fourth graders in urban districts are nearly 30% less likely than the average fourth grader in the United States to have even basic [italics added] reading skills as measured by the National Assessment of Educational Progress (Jerald & Curran, 1998, p. 56).

While schools in urban areas are often those in greatest need of help, untapped community resources often surround these schools. Among the greatest of these community resources are the numerous institutions of higher learning. As a result, many cities have witnessed the development of partnerships between urban school districts and local colleges and universities. President Clinton supported this development by introducing the America Reads Challenge in 1997, a program that challenges institutions of higher education nationwide to increase their commitment to volunteerism within the surrounding local communities (White House, 1997). More recently, President Bush's "No Child Left Behind Act," (2002) provides opportunities for tutoring programs to play a critical role as a supplementary service to assist at-risk children in receiving quality education.

During the Clinton presidency, the growth of these university partnerships accelerated as a result of the America Reads Challenge, launched in 1997 (White House, 1997). Further, President Clinton's 1999 budget proposal included \$140 million to establish programs matching university-based mentors with students in schools that have very high dropout rates and high concentrations of poor students. In support of such programs, former Education Secretary Richard Riley described community and school partnerships as essential components to educational excellence (Sandham, 1998). In this same vein, with the support of current Education Secretary Rod Paige, President Bush has included opportunities for supplemental services in the legislation, including tutoring (Bush, 2001).

Leaders at other levels of government have become increasingly interested in the concept of volunteer tutoring for elementary school children. For example, Governor Gary Locke of Washington called on legislators to create a state corps of 25,000 tutors to help students in grades two through five (Trotter, 1998). The Washington



Reading Corps provides thousands of tutors every year to boost reading achievement in the elementary schools. Since 1998, this program has enabled 11,000-trained volunteers to give intensive tutoring to 22,000 elementary students statewide and has had a positive impact on reading ability (Governor Locke's Major Accomplishments 1997-2002, n.d.).

Volunteer tutor programs can serve an important purpose by providing an opportunity for personal growth for tutors and tutees alike as well as an opportunity for cross-cultural interactions that the participants may not otherwise encounter. Many of these university-based tutoring programs have dual goals of serving both the tutors and the tutees. While there is an emerging research base examining the impact of these programs on the academic performance of the tutees, we know less about impacts on the tutors themselves. As a result, it is also important that researchers systematically analyze the benefits received by the volunteer tutors as well as the benefits for the tutees.

To add to the research base in this area, this paper reports on the results of an evaluation of the West Philadelphia Tutoring Project (WPTP), a large-scale university-assisted partnership focusing on volunteer tutoring of urban elementary students. Unlike most volunteer tutoring studies that solely look at the effects on the tutees (Baker, Gersten, & Keating, 2000; Morris, Shaw, & Perney, 1990; Ritter, 2000, 2001; Vadasy, Jenkins, Antil, Wayne, & O'Connor, 1997), this study explores the effects on the volunteer tutors. Knowing what types of students volunteer is important in terms of actively recruiting new volunteers as well as in targeting certain types of students in order to diversify and increase the tutoring pool. In addition, by discovering why tutors participate and what benefits were derived from this tutoring experience, recruiters of volunteers will have more information to share with potential tutors. The key questions addressed are: (a) What types of students volunteer to tutor, (b) why do these students volunteer to tutor, (c) do they benefit from the volunteer experience, and (d) how do they benefit from their volunteer tutoring experiences?

## Method

### *Participants*

Three hundred seventy-nine students at the University of Pennsylvania chose to participate in a volunteer tutoring program in the 1998-1999 academic year. Participants (volunteer tutors) included 64% females and 36% males; data presented here are drawn only from those participants who completed the year-end tutor surveys (N = 288).

### *Instrument*

The Year-End Tutor Survey (YETS; Ritter, 2000) is a 50-item survey using Likert responses and both closed and open-ended questions that was administered during the final week of the volunteer tutors' participation in WPTP. Internal consistency reliability (Cronbach Alpha) of the 21 survey items related to the type of experience and benefits received by the tutors was .81. Questions focused on the following: tutee school information, demographic data on the tutor, other important background information on the tutor, school and program factors, WPTP and its impact on the tutee, WPTP and its impact on the volunteer tutor, issues on training, personal views and thoughts about the future, importance of other individuals in the WPTP experience, and questions related to whether or not the volunteer tutor would participate in the program again in the future. The information from the year-end survey was augmented by comments made during weekly electronic discussions among the WPTP participants.

## Results

### *The Volunteers: Who Are They?*

The modal University of Pennsylvania tutor is predominantly female, in the first year of WPTP involvement, in the first or second year of study, and majoring in a variety of academic disciplines. Seventy-four percent of the tutors were new participants with no previous WPTP experience. Ninety-one percent of the tutors had previous experience in working with children. Sixty percent of the volunteers had received formal training such as a classes or workshops focusing on working with children. Sixty-nine percent of the volunteers had previously tutored in a program similar to the WPTP program. Finally, 63% of the volunteers shared that when growing up they did not know or have any experience, with people living in similar neighborhoods as the ones in which they tutored.

### *Why Do They Volunteer?*

Program participants indicated a strong interest in working with children. Of the 200 respondents who answered open-ended questions related to reasons for participation, 28% listed working with children as a key reason. A strong interest in children was by far the most popular reason for involvement. While there were several other reasons given by many volunteers, no other response was given by more than 10% of the respondents. Many volunteers described rewarding experiences with tutoring in the past and stressed a sense of moral and social responsibility to participate in community service. Other volunteer tutors chose to tutor to explore the possibility of a career in education and gain some practical teaching experience. Finally, a few volunteers signed up as a way to meet new people and become more involved in university activities.

### *Do They Benefit?*

Ninety eight percent *strongly agreed* or *agreed* that the experience was beneficial to them and 84% shared that they received what they had wanted from the experience. In addition, 86% *strongly agreed* or *agreed* that the experience made them more effective tutors and 63% indicated that they gained more out of the experience than their tutee did. Seventy-nine percent of the volunteers *agreed* or *strongly agreed* that they developed a close relationship with their tutees. In addition, 70% of the tutors shared that they would definitely plan to participate in the program next year.

### *How Do Tutors Benefit?*

Ninety-one percent of the volunteers *strongly agreed* or *agreed* that their volunteer tutoring experience encouraged them to do more volunteer work in the future and 50% of the volunteers *strongly agreed* or *agreed* that the experience had influenced their future career choice. In terms of how the experience affected tutors' view of disadvantaged students, 37% commented that their views had been changed for the better, 58% stated there had been no change, and only 5% stated that their views had been changed for the worse. In addition, 58% of the volunteers indicated that their views regarding programs that benefit disadvantaged students had become more optimistic, while only 11% claimed that their views had become more pessimistic.

Analysis of the open-ended responses of the participants indicated that 30% of the respondents perceived the program to be rewarding because of the feeling that they were making a difference in the lives of children. This feeling of having a positive impact was the most consistent benefit of program participation reported by the

volunteers. Smaller numbers of participants indicated that the program benefited them by exposing them to the field of teaching, by allowing them to meet new friends, and finally by exposing them to people of diverse backgrounds.

#### Conclusions: Were Expectations Met?

In evaluating the expectations of the tutors and the benefits that the tutors derive from the experience, the survey and interview data suggest that volunteer tutors' expectations were indeed met. The university students claimed to volunteer their time to tutor because of their love for children and their desire to help children grow and learn. Data from the survey and interviews indicate that the tutors were pleased to engage in relationships with children with backgrounds that are different than the tutors' backgrounds. After a year of tutoring, the tutors were able to learn more about the field of teaching and education, and many expressed that the tutoring experience had validated their desire to enter the teaching profession. The survey responses indicated that volunteer tutors left the WPTP experience having a sense that they had a positive impact on the students they tutored as well as on the wider community. Volunteer tutors were able to meet other university students involved in the program and to become involved in campus activities. Thus, it is certainly reasonable to conclude that the program succeeded in exposing the tutors to new experiences, which led to personal benefits.

#### Implications

Certainly, there are limitations to a study such as this. The conclusions presented here are reflective of the responses of those tutors who responded to the year-end survey, which represents only three-quarters of the students who volunteered throughout the year. Further, it is possible that the students willing to complete a year-end survey were more satisfied with the program than the others and, consequently, their views may not be representative of the entire group of tutors.

Nonetheless, the available data and the results presented indicate that the tutors from the West Philadelphia Tutoring Project entered the tutoring partnerships with numerous expectations and that most of these expectations were met by the end of the academic year. While most respondents indicated satisfaction with the program, a few tutors, however, shared some concerns with the program operations. These issues related to a lack of structure in some of the sites and a concern that the tutoring was not resulting in academic improvement for the young students. As a result of these matters raised by the tutors, along with program observations and knowledge of the tutoring literature, several recommendations for university-based tutoring programs that may increase program effectiveness and can be incorporated without substantially changing the service-learning venture. The two primary suggestions are: (a) encourage that tutors participate in an orientation to service-learning that would help them feel more connected to the project and to the university, and (b) allow volunteers the option of participating in a shadowing experience where they may observe mentor tutors and instructors, which may aid in future career decision-making. Thus, the tutors who hope to make an effort to effect academic performance would have that opportunity while not diminishing the opportunities of others who prefer a lesser level of involvement.

Additionally, there are implications with respect to the recruitment of tutors into the program. Despite the fact that nearly 100% of the tutors claimed to benefit from the program and 84% claimed to receive exactly what they had hoped from the experience, most of the tutors do not continue on in the program from year to year. During this year, nearly three out of four tutors were new to the program. The tutors have indicated that they are, for the

most part, satisfied with the program. Key program stakeholders acknowledge the importance of the continuity of tutoring partnerships; thus, it follows that the WPTP should make a greater effort to retain existing tutors, for example, through more flexible scheduling. This change might also allow the program to make some progress toward the goal of improving the academic achievement of the students in the program as the added experience will likely lead to more effective tutoring.

Overall, these results add to the existing literature and suggest that the tutors can experience personal growth and enjoy many other benefits from participating as tutors in volunteer tutoring programs. If both the tutors and tutees are experiencing personal growth from such programs, the suggestion is that such programs are worthwhile and should continue. Certainly, program developers should continue to work on strategies and methods such that these partnerships have a greater chance to inspire academic gains for the young tutees. In the meantime, however, it is important that policymakers continue to recognize the real benefits, albeit less tangible benefits enjoyed by the program participants in volunteer tutoring partnerships across the country.

#### References

- Baker, S., Gersten, R., & Keating, T. (2000). When less may be more: A 2-year longitudinal evaluation of a volunteer tutoring program requiring minimal training. *Reading Research Quarterly, 35*, 494-519.
- Bush, G. W. (2001). *Press conferences with President George W. Bush and Education Secretary Rod Paige to introduce the President's education program*. Retrieved January 24, 2003, from <http://www.whitehouse.gov/news/releases/2001/01/print/20010123-2.htm>
- Bush, G. W. (2002). *Next step in education reform*. Retrieved January 24, 2003 from <http://www.whitehouse.gov/infocus/educationreform>
- Governor Locke's major accomplishments 1997-2002*. (n.d.). Retrieved January 24, 2003, from <http://www.governor.wa.gov/accomplish/accomplishments.htm>
- Jerald, C. D., & Curran, B. K. (1998, January 8). By the numbers: The urban picture. *Education Week*, p. 56.
- Morris, D., Shaw, B., & Perney, J. (1990). Helping low readers in grades 2 and 3: An after-school volunteer tutoring program. *The Elementary School Journal, 91*, 133-150.
- Ritter, G. W. (2000). The academic impact of volunteer tutoring in urban public elementary schools: Results of an experimental design evaluation. *Dissertation Abstracts International, 61*(3), 890A. (UMI No. 9965556)
- Ritter, G. W. (2001, April). *The evidence on the effectiveness of volunteer tutoring in the United States: Implications for policy and practice*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Sandham, J. L. (1998, January 28). Where gown meets town. *Education Week*, pp. 26-30.
- Trotter, A. (1998, January 28). Locke promotes 'Reading Corps.' *Education Week*, p. 15.
- Vadasy P. F., Jenkins, J. R., Antil, L. R., Wayne, S. K., & O'Connor, R. E. (1997). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. *Learning Disability Quarterly, 20*, 126-139.

