

The Pennsylvania State University

The Graduate School

Department of Educational Psychology, School Psychology, and Special Education

**A PRELIMINARY INVESTIGATION OF TREATMENT FIDELITY ON
RESEARCH OUTCOMES**

A Thesis in

Educational Psychology

by

J. Kevin Ross

© 2009 J. Kevin Ross

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2009

The thesis of J. Kevin Ross was reviewed and approved* by the following:

Hoi K. Suen
Distinguished Professor of Educational Psychology
Thesis Advisor

Robert L. Hale
Professor of Education

Rayne A. Sperling
Associate Professor
Professor-in-charge

*Signatures are on file in the Graduate School

ABSTRACT

A secondary analysis of data collected from the GE Math Excellence curriculum intervention is conducted to establish whether, or to what extent, the accuracy of implementation was compromised by differential patterns of compliance with instructions on the part of participating students and teachers. Two forms of student compliance were analyzed: 1) whether students identified themselves at the top of a math attitudes questionnaire, and, 2) whether students selected a single favorite subject, as instructed. Teacher supervision of survey completion might suffice to explain the former, while misinterpretation on the part of students might be suggested by the latter.

The first working hypothesis is that students who complied with instructions when completing a math attitudes survey demonstrated more favorable attitudes towards mathematics than did students who did not comply. Results of the data analysis failed to reject the null version of this hypothesis, (i.e., that there were no differences among students' math attitudes vis-à-vis whether or not they complied with instructions for survey completion), suggesting the possibility that teacher supervision adequately explains student compliance.

The second working hypothesis is that students complying with instructions to report a single favorite subject demonstrated more favorable attitudes towards mathematics compared with students who did not comply. Results of the data analysis failed to reject the null version of this hypothesis, (i.e., that there were no differences among students who did/did not select a single favorite subject vis-à-vis whether or not they identified themselves at the top of the page. Aside from teacher supervision of survey completion, the word *favorite* was not interpreted to mean one and only one subject, (i.e., the survey design may not have been suited to students' reading comprehension).

TABLE OF CONTENTS

LIST OF FIGURES.....	v
LIST OF TABLES	vi
ACKNOWLEDGEMENTS	vii
Chapter 1 INTRODUCTION	1
Chapter 2 LITERATURE REVIEW	7
Chapter 3 ANALYSIS AND RESULTS.....	24
Chapter 4 DISCUSSION OF RESULTS	31
REFERENCES.....	37

LIST OF FIGURES

Figure 3-1 : Histogram of MLIKE1 and IDC	27
Figure 2-1 : Histogram of IDC and FSC.....	29

LIST OF TABLES

Table 3-1 : Row and Cell Counts for MLIKE1 and IDC.....	26
Table 3-2 : Chi Square Tests for the relationship between MLIKE1 and IDC	27
Table 3-3 : Row and Cell Counts for IDC and FSC.....	28
Table 3-4 : Chi Square Tests for the relationship between IDC and FSC.....	29

ACKNOWLEDGEMENTS

The contents of this thesis reflect the generous and committed support on the part of the following scholars: Professor Hoi K. Suen and Professor Robert L. Hale. The principal investigators associated with the G. E. Math Curriculum project, Dr. John C. Wise and Dr. Rose M. Zbiek provided encouragement with regard to this secondary analysis of the data collected for their project.

Chapter 1

INTRODUCTION

A secondary analysis is conducted to examine the extent to which problems encountered during assessment of a mathematics curriculum intervention affected the accuracy of its results. Specifically, noncompliance with investigators' instructions concerning completion of a math attitudes survey resulted in missing and incomplete data. Inasmuch as interpretation of intervention efficacy is inextricably linked to accurate implementation, noncompliance poses a threat to validity. Moreover, missing or ambiguous data can increase error variance, (e.g., the ratio between true and reported effect size, under the assumption that reported effect size is not equal to zero). The research question is: What is the cost of noncompliance with instructions, where cost is calculated in terms of the quality of student exposure to the curriculum intervention?

The data towards which this secondary analysis is directed was collected to assess the efficacy of the GE Math Excellence curriculum intervention, hereafter referred to as GEM, (e.g., the GEM intervention, GEM research design, etc.). This chapter presents a rationale as to why such an analysis might be informative, beginning with a summary of the GEM project and its assessments. Possible explanations as to why compliance vis-à-vis classroom-based research might warrant investigation are then presented, followed by an outline of thesis content at the conclusion of the chapter.