**White House, Office of the Press Secretary. (1997). *President Clinton's America reads work-study challenge*. Washington, DC: Author.**

*Gary W. Ritter is an Associate Professor, Kristin K. Higgins is a Doctoral Candidate, and Rebecca A. Newgent is an Assistant Professor in the Department of Educational Leadership, Counseling and Foundations at the University of Arkansas. Correspondence concerning this article should be addressed to Dr. Gary Ritter, Educational Leadership, Counseling and Foundations, 241 Graduate Education Building, University of Arkansas, Fayetteville, AR 72701. Phone: 479-575-4971.*

## **Methodological Issues Section**

I am pleased to offer another section pertaining to methodological issues. This section is intended to provide an opportunity for the readership to learn about and generate a discussion about a variety of issues as they pertain to educational researchers. The current issue contains three articles that address data analysis issues facing educational researchers engaged in quantitative and qualitative research.

Fraas and Drushal's article discusses the use of logistic regression when dependent variables are dichotomous (e.g., passing or failing). Logistic regression is useful in determining the probability of a person belonging to a particular group (e.g., passing the class). As with other quantitative procedures, it is important to communicate findings not only in terms of their statistical significance, but also their practical significance. Fraas and Drushal offer a thorough illustration of the use of Delta- $p$  values to communicate changes in probabilities associated with changes in the independent (predictor) variable. To assist with the implementation of this method, they offer a computer program, which is used in conjunction with Microsoft Excel.

Another important issue faced by researchers is the determination of the validity of the instrumentation they use. Witte's "Learning Styles and Mindstyles: An Examination of Validity" focuses on the validity of two instruments (MMPALT III and Gregorc Style Delineator). Witte describes an approach to support the validity of these instruments using internal consistency reliability, correlations, and confirmatory factor analysis.

Kawulich's "Data Analysis Techniques in Qualitative Research" focuses on the analysis of qualitative data. Many researchers can easily become overwhelmed by the task of analyzing the variety and amount of data that result from transcribed interviews, observations, and artifacts. Kawulich offers a framework for organizing this task and practical guidelines for establishing sound coding and category systems, analyzing data, and reporting findings. Furthermore, they discuss issues and challenges faced by qualitative researchers and teaching qualitative research to others.

I hope this section generates discussion among the readership and generates future submissions for consideration in this section. If you would like to submit a manuscript for this section, or simply discuss some ideas, please contact:

David M. Shannon  
Professor  
Auburn University  
4036 Haley Center - EFLT  
Auburn University, AL 36849-5221  
(334) 844-3071  
[shanndm@auburn.edu](mailto:shanndm@auburn.edu)

**EXPRESSING LOGISTIC REGRESSION COEFFICIENTS AS  
DELTA-P VALUES: PRESENTING INFORMATION TO DECISION MAKERS**

**John W. Fraas  
and  
J. Michael Drushal  
Ashland University**

*A rationale for reporting Delta-p values as a means of assisting researchers and decision makers in gauging the practical significance of independent variables in studies that use logistic regression models is presented. To assist in the implementation of this method, a computer program, which is used in conjunction with the Microsoft® Excel software, was designed to compute these Delta-p values. This program is available from the authors.*

It is not uncommon for program evaluators and researchers to encounter a variety of situations in which the dependent variable of interest is dichotomous (i.e., the variable consists of two categories). For example, Schreiber (2002) attempted to determine if certain student background and academic factors could be used to identify whether the students scored above or below the international mean on an advanced mathematical examination. In a study by McCoach and Siegle (2001) attitude variables were analyzed to determine if they could accurately identify whether a student was a gifted achiever or a gifted underachiever. Hendel (2001) examined if the students' participation in a first-year seminar and various academic and background variables were associated with whether they did or did not return for their second year of college.

When such dependent variables are encountered, it is not uncommon for researchers and program evaluators to use logistic regression analysis, as was the case for the three previously cited studies. Readers of this article who would like to either learn or review the concepts related to linear probability models and logistic regression models may find it helpful to refer to Fraas and Newman (2003), Hair, Anderson, Tatham, and Black (1998), Hosmer and Lemeshow (1989), and Norušis (1999).

One problem confronting these researchers and program evaluators is how to convey the meaning of a significant logistic coefficient in a manner that decision makers find meaningful and useful (Cabrera, 1994). Various researchers have stressed the importance of not only reporting whether parameter estimates, such as logistic regression coefficients, are statistically significant but also the need to provide information that can be used to judge the practical significance of these parameter estimates (Cohen, 1969, 1988; Fisher, 1925; Fraas & Newman, 2000; Levin & Robinson, 2000; Robinson & Levin, 1997; Thompson, 1996). Since a logistic regression coefficient measures the change in the log of the odds of the event occurring associated with a one-unit change in the independent variable, many decision makers may find such a value difficult to use in order to judge the practical significance of the estimated change. Thus the question addressed in this article is: What information regarding statistically significant logistic regression coefficients can researchers provide to decision makers that will allow them to best deal with the concept of practical significance?

A method is presented that expresses a logistic coefficient as a series of probability values where each value represents the change in a person's probability of belonging to the group assigned the value of one in the dependent variable associated with a one-unit change in the independent variable. We believe these changes in the initial probability, which are referred to as Delta-*p* values, provide researchers and decision makers with information that will assist them in judging the practical significance of the logistic regression coefficient of an independent variable. A computer program, which is used in conjunction with the Microsoft® Excel software, was designed to calculate the Delta-*p* values.<sup>1</sup>

### **Interpreting Logistic Regression Coefficients**

In a logistic regression model, the relationship between the binary dependent variable and a given independent variable is assumed to follow a logistic function (Collett, 1991). The logistic function takes the following form:

$$\ln \left[ \frac{\text{Prob}(\text{event})}{1 - \text{Prob}(\text{event})} \right] = B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p$$

where:

1. Prob(event) refers to the probability that the person will belong to the group assigned a value of one in the original dependent variable.
2. The symbols  $X_1$ ,  $X_2$ , and  $X_p$  represent the independent variables.
3. The symbols  $B_0$ ,  $B_1$ ,  $B_2$ , and  $B_p$  refer to the estimated parameters.

To estimate the logistic regression model, the logistic curve is fitted to the actual data. In the process of fitting the logistic curve to the data the values of the dependent variable are transformed to the logarithm of the odds (i.e., the log odds) that a person will belong to the group assigned a value of one in the original dependent variable. As noted by Cizek and Fitzgerald (1999), this transformation results in log odds values that are measured on an equal interval scale. Often these log odds values are referred to as logits, which is the contraction of the terms logistic and units.

The logistic transformation of the dependent variable causes the coefficients estimated for a logistic regression model to differ from those obtained from an Ordinary Least Squares (OLS) regression model. As noted by Cabrera (1994), "unlike OLS, the metric of the individual coefficients under logistic regression is expressed in terms of logits rather than in terms of the original scale of measurement" (p. 245). Thus the logistic coefficient for a given independent variable is interpreted as the change in the log odds associated with a one-unit change in that variable.

#### *An Illustration*

To illustrate the interpretation of a logistic coefficient, consider a study in which the researchers are interested in assessing the relationship between various student and financial factors recorded for students who have applied to a university and whether the students actually matriculated. In this illustrative study the dependent variable identified whether the students did or did not matriculate. Students who did not matriculate were assigned values of zero; while the students who did matriculate were assigned values of one. The following independent variables were included in the logistic regression model:



1. Variable  $X_1$  contained the students' high school grade point averages (HSGPA).
2. Variable  $X_2$  was composed of the students' ACT composite scores (ACT).
3. Variable  $X_3$  indicated the sex of each student (SEX) [ $X_3 = 0$  for a female student; while  $X_3 = 1$  for a male student].
4. Variable  $X_4$  represented the amount of financial aid offered each student (AID). The values were recorded in units of \$1,000.
5. Variable  $X_5$  indicated the amount of financial need established for each student (NEED). The values were recorded in units of \$1,000.

The results of the logistics regression analysis for a model containing these variables are contained in Table 1.

*Table 1*

*Results of the Logistic Regression Analysis*<sup>a</sup>

Variables	b	SE	Wald	df	p	Exp(b)	95% C. I. for Exp(b)		
							Lower	Upper	
$X_1$ (HSGPA)	-.484	.270	3.213	1	.759	.616	.363	1.046	
$X_2$ (ACT)	-.001	.033	.001	1	.974	.999	.936	1.066	
$X_3$ (Gender)	-.548	.207	6.980	1	.008	.578	.385	.868	
$X_4$ (AID)	.128	.045	7.932	1	.005	1.137	1.040	1.243	
$X_5$ (Need)	-.016	.013	1.374	1	.241	.984	.959	1.011	
Constant	1.125	.751	2.245	1	.134	3.080			

<sup>a</sup>  $\Delta$  (-2 log likelihood) = 15.413,  $p = .009$ ,  $n = 525$ .

*Interpretation of the estimated financial aid coefficient.*

Assume the value of interest to the researcher and the policymakers is the AID variable ( $X_4$ ). Decision makers at the university may be keenly interested in the following question: Is a change in a student's amount of financial aid associated with a change in the likelihood the student will matriculate? Typically when addressing this question the researchers would use the results contained in Table 1 and report that the logistic coefficient was statistically significant ( $b = .128$ , Wald value = 7.932,  $p = .005$ ). In addition, the researchers may report and interpret the log odds value ( $b$ ) of .128 and the odds ratio value ( $\exp(b)$ ) of 1.137.

If the researchers interpreted the log odds value ( $b$ ), they would report the log odds value of .128 indicates that an increase in a student's financial aid of \$1,000 (one unit) is associated with an increase of .128 in the log odds value, holding all other variables constant. Other than indicating that an increase in financial aid is associated with an increase in the likelihood the student matriculated, this change in the log odds value is difficult for researchers to use as a means of communicating the practical importance of such changes to decision makers. As noted by Long

(1997), "unfortunately, most of us do not have an intuitive understanding of what a change in the logit [log odds value] means. This requires another transformation" (p. 80).

The use of the odds value ( $\exp(b)$ ) of 1.137 would allow the researchers to present information to decision makers that is somewhat more meaningful. With regard to using the odds ratio value, Cizek and Fitzgerald (1999) stated that: "Information related to the odds--as opposed to log odds--of an event occurring is easier to understand and communicate" (p. 228). This odds value indicates that an increase of \$1,000 (one unit) in financial aid is associated with an increase in the odds of matriculating by a factor of 1.137, holding all other variables constant. Thus a \$1,000 increase in financial aid is associated with an increase of 13.7% in the odds of matriculating, holding all other variables constant.

In addition to reporting and interpreting the odds value of 1.137, the researchers could report and interpret the 95% confidence limits for the odds value, which is part of the SPSS® printout. For the financial aid variable ( $X_4$ ) these limits were 1.040 and 1.243. The researchers could use these limits to indicate to the decision makers that values anywhere between 1.040 and 1.243 are plausible values for the odds. Thus a plausible percentage change in the odds of matriculating for the population of students associated with a \$1,000 change in financial aid may range from 4.0% to 24.3%.

*Interpretation of the coefficient using the Delta-p value.*

We believe additional information can be gleaned from estimated logistic coefficients that will assist researchers and decision makers in judging the practical significance of the coefficient. The calculation and interpretation of the Delta- $p$  value for a given initial probability level and a range of Delta- $p$  values for a range of initial probability levels will provide decision makers with such information.

The Delta- $p$  value estimates the change in the probability of a person belonging to the group assigned a value of one in the dependent variable that is associated with a one-unit change in a given independent variable at a specified initial probability level. Petersen (1985) advocated setting this initial probability value equal to the mean of the dependent variable, which represents the proportion of the sample belonging to the group assigned a value of one in the dependent variable.

Three points should be noted with respect to the calculation of Delta- $p$  values. First, commenting on the use of the Delta- $p$  value in conjunction with a logistic regression coefficient, Petersen (1985) suggested it is a suitable method of estimating the change in the probability associated with the overall change in the independent variable when the initial probability is set equal to the mean of the dependent variable. Second, it is possible to calculate Delta- $p$  value with the initial probability equal to a value other than the mean of the dependent variable. The calculation of a set of Delta- $p$  values would allow the researchers to gauge whether the Delta- $p$  value calculated when the initial probability value set equal to the mean of the dependent variable adequately represents the change in the probability associated with the overall change in the independent variable. Third, since there are no known procedures to estimate the statistical significance of a Delta- $p$  value, St. John (1991) and Cabrera (1994) recommend that Delta- $p$  values be calculated and interpreted only for statistically significant logistics coefficients.

The calculation of the Delta- $p$  value is based on the following equation:

$$P_c = \left[ \frac{\exp\left(\ln\left(\frac{P_i}{1-P_i}\right) + b\right)}{1 + \exp\left(\ln\left(\frac{P_i}{1-P_i}\right) + b\right)} \right] - P_i$$

where:

1.  $P_c$  represents the change in the probability value.
2.  $P_i$  represents the initial probability value. This value is set equal to the proportion of the sample belonging to the group assigned the value of one in the dependent variable (i.e., the sample mean of the dependent variable).
3. The symbol  $\exp$  represents the base of the natural logarithm.
4. The symbol  $b$  represents the logistic regression coefficient for the given predictor variable.

The interpretation of the Delta-p statistic, which is calculated as the initial probability equal to the mean of the dependent variable, depends on whether the nature of the independent variable is continuous or categorical. If the independent variable consists of continuous values, the Delta-p statistic indicates the change in the initial probability of belonging to the group assigned a value of one in the dependent variable associated with a one-unit increase in the independent variable. If the independent variable represents categories, the Delta-p statistic indicates the change in the initial probability of belonging to the group assigned a value of one in the dependent variable if the person is a member of the group assigned a value of one rather than a member of the group assigned a value of zero in the independent variable.

To illustrate the calculation and interpretation of the Delta-p statistic, consider the results of the logistic regression model contained in Table 1. Since the proportion of students who matriculated, which is equal to the mean of the variable used to form the dependent variable, and the logistic coefficient for the continuous independent variable of university financial aid ( $X_4$ ) were .52 and .128, respectively, the Delta-p value is calculated as follows:

$$P_c = \left[ \frac{\exp\left(\ln\left(\frac{.52}{1-.52}\right) + .128\right)}{1 + \exp\left(\ln\left(\frac{.52}{1-.52}\right) + .128\right)} \right] - .52$$

$$P_c = .55 - .52$$

$$P_c = .03$$

The researchers could use this *Delta-p* value to indicate to decision makers that an increase in university financial aid of \$1,000, with the other variables held constant, is associated with an increase of .03 in the initial probability (.52) that a student will matriculate. Decision makers may find this *Delta-p* value useful in judging the practical importance of adjusting the financial aid for potential students in order to increase their likelihood of matriculating.

Two points should be noted, however, regarding the use of a *Delta-p* value. First, the *Delta-p* value is a point estimate. That is, unlike the odds value, confidence limits are not constructed for it. Second, before decision

makers judge the practical significance of changes in the probability of belonging to the group assigned a value of one, they may want to evaluate the changes associated with multiple unit changes in the independent variable (Norusis, 1999).

To illustrate this second point, consider a situation in which increases in financial aid may be of various dollar values, but increases of \$2,000, \$5,000, and \$8,000 are of specific interest to the university's admission staff. The staff may be more interested in these levels of change in financial aid because they are associated with commonly awarded scholarship levels. In this case, researchers may find it informative and helpful to report Delta-*p* values for changes in financial aid of \$2,000, \$5,000, and \$8,000 rather than just for a \$1,000 change.

*Delta-p values generated by a Microsoft® Excel program.*

As a means of encouraging the use of Delta-*p* values by researchers in their reports to policymakers, we developed a computer program that they can use to calculate Delta-*p* values. This program is designed to be used in conjunction with the Microsoft® Excel program.

Before the steps used to execute the program are discussed, it is important to note that it calculates a series of Delta-*p* values for various initial probability value rather than just one Delta-*p* value for one specified initial probability value. The reason why we designed the program to calculate a series of Delta-*p* values is based on the fact that in logistic regression the actual data are not linearly related to the independent variable even though the log-transformed data are linearly related to it. Thus the change in the probability of belonging to the group assigned a value of one in the dependent variable is not constant for various initial probability values. Since the Delta-*p* statistic is calculated for only one initial probability value, it does not allow a decision maker to assess the degree of variability of the changes in the initial probability associated with a specified change in the independent variable for a relevant range of initial probability values. That is, a decision maker may find that the changes in the initial probabilities are practically significant for those people with certain initial probabilities but not for those with other initial probability levels.

The computer program we designed calculates and lists the following: (a) a series of Delta-*p* values for a specified change in the independent variable that corresponds to a series of initial probability values, (b) the Delta-*p* statistic for the initial probability set at the mean of the dependent variable, and (c) the maximum absolute Delta-*p* value for a given change in the independent variable and the initial probability corresponding to this Delta-*p* value. To obtain these Delta-*p* values, six values must be entered into the Microsoft® Excel program on the first sheet labeled "Key Values" (see Figure 1)."

The Calculation of the Changes in Initial Probability		
Values for a Logistic Regression Coefficient		Description of Each Value
Change in the Independent Variable:	1	[the size of change in the independent variable]
Coefficient Value:	0.128	[the logistic regression coefficient for the variable]
Proportion of Cases with a Value of 1:	0.52	[average of the dependent variable]
Minimum Value:	0.25	[the lowest predicted probability for the sample data]
Maximum Value:	0.85	[the highest predicted probability for the sample data]
Interval:	0.05	[interval size between the min. and max. values]

*Figure 1: List of Key Values on Sheet 1 of the Microsoft® Excel Program*

The first value entered into the first sheet of the Microsoft® Excel program is the size of change in the independent variable. Since the change in financial aid was set at \$1,000 (one unit), the value of 1 was entered in the row entitled "Change in the Independent Variable." The second value entered on the first sheet of the Microsoft® Excel program is the logistic regression coefficient estimated for the independent variable. Thus the value of .128, which was the coefficient for the financial aid independent variable ( $X_4$ ), was entered in the row entitled "Proportion of Cases with a Value of 1."

The third and fourth values entered on the first sheet of the Microsoft® Excel program are the minimum and maximum initial probabilities in the series of initial probabilities used to generate a corresponding series of Delta-p values. We recommend that these minimum and maximum initial probabilities be set near the minimum and maximum predicted probabilities produced for the sample analyzed by the logistic regression model. Values outside of these minimum and maximum values, generally, will be of little interest to the researchers and decision makers. The minimum and maximum estimated probability values for the logistic regression model used in our illustration were .27 and .83, respectively. Thus the minimum value entered in the row entitled "Minimum Value" was .25; and the maximum value entered in the row entitled "Maximum Value" was .85. These values were selected for the following two reasons (a) the selected minimum value of .25 was below the predicted minimum probability of .27 while the selected maximum value of .85 was above the predicted maximum probability of .83, and (b) the difference between these two values can be divided into intervals of .05. The logic of this last reason is discussed presently.

Once the minimum and maximum initial probability values are determined, an interval value used to generate a series of initial probability values between the minimum (.25) and maximum (.85) initial probabilities must be specified. We suggest that the use of a .05 interval generally will provide enough information to assess the practical significance of the coefficient. Although the program is designed to handle intervals as small as .01, we believe one should generally use an interval of .05 or .10. The number of probability values generated by intervals smaller than .05 can be overwhelming, while the number of probability values generated by intervals larger than .10 may provide too little information. The interval value of .05 was entered in the row on the first sheet of the Microsoft® Excel program entitled "Interval."

Once these five values are entered into the first sheet of the Microsoft<sup>®</sup> Excel program, the following information is listed on the second sheet labeled "Calculations" (see Figure 2):

1. A series of initial probability values is listed in the column entitled "Initial p values."
2. A series of final probability values once the independent variable had been changed by the amount specified on the first sheet is listed under the column entitled "Final p values."
3. A series of differences between the final probability values and the corresponding initial probability values is listed under the column entitled "Changes in p values." Note that the program provides a graph of these differences between the initial and final probabilities. This graph is located on the third sheet of the Microsoft<sup>®</sup> Excel program (see Figure 3). This third sheet is labeled "Graph."
4. The initial probability set at the mean of the dependent variable and its corresponding  $\Delta$ -p statistic are listed in the row entitled "Delta-p Statistic for."
5. The initial probability value corresponding to the maximum absolute  $\Delta$ -p value is listed in the row entitled "The p value at which the absolute change in p is the largest for the specified change in the independent variable."
6. The actual  $\Delta$ -p value for the maximum absolute  $\Delta$ -p value, which corresponds to the initial probability discussed in Point 5, is listed in the row entitled "Largest absolute change in p expressed as p."
- 7.

Initial p values	Final p values	Change in p values			
0.25	0.275	0.025			
0.30	0.328	0.028			
0.35	0.380	0.030	Delta-p Statistic for	0.52	0.032
0.40	0.431	0.031			
0.45	0.482	0.032			
0.50	0.532	0.032	The p value at which the absolute change in p		
0.55	0.581	0.031	is the largest for the specified change in the		
0.60	0.630	0.030	independent variable		0.484
0.65	0.679	0.029			
0.70	0.726	0.026	Largest absolute change in p expressed as p		0.032
0.75	0.773	0.023			
0.80	0.820	0.020			
0.85	0.866	0.016			

Figure 2. List of  $\Delta$ -p and Corresponding Initial Probability Values on Sheet 2 of the Microsoft<sup>®</sup> Excel Program

### Change in Initial Probability Values

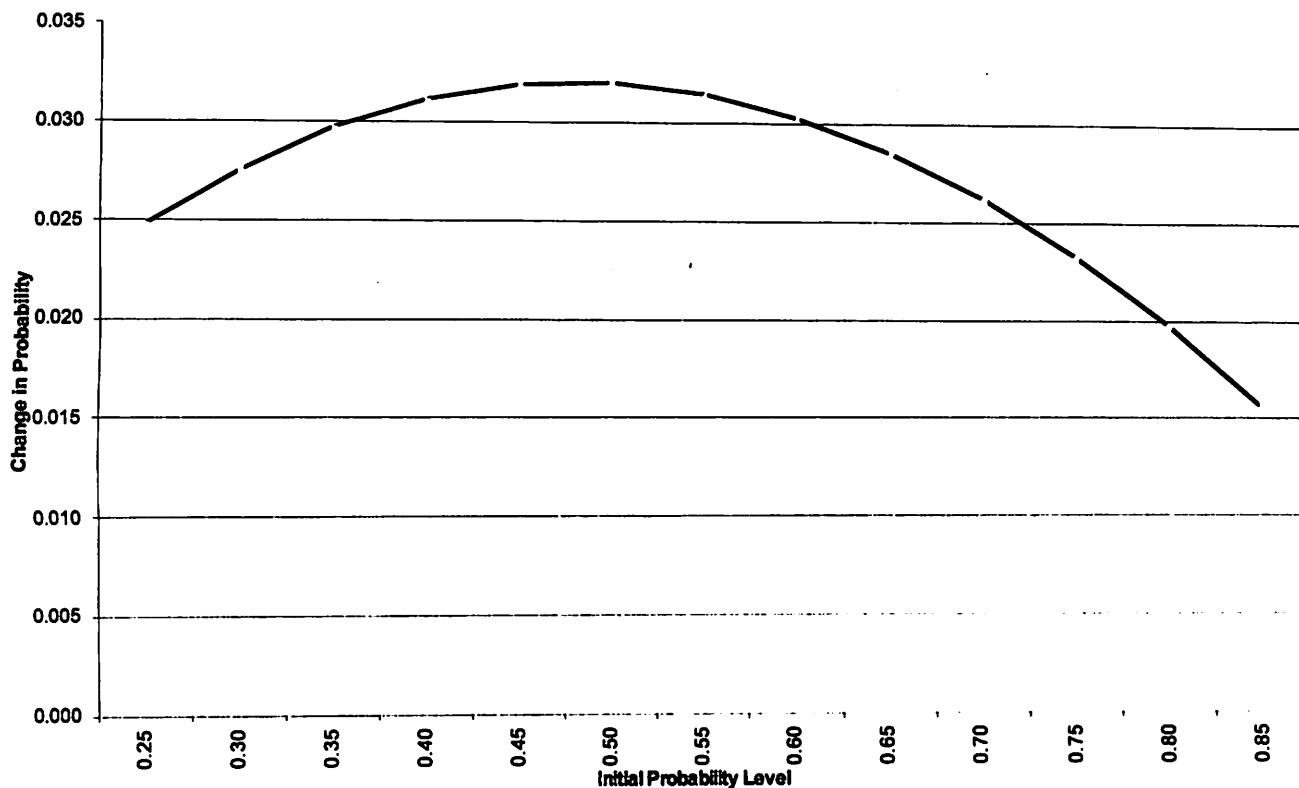


Figure 3: Chart of the Delta-*p* and Corresponding Initial Probability Values Contained on Sheet 3 of the Microsoft® Excel Program.