Summary of GEM

Seventy nine students were assessed with regard to their math-related attitudes, beliefs, and opinions in the context of participation in a mathematics curriculum intervention. A math attitudes questionnaire was constructed for this assessment, hereafter referred to as the MATT. “*Math in a New Technology Context*,” as the curriculum intervention was titled for purposes of interaction with teachers and administrators at the Midtown School District’s schools, sought to improve students’ math achievement test scores through an innovative mathematics curriculum that would also facilitate more favorable beliefs and attitudes concerning mathematics. Through exposure to technological innovations during a series of math classes, (e.g., calculators, cell phones, and radio-tracking devices), students' attitudes and beliefs concerning the importance of math would improve, especially with regard to its importance concerning career choices.

The math attitude scores from the seventy nine students were measured at the beginning of matriculation in the sixth grade. Several characteristics of the GEM research design require explication in order to appreciate the justifications for this secondary analysis. An ongoing process of adjustment to the initial GEM design occurred during implementation, suggesting that interpretations of intervention efficacy be regarded as both formative and summative evaluations of GEM (Popham, 1988). Moreover, several characteristics of the GEM research design suggest that it be regarded as a pilot study, (i.e., insofar as such concerns guide decisions about the appropriate statistical techniques), including the absence of prior validation of the math attitudes questionnaire. Consideration

of the GEM assessment data as a pilot study places emphasis on accurate implementation of the research design, which is a prerequisite to the eventual analysis of intervention fidelity.

The assessment of GEM intervention efficacy included both qualitative and quantitative components, (e.g., interviews with teachers and students, classroom observations, math attitude, and math achievement scores). All sixth grade students enrolled at the Midtown School District (during project year one) were exposed to the GEM intervention. In the absence of either random sampling or assignment, comparisons at the group level will have to rely on statistical controls such as analysis of covariance (ANCOVA). Quasi-experimental designs without randomization, though possible, are inappropriate for this analysis due to teacher attrition. Student level pre- and post-intervention math achievement scores, (e.g., in the form of the Pennsylvania System of School Assessment), were utilized to examine intervention efficacy.

Theoretical Concerns

Regarded from the standpoint of a pilot study, the assessment of GEM intervention efficacy has as its primary focus the issue of accurate implementation of its research design, which is prerequisite to later concerns with treatment fidelity viz. its external validity. One concern with accurate implementation viz. fidelity lies in the extent to which students complied with instructions for the completion of the GEM assessments. Consideration of why student compliance might or might not be informative is necessary for theoretical purposes. Both forms of compliance, student identification and selection of a favorite subject, are extraneous to the measurement of math-related attitudes and beliefs. However, it is possible that either

or both can serve as an indicator(s) of interest, engagement, or commitment to mathematics, thereby distinguishing a qualitatively different exposure to the mathematics classrooms wherein the GEM intervention occurred.

Such qualities might include greater motivation towards learning the mathematics embedded in the technology-enhanced curriculum modules and/or with motivation to learn mathematics in general. Furthermore, motivation might interact with students' self-perception of ability with regard to acquiring mathematical skills, which might be linked to prior math achievement scores. Thus, it is proposed that student compliance with instructions serves as a proxy for qualitatively different forms of engagement with the GEM curriculum, which might mitigate the relationship between exposure to the GEM intervention and subsequent (changes in) math achievement scores. A more valid interpretation of the relationship between exposure to the GEM curriculum intervention and math achievement scores could be obtained by parsing the error variance attributable to noncompliance and/or by controlling for the pre-intervention math achievement scores.

The instructions upon which compliance was assessed included student identification, (e.g., indicating one's name at the top of the MATT questionnaire), and, identification of a favorite subject, one of the items on the MATT. As such, these indicators of noncompliance might suggest problems with administration, (e.g., lack of supervision by teachers to ensure that all questionnaires had been signed) as well perhaps with construction of the survey instrument, (e.g., a comprehensive list of all subjects for the sixth and seventh grades at the exclusion of a write-in response).

Therefore, responsibility for noncompliance need not rest solely upon students. The two-fold goal of this analysis is directed at the concern that compliance might suggest a

qualitatively different exposure to the curriculum intervention as well as to examination of whether problems with adherence to the research design might be implicated in these proxies of noncompliance. This second goal concerns the relationship between accurate implementation and validity; it bears mention that noncompliance was not a rare occurrence.

Alternative interpretations of compliance are also plausible, (i.e., extraneous to either math achievement or exposure to the GEM intervention), such as noncompliance as an indicator of disciplinary problems, or, to deficits in attention. Bearing in mind that the instructions pertained to an assessment of math attitudes rather than of math achievement, (i.e., the importance of measuring math attitudes for investigators' interests might differ from its importance for either teachers or students), and compliance with these instructions might similarly differ from compliance with instructions concerning intervention content; that is, confounding the measurement of an attitude towards math with an attitude towards math attitude assessments, especially when given at the beginning of the school day.

Regarded from the standpoint of adherence to a research design, (i.e., accurate implementation), it is also plausible that compliance, (e.g., signed questionnaire), serves as an indicator of teachers' commitment to the GEM intervention, insofar as they hadn't followed instructions concerning supervision of students' completion of MATT. However, it might also be asking too much to expect that teachers could inspect each questionnaire to ascertain whether a single response was given for identification of a favorite subject, given that the questionnaire was administered during homeroom period on a Monday morning (at the bequest of school administrators), which often includes a variety of administrative tasks and concerns).

Insofar as instructions on the MATT questionnaire, (e.g., how to properly complete the questionnaire), might not have been clear to students, it should be noted that items on the questionnaire had not been previously validated, though its designers thought otherwise. Theoretically, it is possible that student compliance was an artifact of the ‘user-friendliness’ of the MATT instrument, rather than a measure of psychological commitment, whether on the part of teachers, students, or both. Of course, it might be a mixture of all three propositions, such that the topic will be revisited in the discussion of results, at the end of the thesis.

Conclusion

The Literature Review in Chapter Two focuses on prior research on student motivation with regard to self-perceived ability and studies concerned with the assessment and implementation of (classroom-based) curriculum interventions. Chapter Three, Analysis and Results, directs attention to the relationships between (the types of) compliance and global math attitudes. It also presents a rationale for looking at the relation between the two measures of compliance. In Chapter Four, Discussion and Conclusion, the costs of noncompliance with instructions are considered with regard to the assessment of GEM intervention efficacy as a pilot study, including alterations that might be made to the GEM research design. A framework is suggested in which to locate concerns viz. intervention fidelity in order to integrate its measure with that of an evolving interpretation of validity.

Chapter 2

LITERATURE REVIEW

Introduction

Selections from the literature on classroom compliance vis-à-vis accurate implementation of a research design are guided by constraints imposed on this analysis by the GEM research design, and specifically with regard to its assessments; investigators had indirect contact with grade six students for the purpose of math attitude assessment; teachers administered the MATT questionnaire.

Another attribute guiding selections from the literature is the manner or extent to which prior research utilizes an analysis of costs, which is preferred insofar as it translates into a common rubric for the different purposes of this study. James Popham provides a useful framework, within which one might use one or several orientations to cost, for purposes of educational evaluation.

When an evaluator accepts the “benefits forsaken” conception of costs, then several implications are immediately apparent...to calculate the cost of anything, it is necessary to determine the value of *at least one alternative*...costs are irrevocably and exclusively

ted to decisions. It is not teachers' services that constitute a cost; it is the *decision* to use teacher services in one way rather than in another (Popham, 1988).

Among the consequences of the No Child Left Behind legislation (NCLB), observed at the Midtown School District, has been a reallocation of teacher services towards a greater placement in mathematics classrooms. Thus, the opportunity costs not only include less allocation towards instruction in other subjects, (e.g., art), but greater variability among mathematics instructors relative to their experience and training with regard to mathematics.

Two caveats must be made: 1) evaluators include teachers, administrators, and investigators, (i.e., variable interests among stakeholders might yield different estimates or appraisals of costs), and 2) costs calculated in terms of forsaken benefits, (i.e., opportunity costs), require a common *currency*, though it need not be measured in dollars. "A final implication of this conception of opportunity costs is that it tends to make costs occur in the future... such costs must be estimated," (Popham, 1988). Since decisions concerning evaluations are future-directed, it might be useful to think of cost estimates in terms of a line of credit, (e.g., costs incurred viz. future math achievement scores, yielding a positive or negative balance as a measure of cost effectiveness).

One Family – Four Siblings

We have considered, then, four different cost-analytic techniques that educational evaluators may employ in an effort to guide decision makers as they consider the respective merits of alternative programs.

For *cost-effectiveness analysis*, the divisor is each program's performance on an identical measure of effectiveness such as pupils' average test scores. For *cost-benefit analysis*, the divisor is each programs actual or estimated yield in monetary values. For *cost-utility analysis*, the divisor is each program's payoff in terms of judged social utility on a numerical utility scale (Popham, 1988).