A review of the various Delta-*p* values contained on Figure 2 reveals the Delta-*p* value calculated for the initial probability level set equal to the mean of the dependent variable of .52 was .032. An examination of the range in Delta-*p* values calculated for the various initial probability levels allows the decision makers to determine whether the changes in the probability that the event will occur is practically significant across the entire range of initial probabilities or only for a specific subgroup of initial probability values.

The series of Delta-*p* values range from a high of .032 for the initial probability value of .484 to a minimum of .016 for the initial probability of .85. A review of this series of Delta-*p* values by the researchers may lead them to conclude that except for the more extreme initial probability levels, such as .80 and .85, the Delta-*p* value calculated with the initial probability value set equal to the mean of the dependent variable provides a fairly good representation of the overall change in the initial probability for a given change in the independent variable. Thus, in this case, the researchers and decision makers could use this one Delta-*p* value (.032) to assess the practical importance of a change in financial aid of \$1,000 with respect to the corresponding change in the students'

likelihoods they would matriculate. If the series of Delta-*p* values deviate by a substantial amount, as judged by the researchers, they may want to report more than one Delta-*p* value to the decision makers.

Two points should be noted with respect to the calculation of Delta-*p* values. First, a series of Delta-*p* values can be calculated for each statistically significant logistic coefficient. Second, a series of Delta-*p* values can be calculated by the computer program for changes in the independent variable larger than one unit. As previously stated, if researchers and policymakers were specifically interested in financial aid changes of \$2,000, \$5,000, and \$8,000 due to certain types of scholarships, our computer program can be used to generate a series of Delta-*p* values for each of these three changes. To obtain these three series of Delta-*p* values, the researcher would specify the change in the financial aid independent variable as (a) a 2 for a \$2,000 change, (b) a 5 for a \$5,000 change, and (c) an 8 for an \$8,000 change. The series of Delta-*p* values for these three changes are listed in Table 2. Once again, if the series of Delta-*p* values deviate by a substantial amount, as judged by the researchers, they may want to report more than one Delta-*p* value to the decision makers.

Table 2

*Series of Delta-*p* values Corresponding to Three Levels of Increases in Financial Aid*

Initial Probability Level <sup>a</sup>	Change in Financial Aid		
	\$2,000	\$5,000	\$8,000
.25	.051	.137	.231
.30	.056	.148	.244
.35	.060	.155	.250
.40	.063	.158	.250
.45	.064	.158	.245
.50	.064	.155	.236
.55	.062	.149	.223
.60	.060	.140	.207
.65	.056	.129	.188
.70	.051	.116	.167
.75	.045	.101	.143
.80	.020	.020	.020
.85	.016	.016	.016

<sup>a</sup> The Delta-*p* values for the initial probability of .52 (mean of the dependent variable) are .063, .153, and .231 for changes in financial aid of \$2,000, \$5,000, and \$8,000, respectively.



## Summary

We believe the values generated by the method and the computer program described in this paper will provide researchers and decision makers with information that will assist them in judging the practical significance of a logistic regression coefficient in a number of ways. First, decision makers may find the changes in probability values easier to use when judging the practical significance of the coefficient as compared to the log odds ratio value (i.e., the logistic regression coefficient) or even the odds value ( $\exp(b)$ ). Second, the series of changes in the initial probability values associated with a specified change in the independent variable will allow decision makers to judge whether the Delta- $p$  value provides enough information to judge the practical significance of the coefficient. Third, if a decision maker believes the practical significance of the coefficient should be judged in terms of changes in the independent variable that are greater than one unit, the program can easily generate the essential values that reflect such a change.

Although we have presented a computer program that allows researchers to readily compute Delta- $p$  values for given changes in a specified independent variable, not all of the issues related to Delta- $p$  values have been addressed. One such issue is: What is the best way to express logistic coefficients to assist researchers and policy makers in assessing practical significance when interaction effects are statistically significant? Concepts and procedures remain to be developed with respect to this issue.

The need to address the practical significance of statistically significant findings in a logistic regression analysis is an important task for researchers and decision makers to undertake. In spite of the fact that additional issues need to be addressed with respect to the use of Delta- $p$  values, we believe the implementation of the method presented in this article will begin to assist researchers and decision makers in the task of assessing the practical significance of logistic regression coefficients.

## References

- Cabrera, A. F. (1994). *Logistic regression analysis in higher education: An applied perspective*. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 225-254). New York: Agathon Press.
- Cizek, G. H., & Fitzgerald, S. M. (1999). An introduction to logistic regression. *Measurement and Evaluation in Counseling and Development*, 31(4), 223-244.
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Collett, D. (1991). *Modeling binary data*. New York: Chapman Hall.
- Fisher, R. A. (1925). *Statistical methods for research workers*. London: Oliver & Boyd.
- 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.

- Fraas, J. W., & Newman, I. (2003, February). *Ordinary least squares regression, discriminant analysis, and logistic regression: Questions researchers and practitioners should address when selecting an analytic technique*. Paper presented at the annual meeting of the Eastern Education Research Association, Hilton Head Island, SC.
- Hair, H. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Hendel, D. D. (2001, April). *The relative contribution of participating in a first-year seminar on student satisfaction and retention into the sophomore year*. Paper presented at the annual meeting of the American Educational Association, Seattle, WA.
- Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons.
- Levin, J. R., & Robinson, D. H. (2000). Rejoinder: Statistical hypothesis testing, effect-size estimate, and the conclusion coherence of primary research studies. *Educational Researcher*, 29(1), 34-36.
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.
- McCoach, D. B., & Siegle, D. (2001, April). *Why try? Factors that differentiate underachieving gifted students from high achieving gifted students*. Paper presented at the annual meeting of the American Educational Association, Seattle, WA.
- Norušis, M. J. (1999). *SPSS regression models™ 10.0* [computer software]. Chicago, IL: SPSS, Inc.
- Petersen, T. (1985). A comment on presenting results from logit and probit models. *American Sociological Review*, 50(1), 130-131.
- Robinson, D. H., & Levin, J. R. (1997). Reflections on statistical and substantive significance, with a slice of replication. *Educational Researcher*, 26(5), 21-26.
- Schreiber, J. B. (2002). Scoring above the international average: A logistic regression model of the TIMSS advanced mathematics exam. *Multiple Linear Regression Viewpoints*, 28(1), 22-30.
- St. John, E. P. (1991). What really influences minority attendance? Sequential analyses of the high school and beyond sophomore cohort. *Research in Higher Education*, 32(2), 141-158.
- Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, 25(2), 26-30.

#### Endnote

<sup>1</sup> A copy of the program can be requested from the authors of this article by sending such a request to [jfraas@ashland.edu](mailto:jfraas@ashland.edu)

*John W. Fraas is a Professor of Business Administration at Ashland University and J. Michael Drushal, is Chairperson of the, Department of Business Administration at Ashland University. Correspondence concerning this article should be addressed to Dr. John W. Fraas, School of Business Administration, Ashland University, 401 College Avenue, Ashland, OH 44805.*

## LEARNING STYLES AND MINDSTYLES: AN EXAMINATION OF VALIDITY

Maria Martinez Witte  
Assistant Professor, Adult Education  
Auburn University

*This study sought to examine the discriminant validity of the Multi-Modal Paired Associates Learning Test III (MMPALT III) and the Gregorc Style Delineator. The MMPALT III assesses attributes from the sensory modality domain and the Gregorc Style Delineator assesses attributes from the cognitive domain. The MMPALT III is a performance-based learning style instrument identifying an individual's ability to acquire information through print, aural, interactive, visual, haptic, kinesthetic, and olfactory mediums. The Gregorc Style Delineator is a self-assessment instrument identifying an individual's dominant, intermediate, or low mediation abilities classified as: Concrete Sequential, Abstract Sequential, Abstract Random, and Concrete Random. Results of this study concluded that the MMPALT III measured constructs separate from the Gregorc Style Delineator. These instruments were designed using different conceptual structures and demonstrate discriminant validity. Empirical results indicate the instruments measure separate although minimally related constructs.*

### Introduction

The development and measurement of learning styles has been hampered by the lack of appropriately developed instruments (James & Blank, 1993). Even though there is general agreement on the uniqueness of learning styles, it is doubtful that learning style preferences can be measured reliably using currently available learning style instruments (Terry, 2002). Lemire (2002) also identified a concern as to weaknesses in validity and reliability in the use of learning styles instruments. Instrument validity refers to the extent to which the instrument measures what it purports to measure or how accurate it is and reliability refers to the extent to which results are consistently obtained (Hittleman & Simon, 2002). Establishing valid and reliable learning styles instruments begins with an understanding of validity. Carmines and Zeller (1979) stated, "an indicator must be more than reliable if it is to provide an accurate presentation of some abstract concept. It must also be valid" (p. 12).

#### *Purpose*

This study investigated the relationship of two learning styles instruments' constructs, the Multi-Modal Paired Associates Learning Test III (MMPALT III) and the Gregorc Style Delineator, through the examination of subtest scores. The MMPALT III measures attributes from the sensory modality domain. The Gregorc Style Delineator measures attributes from the cognitive domain. Although efforts to establish construct validity of the Multi-Modal Paired Associates Learning Test III (MMPALT III) have yet to be completed, Harrison (1997), Reno (1997), and Witte (1999) addressed concurrent validity in their research. By establishing relationships between the MMPALT III and the Gregorc Style Delineator, this study contributes further to establishing an estimate of validity for the MMPALT III.

#### *Types of Validity*

There are primarily four areas used to investigate an instrument's validity. These areas consist of collecting

evidence of (1) content, (2) face, (3) criterion (which includes concurrent and predictive), and (4) construct validity. Content validity is provided “when an instrument’s creators demonstrate that the specific items or questions represent an accurate sampling of specific bodies of knowledge (i.e., curricula or courses of study)” (Hittleman & Simon, 2002, p. 112). Face validity refers to the extent in which an instrument appears, through that which is visible and observable, to measure specific information (Hittleman & Simon, 2002). Criterion-related validity refers to the degree to which instruments provide accurate measurements when compared to scores on a relevant criterion variable. There are two kinds of criterion-related validity: concurrent and predictive (Huck, 2000). Concurrent validity is determined by “relating the test scores of a group of subjects to a criterion measure administered at the same time or within a short interval of time” (Borg & Gall, 1989, p. 254). The extent to which the results correlate is an indication of concurrent validity (Hittleman & Simon, 2002). Predictive validity refers to ability of the instrument to predict a target population’s performance after some future event (Hittleman & Simon, 2002).

Construct validity is the extent to which a particular instrument can be shown to measure a hypothetical construct (Gall, Borg, & Gall, 1996). In human behavior hypothetical construct examples are intelligence, anxiety, or creativity. These would be considered hypothetical constructs because they are not directly observable but rather are inferred on the basis of their observable effects on behavior (Borg & Gall, 1989). Pedhazur and Schmelkin (1991) noted “construct validation is concerned with validity of inferences about the unobserved variables (the constructs) on the basis of observed variables (their presumed indicators)” (p. 52). They also observed that “construct validation is a never ending enterprise” (p. 59). Carmines and Zeller (1979) identified construct validation as involving three distinct steps:

1. The theoretical relationship between the concepts themselves must be specified.
2. The empirical relationship between the measures of the concepts must be examined.
3. Finally, the empirical evidence must be interpreted in terms of how it clarifies the construct validity of a particular issue.(p. 23)

Campbell and Fiske (1959) identified four criteria for assessing convergent and divergent validity. They refer to coefficients being statistically significant and sufficiently different; coefficients being higher than any other correlations; coefficients being higher than correlations between that factor and any other factor evaluated by the same group of raters; and the establishment of similar correlational patterns among all rating factors. Huck (2000) also proposed methods that can be used to establish a degree of construct validity. Two of these methods were used in this study to assist in establishing a degree of construct validity on the Multi-Modal Paired Associates Learning Test (MMPALT III). Huck’s (2000) construct validity criteria were as follows:

- (1) provide correlational evidence showing that the construct has a strong relationship with certain measured variables and a weak relationship with other variables, with the strong and weak relationships conceptually tied to the new instrument’s construct in a logical manner; (2) show that certain groups obtain higher mean scores on the new instrument than other groups, with the high- and low-scoring groups being determined on logical grounds prior to the administration of the new instrument; or (3) conduct a factor analysis on scores from the new instrument. (p. 104)

Confirmatory factor analysis can begin with an a priori hypothesized model and deductively ascertain its feasibility

by offering a more definitive empirical evidence of the underlying factor structure of a scale than is afforded by exploratory factor analysis (Tabachnick & Fidell, 2001). If correlational evidence supports the researcher's predictions (e.g., to be low or high), they are said to establish convergent and discriminant validity (Huck, 2000).