These forms of cost-analytic techniques are brought to bear on the results of the secondary analysis, and are identified in literature vis-à-vis classroom compliance, where applicable. It is suggested that the current investigation lends itself to the cost-effectiveness technique, though perhaps to the cost-utility technique, as well.

Theoretical Concerns Concerning Classroom Compliance

The attributes by which compliance was measured suggest an interactionist approach to this variable. Given that attribution for responsibility, viz. compliance with instructions, may be directed at each category of GEM participant, it seems reasonable to approach classroom-based intervention and assessment as occurring in the process of transactions among students, teachers, and investigators. John Hetherington provides a useful summary of alternatives to positivistic methodologies, (viz. the contributions of Thomas Kuhn and Imre Feyerabend), which bear upon a transactional approach to the (study of) the classroom.

Transactional methods posit that scientific data are literally created from the interactive interchange between an investigator and participant. Constructivism, the predominant contemporary transactional approach...is unique in its denial of the existence of a single reality. Constructivism argues that multiple realities are socially constructed...None of these constructed realities is more or less a true reflection of reality, but some may be more or less informed or sophisticated (Hetherington, 1995).

Given the dire consequences for school districts not meeting demands for improved achievement scores, one might speculate that the NCLB legislation has rendered such a relativistic approach tenuous; transactions among administrators, teachers, students, and investigators might be structurally more uniform, (i.e., objectively accessible). To paraphrase W. I. Thomas, what the NCLB legislation defines as real has become real in its consequences.

Yet a certain willingness to abandon (pretenses towards) objective assessment is at issue with regard to use of the constructivist paradigm. Should one find the price of such a commitment, (i.e., to a relativistic interpretation), too high, one might still benefit from a transactional approach to classroom compliance, (e.g., power relationships in the context of student-teacher interactions). At the same time, a transactional approach is compatible with *stakeholder theory*, as discussed below.

Donetta J. Cothran and Catherine D. Ennis provide evidence of the utility of such an approach in their analysis of conflicts of interest, reciprocal power, and negotiations in the classroom.

Conflicts of Interest and Power Relationships

The interactive nature of the teaching process is built on a social relationship between teachers and students. Social relations reside on a continuum that extends from consent to domination (Burbules, 1986). Consent is the ideal relation because it is based on a commonly recognized purpose on which all members agree, resulting in few conflicts of interest. At the other end of the continuum, domination has a very high degree of conflict with little concern for group relations. Consent and domination rarely occur at the school level. Instead, a range of compliant behavior is manifested in schools (Cothran, 1997).

For present purposes, it is sufficient to note that classroom compliance might oscillate, (e.g., a student might be more or less compliant with one teacher vs. another, or, in a classroom with her friends vs. in one with peers).

Reciprocal Power and the Negotiated Classroom

Both teachers and students held resources desired by the other, and consequently both held power that they used to influence the nature of the class. For example, students knew that teachers needed their compliance, and they used that knowledge to negotiate for the opportunity to engage in their favorite content and spend time with friends. In turn, the teachers knew that students valued grades and did not value an educational focus as much as other aspects of the class (Cothran, 1997).

Among factors proposed to manifest in the form of compliance with instructions at the student-level of analysis, motivation and ability have been contrasted with attention (span) and discipline, as plausible explanatory variables. Review of literature with regard to motivation is considered next insofar as it pertains to metacognition; that is, motivation towards academic achievement might be coupled with help-seeking behavior, on the part of students in recognition of obstacles to successful learning, (constraints on ability, missing information, or misunderstood instructions).

Academic help-seeking behavior is the culmination of a complex sequence of cognitive and metacognitive processes that have resulted in the identification of a problem and a decision to seek a solution. Current models of this sequence include motivational factors, achievement goals, attitudes toward seeking help and attributions for the cause(s) of the problem among other processes. The extensity of theoretical connections between these factors, goals, and processes suggests the possibility that gender might mitigate the selection of help seeking among various alternative choices for problem resolution

As is the case with class participation in general, a request for help from the teacher can be motivated by a variety of goals and beliefs. Attempts to draw attention to oneself by distracting the teacher with an irrelevant question are decidedly different from requests for information that enable the student to solve the problem himself. The concept of *adaptive help seeking* has been used to distinguish among various goals and motivational factors underlying a particular request for assistance (Newman, 1990; Ryan and Pintrich, 1997).