#### Method

##### *Sample*

Statistical power and effect size were primary considerations in selecting an appropriate sample size. A sample size of 80 participants was selected to fill a quota sample and resulted in an expected medium effect size of  $r = .30$ ,  $p < .01$  (Cohen, 1988).

The sample size of 80 was supported by previous research (Harrison, 1997; Reno, 1997). The sample was stratified by age, educational attainment level, and gender variables that were consistent with previous MMPALT research and to assist in future confirmatory factor analysis efforts. Age characteristics consisted of two groups: younger participants 18 to 30 years of age and older participants 50 years of age and older. Two educational levels were used: basic, indicating the participant had a high school diploma or GED and no baccalaureate, and advanced, indicating the participant had a high school diploma or GED and at least a bachelor's degree (see Table 1).

*Table 1*

*Distribution of Participants by Age and Educational Level*

Age 18-30 Years			
Basic Education		Advanced Education	
Male Age	Female Age	Male Age	Female Age
18	18	22	22
18	18	22	23
19	19	24	24
20	19	25	25
21	20	25	26
23	21	27	26
25	21	28	27
26	22	29	28
28	23	30	28
29	23	30	29
Age 50 + Years			
Basic Education		Advanced Education	
Male Age	Female Age	Male Age	Female Age
50	50	50	50
51	55	51	52
52	56	52	52
54	56	55	53
55	58	56	54
58	58	56	56
61	59	57	59
63	60	58	59
67	61	63	61
68	63	64	64

N = 80

*Procedure*

This study examined the relationship among seven subtests of the MMPALT III identified by Cherry (1981) and four mediation abilities of the Gregorc Style Delineator (Gregorc, 1985). The four mediation abilities are: Concrete Sequential (CS), Abstract Sequential (AS), Abstract Random (AR), and Concrete Random (CR) (Gregorc, 1985).

Following the Gregorc Style Delineator, administration of the MMPALT III subtests (Institute for Learning Styles Research, 1996) began with the group-administered print, aural, and visual subtests. The order of the remaining subtests was interactive, kinesthetic, haptic, and olfactory.

**Results**

The results of this study were analyzed using two of Huck's (2000) criterion to further establish an estimate of validity and to contribute to establishing relationships between the MMPALT III and Gregorc Style Delineator

instruments. Huck's (2000) first criterion for establishing an estimate of construct validity requires providing correlational evidence demonstrating that a strong relationship exists with certain measured variables and a weak relationship exists with other variables. Results from this study were analyzed by reviewing instrument reliability.

The relationship between the instruments was analyzed using correlation and bivariate correlation analysis. Bivariate correlation coefficients were computed between the seven subscales of the MMPALT and the four factors of the Gregorc Style Delineator (see Table 2). The p-value of less than .002 ( $.05/28 = .0018$ ) was required for significance using the Bonferroni approach to control for Type I error across the 28 correlations. None of the 28 correlations was statistically significant with the coefficients ranging from .024 to .280 (weak coefficients). In general, the results suggest that the dimensions of the MMPALT III and the Gregorc Style Delineator demonstrate discriminant validity.

Table 2

*Bivariate Correlations Among the Subset Subscales*

	Print	Visual	Aural	Inter	Kine	Haptic	Olf
Bivariate correlations							
CS	-.102	-.173	.020	.002	-.169	-.240	-.224
AS	-.259	-.176	-.045	-.030	-.109	-.131	-.168
AR	.280	.208	.134	.078	.197	.206	.212
CR	.048	.082	-.173	-.028	.137	.167	.167

$p < .002$  bivariate correlations;  $N = 80$

The internal consistency of the MMPALT III subtests was estimated using a Cronbach's Alpha technique. Table 3 contains the Cronbach's Coefficient values by subtest. Subtest coefficient alphas ranged from a high of .744 for Print, to a low of .554 for Olfactory. These data support earlier findings of Reno (1997) who reported subtest alpha values for Print (.737), Aural (.761), Interactive (.795), Visual (.790), Haptic (.774), Kinesthetic (.544), and Olfactory (.390).

Table 3

*Cronbach's Coefficient Alpha for MMPALT III Subtests*

Subtest*	M	SD	
Print	4.11	2.74	.77
Aural	4.13	2.46	.67
Interactive	5.48	2.62	.71
Visual	5.85	2.66	.73
Haptic	6.44	2.50	.70
Kinesthetic	3.88	2.31	.64
Olfactory	0.50	1.29	.55

\*each subtest contains 10 paired items;  $N = 80$

Huck's (2000) second criterion for establishing an estimate of construct validity requires demonstrating that certain groups obtain higher mean scores on the new instrument than other groups, with the high- and low-scoring groups being determined on logical grounds prior to the administration of the instrument. This criterion was not investigated in this study.

Huck's (2000) third criterion for establishing an estimate of construct validity requires conducting a factor analysis on instrument scores. The confirmatory factor analysis begins with an a priori hypothesized model and deductively ascertains its feasibility by offering a more definitive empirical evidence of the underlying factor structure of a scale than is afforded by exploratory factor analysis (Tabachnick & Fidell, 2001).

A low correlation coefficient was hypothesized to demonstrate discriminant validity between the constructs. The results indicated an empirically supported likelihood that these two instruments are measuring separate although minimally related constructs.

Because missing data appeared to be randomly scattered among the variables, a full information maximum likelihood (FIML) imputation was performed to estimate missing data. The factor structures were also examined using a confirmatory factor analysis. A series of models was tested in the following order: (a) a single-factor *g* model in which all items were free to load on only one common factor; (b) a correlated two-factor model in which each factor was correlated to the other; and (c) an orthogonal two-factor model in which each factor was set to be independent of each other. The first two models were included to aid in the assessment of the orthogonal two-factor model.

The models were examined by AMOS version (4.0) maximum likelihood factor analysis (Arbuckle, 1999). The models were evaluated by a variety of fit measures that are classified as absolute, relative, parsimonious, and population discrepancy. Absolute fit measures assess how well the proposed interrelationships among the variables match the interrelationships among the actual interrelationships. The measure of absolute fit used in this study was the chi-square test because AMOS does not provide other absolute measures when missing data are estimated with the FIML imputation procedure. Measures of relative fit compare the hypothesized model to the null model. The relative fit measures employed in this study were the Comparative Fit Index (CFI) (Bentler, 1990), and the Tucker-Lewis Index (TLI) (Bentler & Bonett, 1980).

Measures of parsimonious fit were employed to determine if the overall fit of the model has been accomplished by overfitting the data. The parsimonious fit measure in this study was the chi-square divided by the degrees of freedom. Lastly, population discrepancy measures are estimates from the sample coefficients to the population coefficients. The population discrepancy measure in this study was the Root Mean Square Error of Approximation (RMSEA) (Browne & Cudeck, 1993). Models were compared by examining differences in values of chi-square to identify statistically significant variations among the models. The fit indices for the three models are presented in Table 4.



Table 4

*Chi-square and Goodness of Fit Indices for Confirmatory Factor Models*

Factor Model	X <sup>2</sup>	df	X <sup>2</sup> /df	CFI	TLI	RMSEA
Single (g)	244.50* 44	5.57	.904	.856	.240	
Correlated	158.41* 43	3.68	.945	.915	.184	
Orthogonal	160.26* 44	3.64	.944	.916	.183	

\* p < .05; N = 80

The single-factor g model in which all items were free to load on only one common factor was the poorest fitting model and was thus eliminated from further consideration. The chi-square test for differences failed to reveal any statistically significant differences between the correlated two-factor model and the orthogonal two-factor model. The factor loadings for both the correlated and orthogonal models of the MMPALT III and Gregorc Style Delineator are provided in Table 5 and Figure 1. All items loaded significantly on their respective factors, p < .05.

Because of the low correlation coefficient between the two dimensions of the MMPALT III and the Gregorc Style Delineator (r = -.188, p < .05) the orthogonal two-factor model was deemed superior to the correlated model. The orthogonal model yielded acceptably high goodness of fit indices for both the CFI (.944) and the TLI (.916) (Bentler, 1990). The RMSEA achieved an index of .183 indicating a poor fit of the model in relation to the degrees of freedom. These results strongly suggest that the MMPALT III and the Gregorc Style Delineator demonstrated discriminant validity.

Table 5

*Factor Loadings of the Correlated & Orthogonal Models of the MMPALT III and Gregorc Style Delineator*

Item	MMPALT III (Correlated)	MMPALT III (Orthogonal)	Gregorc Style Delineator (Correlated)	Gregorc Style Delineator (Orthogonal)
Print	0.743	0.740		
Visual	0.678	0.674		
Aural	0.714	0.726		
Interactive	0.630	0.636		
Kinesthetic	0.483	0.478		
Haptic	0.527	0.519		
Olfactory	0.401	0.393		
Concrete Sequential			0.852	0.851
Abstract Sequential			0.575	0.570
Abstract Random			-0.570	-0.554
Concrete Random			-0.803	-0.814

Note: All items significantly loaded on their respective factors (p < .05); N = 80

### Conclusions and Recommendations

This study provided evidence that the MMPALT III, as a measure of learning styles, and the Gregorc Style Delineator, as a measure of cognitive mindstyles, accomplish the purpose for which each was designed. Therefore, the instruments can be used to examine independent

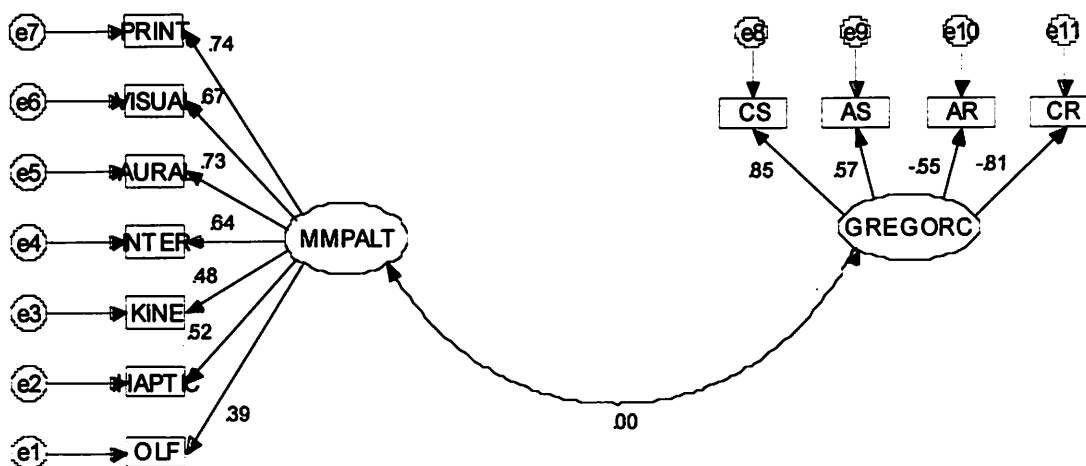


Figure 1. Correlated & Orthogonal Models of the MMPALT III and the Gregorc Style Delineator

### Conclusions and Recommendations

This study provided evidence that the MMPALT III, as a measure of learning styles, and the Gregorc Style Delineator, as a measure of cognitive mindstyles, accomplish the purpose for which each was designed. Therefore, the instruments can be used to examine independent constructs. The MMPALT III and the Gregorc Style Delineator are not interchangeable instruments measuring constructs from the same domain. These instruments were designed using different conceptual structures and demonstrate discriminant validity.

Based on the results of this study, it is recommended that other cognitive instruments (e.g., analogical reasoning ability, inquiry mode analysis, thinking styles) be used in conjunction with the MMPALT III to investigate the relationships of constructs and cognitive processes as evaluated by that specific instrument. Further validity investigations should begin the research study with the construct criterion as part of the overall design. Cronbach and Meehl (1955) identified goals to assist the researcher in design clarity. These included, “(a) what interpretation is proposed, (b) how adequately the writer believes this interpretation is substantiated, and (c) what

evidence and reasoning lead him to this belief' (p. 297).

## References

- Arbuckle, J. L. (1999). *Amos 4.0 user's guide* [Computer software]. Chicago: Small Waters Corporation.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588-606.
- Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (5<sup>th</sup> ed.). New York: Longman.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.
- Carmines, E., & Zeller, R. (1979). *Reliability and validity assessment*. London: Sage.
- Cherry, C. E., Jr. (1981). The measurement of adult learning styles: Perceptual modality (Doctoral dissertation, University of Tennessee, Knoxville, 1981). *Dissertation Abstracts International*, *42*(09A), 3852.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (3<sup>rd</sup> ed.). New York: Academic Press.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.
- Gall, M., Borg, W., & Gall, J. (1996). *Educational research: An introduction* (6<sup>th</sup> ed.). New York: Longman.
- Gregorc, A. F. (1985). *Gregorc style delineator*. Maynard, MA: Gabriel Systems.
- Harrison, W. D. (1997). Adult student learning style and academic major choice: A technical institute profile (Doctoral dissertation, University of Alabama, Tuscaloosa, 1997). *Dissertation Abstracts International*, *58*(06A), 2026.
- Hittleman, D., & Simon, A. (2002). *Interpreting educational research: An introduction for consumers of research* (3<sup>rd</sup> ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Huck, S. (2000). *Reading statistics and research* (3<sup>rd</sup> ed.). New York: Longman.
- Institute for Learning Styles Research. (1996, August). *Multi-modal Paired Associates Learning Test III procedures*. Knoxville, TN: Author.
- James, W., & Blank, W. (1993). Review and critique of available learning-style instruments for adults. In D. Flannery (Ed.), *Applying cognitive learning theory to adult learning: New directions for adult and continuing education* (No. 59, pp. 47-57). San Francisco: Jossey-Bass.
- Lemire, D. (2002). What developmental educators should know about learning styles and cognitive styles. *Journal of College Reading and Learning*, *32*(2), 177-182.

- Pedhazur, E. J., & Schmelkin L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Reno, H. (1997). A comparison of the perceptual learning styles of native Hispanic adults and European-American adults as measured by the Multi-Modal Paired Associates Learning Test (Doctoral dissertation, University of South Florida, Tampa, 1997). *Dissertation Abstracts International*, 58(11A), 4159.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4<sup>th</sup> ed.). Needham Heights, MA: Allyn & Bacon.
- Terry, M. (2002). Translating learning style theory into developmental educational practice: An article based on Gregorc's cognitive learning styles. *Journal of College Reading and Learning*, 32(2), 154-177.
- Witte, M. M. (1999). A validation study of the relationship of learning styles and mindstyles as indicated by the Multi-Modal Paired Associates Learning Test III and the Gregorc Style Delineator (Doctoral dissertation, University of South Florida, Tampa, 1999). *Dissertation Abstracts International*, 60(08), 2770A.

*Maria Martinez Witte is an Assistant Professor of Adult Education at Auburn University in Auburn, AL. Correspondence concerning this article should be addressed to Dr. Maria Martinez Witte, Educational Foundations, Leadership and Technology, 4036 Haley Center, Auburn University, AL 36849-5221.*

## DATA ANALYSIS TECHNIQUES IN QUALITATIVE RESEARCH

Barbara B. Kawulich  
State University of West Georgia

*The explanation of how one carries out the data analysis process in qualitative research is an area that is sadly neglected by many researchers. This paper presents a variety of data analysis techniques described by various qualitative researchers, including LeCompte and Schensul, Wolcott, and Miles and Huberman. It further shares several activities used to teach qualitative data analysis techniques to students of qualitative research methods.*

### Introduction

Since the 1976 Sieber study, which found that less than 10% of the pages of frequently used qualitative textbooks discussed the topic of data analysis, more recent texts have taken care to elaborate on the methods used to analyze data (Huberman & Miles, 1998). It is important for qualitative researchers to have the opportunity to share the techniques they use to analyze data to add more options to our repertoire of techniques or to refresh our skills in conducting research. The purpose of this article is to discuss a variety of ways that qualitative researchers analyze data and to share techniques for teaching data analysis to students enrolled in qualitative research classes.

There are many different techniques for analyzing qualitative data. Novice researchers may feel overwhelmed by the variety; however, it may be helpful for them to understand that there is no prescribed way to address the process. The ways that they choose to analyze data should stem from a combination of factors, which include the research questions being asked, the theoretical foundation of the study, and the appropriateness of the technique for making sense of the data. Analyzing qualitative data typically involves immersing oneself in the data to become familiar with it, then looking for patterns and themes, searching for various relationships between data that help the researchers to understand what they have, then visually displaying the information and writing it up. There are more techniques for analyzing data than are presented in this article. Further, this article does not address the software available for assisting the researcher in the data management and analysis processes. Today's computer age has provided a variety of software to help qualitative researchers. The software has enhanced the qualitative researcher's ability to distill massive amounts of data in a much shorter time than was possible without its use. Software provides a useful tool for organizing data, in particular. There are a variety of software programs available which augment the coding and analysis of data from text, audio, and video sources. Whether one chooses to use the assistance of computer software, the same processes for analyzing data should occur. Using a variety of techniques enables the researcher to make appropriate connections that explain the full story being told.

The citations used in this article are only a few of the available sources that provide a discussion of analysis techniques. They were selected, because they are a few of the ones I use in teaching data analysis. Student feedback has influenced me to continue to use these sources as teaching tools, as the

sources cited herein are well presented, are easily read and understood by students, and, to some degree, provide an overall view of qualitative data analysis.

### **What is Data Analysis?**

LeCompte and Schensul (1999) define analysis as the process a researcher uses to reduce data to a story and its interpretation. Data analysis is the process of reducing large amounts of collected data to make sense of them. Patton (1987) indicates that three things occur during analysis: data are organized, data are reduced through summarization and categorization, and patterns and themes in the data are identified and linked. LeCompte and Schensul (1999) suggest that data analysis be done as data are collected in the field, as soon as possible after the data have been collected, both while the researcher is still in the field, and later, when the researcher is no longer in the field. They describe in-the-field analysis as including inscription, description, and transcription. They also suggest that analysis may be conducted in both a top down fashion and a bottom up fashion.

### **Approaches to Analysis**

Merriam (1998) discussed several approaches to data analysis, including ethnographic analysis, narrative analysis, phenomenological analysis, and constant comparative method. Ethnographic analysis involves identifying categories related to a culture's economy, demographics, human life, particularly family, education, and health care issues, and the environment. She describes narrative analysis as it is used in several fields of study. For example, sociological/sociolinguistic models relate narrative to the social context. Psychological approaches focus on memorization in storytelling, with particular emphasis on understanding, recall, and summarization. Anthropological models emphasize how stories vary across cultures, looking at customs, beliefs, values, and social context of narratives. Literary models focus on grammar, syntax, and plot of narratives; ideological perspectives, like feminist theory, critical theory, and postmodernism, may be used to analyze and interpret narratives. Phenomenological analysis includes an epochal approach, which involves laying out one's assumptions about the phenomenon under study, bracketing, imaginative variation (looking at the phenomenon in various ways), and first and second order knowledge. Constant comparative method assigns codes that reflect the conceptual relationships (Merriam, 1998).

Bernard (2000) also suggests several approaches to data analysis, including hermeneutics or interpretive analysis, narrative and performance analysis, discourse analysis, grounded theory analysis, content analysis, and cross-cultural analysis. In hermeneutics or interpretive analysis, the researcher "continually interpret(s) the words of those texts to understand their meaning and their directives" (p. 439). The focus of narrative and performance analysis is to discover repeated similarities in people's stories. Discourse analysis involves looking closely at how people interact with each other. Bernard describes grounded theory as "a set of techniques for (1) identifying categories and concepts that emerge from text, and (2) linking the concepts into substantive and formal theories" (p. 443). In content analysis, Bernard presents a more quantitative approach to data analysis by creating codes, applying the codes to texts, testing the

intercoder reliability when more than one coder is used, creating a matrix or table of units of analysis by variable, and conducting statistical analysis of the matrix. Cross-cultural analysis emphasizes the application of codes found in the Human Relations Area Files (HRAF) at Yale University. The approaches to analysis presented by Merriam (1998) and Bernard (2000) are illustrated in Table 1.

*Table 1*

*Approaches to Analysis Presented by Merriam (1998) and Bernard (2000)*

Author	Approaches
Merriam (1998)	Ethnographic Analysis Narrative Analysis Phenomenological Analysis Constant Comparative Analysis
Bernard (2000)	Hermeneutic/Interpretive Analysis Narrative/Performance Analysis Discourse Analysis Grounded Theory Analysis Content Analysis Cross-cultural Analysis

### **Coding and Category Guidelines**

While the emergent design of qualitative research lends itself to analyzing data as it is collected in the field, there are certain tasks that must be accomplished before in-depth analysis can occur. Coding of transcribed data from interviews and field notes, for example, is an initial step in this process. Merriam (1998) describes the process of data analysis as being a complex action of moving back and forth between data and concepts, between description and interpretation, using both inductive and deductive reasoning. Merriam also describes Glaser and Strauss' (1967) grounded theory approach, constant comparative method, in which categories, properties, and hypotheses are used to provide conceptual links between and among categories and properties. Constant comparative method is used to analyze data by assigning codes that reflect various categories and properties to units of data through sorting them into groups of like substance or meaning. Categories should:

- reflect the purpose of the research,
- be exhaustive,
- be mutually exclusive,
- be sensitive to category content, and
- be conceptually congruent (Merriam, 1998).

The number of categories used should be guided by the frequency of mention by participants, the audience for whom the study is conducted, and the uniqueness of the category (Guba & Lincoln, 1981).

The coding process generates categories that must be fleshed out by seeking relevant data bits that inform the category. Boyatzis (1998) stated that a good code has these five elements: (1) a label (i.e., a name); (2) a definition of what the theme concerns (that which characterizes the theme); (3) a description of how to know when the theme occurs (those aspects that let you know to code a unit for that theme); (4) a

description of any qualifications or exclusions to the identification of the theme; and (5) a listing of examples, positive and negative, to eliminate confusion. The label should be developed last and should be conceptually meaningful, clear and concise, and close to the data.

There are several approaches to the coding process. For example, theory driven coding begins with the researcher's theory of what occurs and the formulation of the indicators of evidence that would support the theory. The elements of the code are derived from the hypothesis or the elements of the theory. Prior research also is used to develop coding schemes. Research driven codes are those used by others, and their findings provide the most direct help in developing codes from prior research. Using someone else's codes may require that one look at inter-rater reliability. A third approach to coding, data driven coding, involves inductive code development based on the data collected in the study on which one is working.

Data are analyzed on an ongoing basis. The constant comparative method of data analysis is typically used for theory building, but is appropriate for much of the inductive data analysis that takes place in qualitative research. This method involves transcribing all data sources, including field notes, into raw data. All data transcriptions should be photocopied with originals stored in a safe place, unaltered. Photocopies of raw data are then used for analysis (Strauss & Corbin, 1990).

Strauss and Corbin (1990) provided direction for researchers to develop grounded theory using three types of coding: open, axial, and selective coding. They contend that concepts are the building blocks of theory. Open coding is the process by which concepts are identified and developed in terms of their properties and dimensions. This includes asking questions about the data, making comparisons for similarities and differences between incidents, events, or other phenomena. Similar ones are grouped to form categories. Axial coding is the set of procedures used to put data back together in different ways after open coding, by making connections between categories. This is achieved through use of a coding paradigm that includes (intervening) conditions, context, action/ interactional strategies, and consequences. They describe axial coding as the process of relating subcategories to a category, which is developed in terms of the causal conditions that give rise to it, its dimensional location in terms of its properties, the context, the action/interactional strategies used to handle, manage, and respond to the phenomenon in light of that context, and the consequences of such action. Selective coding is the process of selecting the core category, systematically relating it to the other categories, validating those relationships, and filling in categories that need further refinement and development. The core category is the central phenomenon around which all the other categories are integrated (Strauss & Corbin, 1990).