This form of help seeking is characterized by an intrinsic motivation to acquire sufficient understanding of a problem, often by way of hints, such that help will no longer be needed to solve it or a similar problem in the future. Whereas, extrinsic motivations are directed at immediate rewards, or, used to expeditiously solve a problem (in order to complete an assignment), with little concern for what might happen should a similar request be made on the student in the future (Ryan and Pintrich, 1997; Butler and Neuman, 1995).

Help-seeking behavior stands out among other learning strategies as necessarily social in terms of access (Newman, 1991), which lends itself to the transactional approach of this analysis. Moreover, compared with rehearsal, drilling, and self-monitoring, this strategy is not entirely within the control of the student, (e.g., access to helpers might be blocked by other scheduled activities for either the student or the teacher). The type of help sought, ranges from *executive aid* (used to expeditiously complete an assignment) to *instrumental aid* (used to solve a problem through understanding (such that a similar problem can be solved in the future), which can be seen as more or less contributing to the student's development as a self-regulated learner (Newman, 1991; Butler and Neuman, 1995).

The self-regulated learner has been characterized as actively involved in learning strategies that result in academic achievement in a context of self-monitoring and self-control. The cognitive and metacognitive skills needed to accomplish self-direction are presumed to build during elementary school and include information-processing techniques. Thus, one might predict a shift in help seeking, from executive to instrumental, as the child ages. The evidence for such a shift is unclear; indicating that mastery or intrinsic motivation is mediated by additional factors some of which might be situational (Butler and Neuman, 1995).

Cost-effectiveness might be used to analyze help-seeking behavior vis-à-vis compliance with instructions to complete the MATT questionnaire, (i.e., there is an identical measure of effectiveness), though it would not lend itself to the distinction between intrinsic and extrinsic forms of motivation. A cost-benefit analysis is suggested when costs might entail loss of face, (e.g., self-esteem), or when the counterpart to help-seeking, (i.e., unsolicited help is provided), is considered with regard to motivation.

Among the various resources a child might draw upon, seeking help from teachers or peers entails possible risks to self-esteem as well as potential benefits, (e.g., increased achievement scores). The fear of asking a stupid question has silenced many a well-intentioned student. The motivation to seek help is related to achievement goals and includes an implicit belief that available help will prove efficacious in solving the problem. It's not clear that such a belief generalizes, (i.e. extends from one classroom or teacher to another (Ames and Lau, 1982).

Aside from its inherently social context, help-seeking behavior has a counterpart that also distinguishes it from other learning strategies by suggesting a reciprocity in the form of help offered, (e.g., by the teacher to facilitate compliance with instructions). In this instance, identification of an obstacle to successful learning by the teacher can affect student motivation with regard to metacognitive appraisals. Unsolicited help might be offered to (or imposed upon) a student once a problem had become identified perhaps through a failed exam or less-than-successful completion of an assignment. The relationship between other vs. self-identification of the problem might inform the impact such evaluations have on prior attitudes and attributions related toward help seeking (Graham and Barker, 1990).

Previous research has demonstrated gender differences in attribution processes and achievement outcomes (Ickes and Kidd, 1976). Gender effects have also been documented in studies of attitudes toward help seeking (Ames and Lau, 1982). To the extent that gender effects are apparent with regard to help-seeking behavior, it suggests a possible interaction, for purposes of the GEM pilot study, (e.g., between compliance with instructions and gender).

Gender differences have been demonstrated in the reaction to unsolicited help.

To test whether there were differences in ability attributions for poor performance on the pattern-completion task; an analysis of variance was performed with gender and whether help was given as independent variables and ratings on “ability level” as the dependent variable. A significant main effect of whether help was received, $F(1, 39)$

= 4.14 ($p=.05$) and a gender- by-help interaction, $F, (1, 39) = 3.97$

($p=.05$) emerged (Daubman and Lehman, 1993).

For our purposes, this gender difference hangs together with gender-based attribution patterns, but had a surprising impact on related behaviors after the help was received (Feather and Simon, 1971). Women who received help performed better on a subsequent task than women who did not receive help. By contrast, men who received help performed worse on a subsequent task than men who had not received help. Whether this difference can be explained by threats to self-esteem associated with receiving help (though it was not sought) or whether men “devalued” the assistance by spending less time, (i.e. reflecting its lack of importance), is unclear. At the very least, this difference suggests that generalizing attitudes towards help based on a specific form of help is problematic and might involve different processes for men than for women. Moreover, since the unsolicited help was received from a peer (as compared to a professor or teaching assistant), it poses an additional dimension: the status of the helper.