It is important to identify patterns and to group the data accordingly, because this is what gives the theory specificity. One is then able to say: Under these conditions (listing them), this happens; whereas under these conditions, this is what occurs. Once the data are related not only at the broad conceptual level, but also at the property and dimensional levels for each major category, one has the rudiments of a theory. Validating one's theory against the data completes its grounding.

The process of coding data includes looking for patterns and themes. DeSantis and Ugarriza (2000) defined a theme as "an abstract entity that brings meaning and identity to a recurrent experience and



its variant manifestations. As such, a theme captures and unifies the nature or basis of the experience into a meaningful whole” (p. 362). They suggest that a theme involves five aspects:

1. the overall entity or experience;
2. the structure or the basis of the experience;
3. the function or the nature of the experience into a meaningful whole;
4. the form or the stability or variability of the various manifestations of the experience; and
5. the mode or the recurrence of the experience.

They also elaborate on the four criteria they believe to be fundamental to defining themes: the theme’s emergence from data, abstract nature, iteration or recurrence of patterns of behavior, and levels of identification (which would include categories, domains, or taxonomies).

### **Analysis Issues**

Qualitative researchers have shared a variety of strategies used to analyze qualitative data. For example, to analyze data, LeCompte and Schensul (1999) suggest:

- looking at the theoretical framework. The theoretical underpinning provides the lens through which the data are viewed and helps the researcher to situate the results in the theory, which helps to facilitate the understanding of the data within that theoretical perspective.
- reviewing the research questions. The research questions in qualitative studies are used to guide the design and implementation of the study. They are the questions one wants answered by the study; hence, it is important to view the data in terms of ensuring that sufficient data were collected to enable the researcher to answer the questions posed within the study.
- creating vignettes. These snapshots provide an overview or summary of data, encouraging the researcher to organize the data into smaller segments that help to develop understanding.
- writing some history. The background and history of the topic under study is important in assisting the researcher to interpret the data. In many studies of cultures, for example, it is important to understand the past as a tool for understanding the present.
- describing a social process. Watching how participants interact is sometimes an excellent tool for helping the researcher to understand data that have been collected through observation. What talks to whom? What are the relationships among the participants being observed?
- creating summaries of interviews or survey results. Peer review is an effective way to help researchers begin to make sense of the data by encouraging the researcher to summarize what has been found thus far.
- creating collections of quotations. Organizing quotations related to the same topic can help the researcher to recognize patterns found in the data.

- creating a conceptual framework and drawing it. Providing a visual that encompasses the concepts one has found may be helpful to foster understanding of the data.
- writing in a narrative form. Sometimes just sitting down and beginning to write is a helpful tool for the researcher to organize his/her thoughts about the results.
- borrowing the narrative form of those from whom one is learning. This is especially important when dealing with marginalized populations; writing in an emic voice, using the participants' words, brings life to the results that the researcher's etic voice cannot accomplish.
- developing a metaphor. Presenting the data in various ways verbally is sometimes helpful.
- describing functions or organization structure of a group. This is similar to describing the social processes, in that, both social processes and organizational structure and function involve how people interact and relate to each other. Understanding those relationships facilitates one's understanding of the data.
- writing up the critical events in chronological order. This is an effective tool to help the researcher lay out the story from beginning to end.
- making a list of the most important facts. Like the chronology, such a list would help present the story's most important points.
- drawing a visual display of data. There are a variety of ways to display data, including charts, graphs, tables, figures, snapshots, vignettes, video, interactive plays, poems, and concept maps, for example.

LeCompte and Schensul (1999) further the analysis process with these strategies for interpreting the data: engaging in speculation, reviewing the research questions, reviewing relevant theories, contrasting the views of insiders with those of outsiders, seeing relevance to program/policy, evaluating the project, and considering the audience.

Wolcott (1994) describes analysis as relying on "agreed-upon knowledge, the recognition of mutually recognized properties or standards" which are "inherently conservative, careful, systematic" (p. 25). He presents the question of what should be analyzed or what to analyze for. He advocates the following ideas for doing initial analysis:

- Highlight the findings by emphasizing some data more than other data, summarizing what is important to the study.
- Display the findings in graphic representations, including visual displays like concept maps, photographs, videotapes, posters, or computerized drawings.
- Use and report "systematic" fieldwork procedures, such as those presented by Spradley (1972/1988).
- Flesh out whatever analytical framework guided the data collection, such as discourse analysis or content analysis.

- Identify patterns in the data, looking for relationships among the data, “what goes with what” (p. 33), carefully reporting what has been learned without use of generalizations.
- Compare with another case that is known.
- Evaluate by comparing with a recognized standard or include how those included in the research evaluate what is happening.
- Position the results in a broader analytical framework, drawing connections and relating them to theory.
- Critique the research process to determine how one knows what he/she knows; that is, with how much assurance can one say what he/she is saying?’
- Propose a redesign for the study, using what has been learned from the problems encountered in the study (Wolcott, 1994, 29-36).

Wolcott provides several suggestions for interpretation, including:

- Extending the analysis by asking questions derived from one’s investigation of the data.
- Making inferences using inductive reasoning.
- Stopping when one comes to the end, noting what needs to be done next.
- Taking the suggestions of one’s committee members, colleagues, or editors in whose journals one intends to publish.
- Looking at theory for both analysis and interpretation; the analytical analysis provides structure, while the interpretation provides a way to link to theory for explanatory purposes; one should particularly relate his/her results to competing theories and resolve the conflict in the discussion.
- Refocusing the interpretation on the basis of the tradition in which one is working.
- Connecting with one’s own experience, sharing what one “makes of it all” (p. 44).
- Analyzing the interpretive process, explaining what one did and what prohibited his/her further progress.
- Interpreting the analytical process, using counterintuitive ideas, looking for new ways to express one’s results and thoughts.
- Exploring alternative formats, including poetry, fiction, performance, and so forth. (Wolcott, 1994, 40-46).

Miles and Huberman (1994) also provide a variety of ways to ensure that the researcher is drawing and verifying conclusions correctly. They suggest that researchers use a variety of the following tactics for generating meaning from the data: Noting patterns and themes; seeing plausibility; clustering; making metaphors; counting; making contrasts/comparisons; partitioning variables; subsuming particulars into the general; factoring; noting relations between variables; finding intervening variables; building a logical chain of evidence; and making conceptual/theoretical coherence. They also suggest the following tactics for testing or confirming findings: Checking for representativeness; checking for researcher effects; triangulating; weighting the evidence; checking the meaning of outliers; using extreme cases; following up

surprises; looking for negative evidence; making if-then tests; ruling out spurious relations; replicating a finding; checking out rival explanations; and getting feedback from informants.

Numbers are typically associated with quantitative means of data collection; however, numbers can also be useful in analyzing qualitative data. The use of numbers or counting can be used to provide frequency counts to generate meaning or show the complexity of a theme. Counting is advocated by Beck (2003) as a tool for identifying patterns in data. Beck also suggests that numbers may be used to document, verify, and test one's interpretations or conclusions. In teaching nursing students to analyze data from written descriptions of patient advocacy, she listed six steps for data analysis, which include: Reading students' written descriptions; extracting significant statements or phrases that are directly related to the phenomenon under study; developing meaning from these statements; organizing the meanings into clusters of themes; using the data analysis results to write an exhaustive description; and using member checks to validate the description written. Member checking, as described by Lincoln and Guba (1985), involves testing the analytic categories, themes, interpretations, and conclusions with the participants of the study to establish credibility. This involves telling their stories in a way that enables them to recognize themselves in the writing, while telling the bigger story that encompasses the stories of all of the participants into a whole.

Jones (2002) highlights the importance of situating the researcher in the research. She notes that positionality is an integral aspect that needs to include the influence that the researcher's presence or position has on who and what can be known. Jones shares Janesick's (2000) belief that observations and interviews are not sufficient for research to be qualitative; the researcher also must "interpret the beliefs and behaviors of the participants" (p. 387). Further, she suggests that identifying themes is not sufficient for analysis. She advocates deriving meaning from engaging in an inductive analytic process that includes an understanding of the exact words and behaviors of the participants, using a number of analytic strategies. The appropriate strategies for analysis, she suggests, are those that produce findings that assist the researcher in developing a deep understanding of the phenomenon under study, with the story emerging from the words and behaviors of the participants and from the contexts in which they occur. These findings should then be combined with the insights, intuitive ideas, creativity and artistry of the researcher (Jones, 2002).

Doucet and Mauthner (1998) use a voice centered relationship model to analyze interview data, which includes several readings of the interview text. To begin, they read through the interview, looking for plot and story, identifying recurring patterns, words, phrases, and contradictions in the text. They then advocate relating the participant's story to that of the researcher's own experience, including the relationship between the researcher and the participant. This involves looking at one's assumptions and at perspectives that may affect one's interpretation. This is the locating of self in the story as related to the participant. Their second reading of the interview focuses on how the participant speaks about herself and her world. They then write up each participant's story in case study form. Doucet and Mauthner (1998) noted,

This approach is fundamentally different to the thematic organization characteristic of most methods of data analysis, including those assisted by computer programs. It delays the reductionist stage of data analysis when transcripts are cut up into themes and aggregated. This process shifts data analysis away from traditional coding, which implies fitting a person into a pre-existing set of categories, whether those of the researcher or those of established theoretical frameworks.

Doucet and Mauthner (1998) use this method to describe individual participants' voices, thereby helping them to maintain differences between the participants. They then move away from individual focus to looking at the data as a whole, breaking each transcript into overlapping themes and sub-themes. The analysis of the data therefore involved organizing the data in different ways (tapes; verbatim transcripts; 4 readings; case studies; summaries; themes) in order to tap into different dimensions of the data sets. It also involved a dialectical process of moving between different ways of organizing or representing the data, and between the details and particularity of each one of the individual respondent's experiences, and the overall picture of the samples as wholes.

A qualitative content analysis approach to data analysis involves reading and re-reading transcripts, looking for similarities and differences that enable the researcher to develop themes and categories. Ways to mark the text include coding paragraphs or other units of analysis; highlighting units of analysis with different colored ink, arranging the data into themes; cutting up transcripts and putting them in thematic folders; using a card index system, noting line numbers for easier cross-referencing; and using computer software to assist in data management activities, like sorting.

Pope, Ziebland, and Mays (2000) also provided strategies for analyzing data, using the framework approach, which includes becoming familiar with the raw data by immersing oneself in it; developing a thematic framework in which one has identified all the key issues, concepts, and themes; indexing all of the data in textual form by coding transcripts or short text descriptors; charting the data using summaries of experiences; mapping and interpretation of data using charts to define concepts, and mapping the range and nature of the phenomena, creating typologies and finding association between themes to find explanations and develop findings.

Miles and Huberman (1994) presented a series of sequential steps for conducting data analysis. These steps include developing codes and applying them to textual data, like field notes from observations, interviews, or documents/artifacts. Then they suggest noting the researcher's reflections on the observations and interviews. Patterns, themes, relationships between themes are then identified. The next step is conducting an investigation of common and different aspects and developing interpretations of findings, followed by verifying the interpretations through member checks, peer review, and triangulation. Additional means of data analysis suggested by Miles and Huberman (1994) include triangulating, using at least three different pieces of data from three different points of view. Triangulation may be viewed in terms of using a variety of both data collection methods and sources. They also suggest constantly comparing earlier data with later data as they are collected, using different bases for comparison, and categorizing and sorting data, using visual displays to enable the researcher to look at the ways that data

develop into categories. One might also look at how the codes develop categories. Miles and Huberman further advocate ordering data and reordering them in various ways, such as by chronology, importance, or frequency, for example. Contrasting data to determine what fits the researcher's assumptions or others' findings is another suggestion for analyzing data they share, along with developing one's own hunches based on his/her instincts about what is happening. Another means they suggest for analyzing data is restating the question to fit the data one has collected. Equally important is providing a visual representation of the data, looking for metaphors that are envisioned as the researcher looks at the diagrams, sketches, charts, and so forth. Other data analysis techniques suggested by Miles and Huberman include summarizing and distilling the data into smaller segments that capture the flavor and meaning of each piece of data; using peer review to discuss one's findings with others, as having to condense the findings into a few words or sentences will help researchers to distill what they have learned from the data; member checking, which may be used to verify information collected, but it also provides a powerful tool for checking out one's interpretations of the participants' stories; stepping away from the data, then returning to it refreshed and with a different perspective on what one is seeing; reiterating one's theoretical lens used to view the data – that is, how does the theory frame the findings (Miles & Huberman, 1994)?

### **Reporting Findings**

In reporting data findings, one does not analyze the data on a question by question basis. Instead, one summarizes key themes, using selected quotes to illustrate findings. Merriam (1998) notes that findings may be presented in different levels of analysis, such as organized, descriptive accounts, themes found repeatedly throughout the data, or as models or theories that explain the data. Merriam (1998) discussed several levels of data analysis. At the basic level, a narrative may be used to present data that have been organized chronologically or topically. The next higher level of analysis involves using concepts developed from the themes and categories to describe phenomena. The third level of analysis Merriam mentioned involves “making inferences, developing models, or generating theory” (p. 187). It is at this phase of analysis that the moving back and forth between the details and a landscape view of concepts occurs, as described by Miles and Huberman (1994). One should present a summary of the findings, using the literature and member checks to substantiate or negate one's findings.

### **Analysis in Action**

Because there are so many techniques described in this article, I sought a means for organizing them to make it easier for readers to identify techniques that they might find useful in their research. Applying the procedures suggested by the authors mentioned in this article, I read and reread their suggested techniques and grouped them into categories. The categories that emerged for me – other researchers may view them differently – included five categories or stages of analysis, extending those described earlier in this paper by Merriam (1998). They include:

1. **Narrative** – These techniques include suggested actions that are preparatory to the coding process or which help the researcher to derive meaning through narrative approaches, such as narrative reading or writing of textual data.

2. **Coding** – These techniques apply to those actions that involve organizing and reorganizing the data into categories that enable the researcher to identify relationships between and among categories.
3. **Interpretation** – These techniques are those actions that enable the researcher to make meaning from the narrative and coding activities and facilitate the researcher's understanding of the conceptual framework generated through the coding process.
4. **Confirmation** – These techniques enable researchers to be confident that the interpretations they have derived are from the data and not from researcher construction.
5. **Presentation** – These techniques are those actions that researchers may use to present the findings to a particular audience in a cogent manner and which serve as a tool for further analysis and explanation.

In Table 2, the techniques discussed in this article by LeCompte and Schensul (1999), Wolcott (1994), Beck (2001), Doucet and Mauthner (1998), and Miles and Huberman (1994) are presented as I viewed their fit into these categorical stages.

Table 2

*Summary Table of Techniques Used at Various Stages of Analysis*

STAGE	LeCompte & Schensul (1999)	Wolcott (1994)	Beck (2003)	Doucet & Mauthner (1998)	Miles & Huberman (1994)
Narrative	<p>Review research questions</p> <p>Write some history</p> <p>Describe a social process</p> <p>Create summaries of interviews</p> <p>Create collections of quotations</p> <p>Describe functions/ structures of group</p> <p>Write up critical events chronologically</p> <p>Make list of important facts</p>	<p>Connect to your own experience</p>	<p>Read written descriptions</p>	<p>Relate participant's story to your own experience</p> <p>Locate self in the story as related to participant(s)</p> <p>Look at how participants speak about self and their world</p>	<p>Making metaphors</p> <p>Note reflections on collected data</p>
Coding	<p>Create vignettes</p> <p>Create a conceptual framework</p>	<p>Identify data patterns</p> <p>Extend analysis by asking questions derived from the data</p>	<p>Develop meaning from the statements</p> <p>Organize meanings into clusters of themes</p>	<p>Break down text transcripts into overlapping themes and sub-themes</p> <p>Organize data in different ways</p>	<p>Note patterns and themes</p> <p>Cluster</p> <p>Partition variables</p>



				to tap into different dimensions of data sets	<p>Subsume particulars into the general</p> <p>Factor</p> <p>Note relations between variables</p> <p>Find intervening variables</p> <p>Follow up surprises</p> <p>Develop codes and apply to textual data</p> <p>Identify patterns, themes, relationships between themes.</p> <p>Conduct an investigation of common/different aspects</p> <p>Categorize and sort data</p> <p>Order and reorder data by chronology, importance, frequency</p>
Interpretation	<p>Develop a metaphor</p> <p>Look at theoretical framework; review relevant theories</p> <p>Engage in speculation</p> <p>Look for relevance to program/policy</p>	<p>Relate to theory</p> <p>Refocus on the basis of your tradition/discipline</p> <p>Evaluate against a standard or against participants' interpretation</p>	<p>Extract significant statements related to the phenomenon under study</p>	<p>Shift focus from individual cases to groups</p>	<p>Look for plausibility</p> <p>Build a logical chain of evidence</p> <p>Make conceptual/theoretical coherence</p> <p>Weight evidence</p>

	Evaluate the project	<p>Position results in a broader analytic framework</p> <p>Make inferences using inductive reasoning</p> <p>Flesh out analytical framework</p>			<p>Check meaning of outliers</p> <p>Use extreme cases</p> <p>Make if-then tests</p> <p>Rule out spurious relations</p> <p>Develop interpretation of findings</p> <p>Contrast data to determine what fits your assumptions or others' findings</p> <p>Develop hunches</p> <p>Restate question to fit data</p>
Confirmation	Contrast insider views with outsider views	<p>Critique the research process</p> <p>Report systematic fieldwork procedures</p> <p>Propose a redesign of the study</p> <p>Stop when you come to the end, asking what needs to be done next?</p> <p>Compare to a known case</p> <p>Analyze the interpretive process</p>	<p>Use member checks to validate the written description</p> <p>Use numbers to document, verify, and test interpretations</p>	Look at one's assumptions	<p>Triangulate</p> <p>Count</p> <p>Make contrasts and comparisons</p> <p>Check for representativeness</p> <p>Check for researcher effects</p> <p>Look for negative evidence</p> <p>Replicate a finding</p> <p>Check out rival explanations</p> <p>Get feedback from participants</p>

					<p><b>Verify interpretations by member checks, peer review, triangulation</b></p> <p><b>Constantly compare earlier data with later data using different bases for comparison</b></p>
<p><b>Presentation</b></p>	<p><b>Consider the audience</b></p> <p><b>Draw visual display</b></p> <p><b>Write in narrative form, borrowing form from participants</b></p>	<p><b>Emphasize important data</b></p> <p><b>Take suggestions from editors/ committee/ colleagues</b></p> <p><b>Display findings graphically</b></p> <p><b>Explore alternative formats for presentation</b></p>	<p><b>Use data analysis results to write an exhaustive description</b></p>	<p><b>Write up results in case study form</b></p>	<p><b>Use visual displays</b></p>

### **Challenges Faced in Qualitative Data Analysis**

Beck (2003) noted common pitfalls that beginning researchers make in analyzing qualitative data. These pitfalls include data shuffling, premature closure, and overly delayed closure. She explained that sometimes it is easier for novices to sort data than to move through the process of analysis and interpretation. Typical pitfalls she noted were related to closure, such as when the researcher stops collecting data before all categories are saturated or before enough data are collected or “a solid conceptual outcome is created.” She noted that analysis of qualitative data is complex and requires that data be organized and reorganized, presented and represented. Lincoln and Guba (1985) pointed out that knowing when to quit collecting data and concentrate on simply analyzing collected data is determined by the researcher’s having exhausted sources, saturation of categories, emergence of regularities in the data, and over-extension.

### **Teaching Qualitative Data Analysis**

There are numerous ways to teach qualitative data analysis. One of the ways data analysis can be taught to students is to have them participate in a group research project in which they each collect data addressing the same research questions, then have them code and analyze the data individually and in groups. In the past, I have used an approach to teaching data analysis, beginning with the discussion of interviewing and observation skills. When teaching interviewing, for example, I have the students develop an interview guide that they use individually to collect data. They are given a couple of weeks to complete the interview, transcribe it, and initially code it. They bring their coded transcripts to class, where they work in small groups to develop a coding scheme that addresses each group member’s interview, but which also is representative of the small group’s collective data. Each small group then shares its coding scheme, after which they develop an overriding coding scheme that addresses all interviews conducted by class members. Through this process the analysis is taking place using the techniques addressed in the previous section of this paper. The small groups then write up their analysis of the data and present it visually to the class.

Another way used by professors to teach data analysis involves putting a basket of candies of various types on a centrally located table. Student grab a handful of candy and sorts/code it in various ways, then write a small summary of what they found. This has been conducted with buttons, having students sort buttons of various types, colors, materials, and so forth.

One of the most fun ways my students have learned about the coding and analysis aspects of qualitative research involves what I call the “Fantasy Exercise.” In this activity, students write on a Post-it note their personal answers to several questions and tape them to the board or wall. After reiterating that this is a fantasy, I ask them to answer the following questions, with their answers to each question put on a separate Post-It note, omitting their identifying information. The questions are: Given all of the money you need, where in the world would you like to go for a month-long vacation? With whom would you like to go on this vacation (it cannot be a spouse, partner, family member, or friend; it must be someone living)?

What would you like to do on this vacation with this person? The answers have provided some hilarity in this activity, and such levity eases their nervousness about their ability to do data analysis. They then look at various ways that the notes can be arranged and rearranged to make sense of the data, and they build a visual display of the analysis to illustrate their findings.

### Summary

The process of analyzing qualitative data varies from one study to another, depending on how the researcher is guided by the research questions, the theoretical framework of the study, and the appropriateness of the techniques for making sense of the data. The purpose of analysis is to interpret and, hence, convert the data into a story that describes the phenomenon or the participants' views, using the emic perspective. The process typically involves collecting data that will inform the study, breaking down the data into various categories and making connections between these categories in terms of relationships among them, then visually displaying the interpretation and writing it up for dissemination. While the discussion of data analysis techniques provided in this article is by no means complete, it may provide novice researchers with some ideas for activities that will help them to begin analyzing data. Further, it is my hope that veteran researchers will find such a review of techniques useful for both conducting and teaching data analysis.

### References

- Beck, C. T. (2003). Initiation into qualitative data analysis. *Journal of Nursing Education, 42*(5), 231.
- Bernard, H. R. (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage Publications.
- DeSantis, L., & Ugarriza, D. N. (2000). The concept of theme as used in qualitative nursing research. *Western Journal of Nursing Research, 22*, 351-372.
- Doucet, A., & Mauthner, N. (1998). *Voice, reflexivity, and relationships in qualitative data analysis: Background paper for workshop on "Voice in Qualitative Data Analysis."* Retrieved August 15, 2004, from [http://www.coe.uga.edu/quig/proceedings/Quig98\\_Proceedings/doucet\\_mauthner.html](http://www.coe.uga.edu/quig/proceedings/Quig98_Proceedings/doucet_mauthner.html)
- Glaser, B. G., & Strauss, A. L. *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation*. San Francisco: Jossey-Bass.
- Huberman, A. M., & Miles, M. B. (1998). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 179-210). Thousand Oaks, CA: Sage Publications.

- Janesick, V. J. (2000). The choreography of qualitative research design: Minuets, improvisations, and crystallization. In N. K. Denzin & Y. S. Lincoln (Eds), *Handbook of qualitative research* (2<sup>nd</sup> ed., pp. 379-400). Thousand Oaks, CA: Sage.
- Jones, S. R. (2002). (Re)writing the word: Methodological strategies and issues in qualitative research. *Journal of College Student Development*, 43(4), 461-474.
- LeCompte, M. D., & Schensul, J. J. (1999). *Analyzing and interpreting ethnographic data*. Walnut Creek, CA: AltaMira Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. (2<sup>nd</sup> ed.). London: Sage.
- Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. London: Sage.
- Pope, C., Ziebland, S., & Mays, N. (2000). Qualitative research in health care: Analyzing qualitative data. *British Medical Journal*, 320, 114-116.
- Spradley, J. P., & McCurdy, W. (Eds.). (1988). *The cultural experience: Ethnography in complex society*. Prospect Heights, IL: Waveland Press.
- Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, interpretation*. Thousand Oaks, CA: Sage.

*Barbara B. Kawulich is an Assistant Professor in the Department of Educational Leadership and Professional Studies at State University of West Georgia. Correspondence concerning this article should be addressed to Dr. Barbara B. Kawulich, Department of Educational Leadership and Professional Studies, State University of West Georgia, Education Annex, Carrollton, GA 30118-5160.*