Will attitudes toward this form of (unsolicited) help generalize to help seeking requests? Attitudes and attributions toward help seeking are affected by the perceived efficacy of the help (Ames and Lau, 1982; Butler and Neuman, 1995). It's conceivable that attitudes toward help in general (sought or unsolicited) are changed by prior experiences with help. Both forms of helping behavior are embedded in social contexts that might include costs to self-esteem, (e.g., of publicly requesting help and/or enduring it when unsought). There's evidence to suggest these links between attitudes toward help and actual help can be acquired

vicariously, as well. A student watches a peer squirm in discomfort as unsought help is offered or perhaps empathizes with a student whose request for help has negative consequences; both are viable cascade effects as seen within modeling theory.

A discussion of these gender effects in the context of the attitude-behavior relationship has implications for modeling attribution processes. The attitude-behavior relationship has vexed social scientists for much of the last century such that attitudes as constructs acquired greater elaboration and latent characteristics were linked to other cognitive and affective processes such as beliefs and values. The inherent flexibility that such elaboration gives to the construct can be meaningful, ranging from the use of highly specific attitude targets, (e.g., asking help from Mrs. Jones during math class) to the use of categorical forms of attitudes (e.g., stereotypes regarding teachers). However, these extensions into non-affective domains such as beliefs and perceptions might pose problems for a path model (e.g., in determining where cognitive components of attitudes end and where similar components of attributions begin).

The ability to predict a student's behavior from prior knowledge of her attitude is only one way to approach this relationship. The behavior might precede the attitude formation, at least in terms of its specificity, and perhaps in regard to attitude change as well; in other words, it's just as conceivable that knowledge of behavior predicts a related attitude. A student with a transient attitude toward help seeking might have an experience with help-seeking behavior that feeds back onto the attitude itself, perhaps serving to concretize or even alter the earlier version.

Other than verbal self-report (of attitudes), related behaviors arguably provide the best operational definition of the attitude target. In addition to asking students their feelings about asking for help, we need to know whether or not they take advantage of opportunities for help (Ames and Lau, 1982. In the case of attitudes toward help-seeking behavior, then, we can see the need for an additional component, e.g., the conative component refers to the intention to actually seek help given the positive evaluation of help).

Ryan and Pintrich attempted to measure the mediating properties of attitudes toward help-seeking behavior and the actual behavior; in “Should I Ask for Help: The Role of Motivation and Attitudes in Adolescents’ Help-Seeking in Math Class,” in which a path model is explicated that allows for this conative component. By first differentiating between intrinsic and extrinsic motivations towards learning, direct and indirect effects on help-seeking behavior were compared. While attitudes toward help seeking were as predictive as the motivational variables with regard to students’ self-reported help seeking, there are several interesting dilemmas that emerged in the course of their analysis. Since the motivations had to be defined in such a way as to differentiate between students under the vulnerability hypothesis, intrinsic and extrinsic orientations to learning were ultimately based on perceptions of cognitive and social competence. Upon closer examination, it appears possible that such perceptions are based on attributions of causality, “this hypothesis proposes that students with high perceptions of cognitive competence do not attribute their need for help to lack of ability, and, thus are more likely to seek help when they need it,” (Newman, 1990).

For present purposes, we will need to revise the path model first articulated by Newman and used by Ryan & Pintrich (as well as by Butler and Neuman, 1995; and Ames and Lau, 1982) to posit a sequence from attribution through attitudes and intentions to behavior.

This revision poses no great obstacle until it is tested in reverse, that is, do experiences with helping behavior strengthen or modify attitudes (toward help seeking in particular, and perhaps toward help in general); is it credible to think these attitudes then inform or feedback on attributions and motivations? If such is the case, is it possible that attitude-behavior links are concurrent with attributions (as opposed to sequential, as suggested by the path model) and that intentions concerning future behavior are therefore as affected by attributions as by attitudes? Perhaps the obfuscating potential of such a model can be seen should we include attitudes toward intentions (to seek help), attributions about the cause of an intention to ask for help, different students' inclinations toward reticence, etc. The point may well be that processes occurring in each of these components (attitudes, attributions, intentions, and behavior) extend into one another, and it might prove difficult to tease apart the effects of each when they are concurrent.

A review of "An Attributional Analysis of Student Help-Seeking in Academic Settings" by Russell Ames and Sing Lau might help us to understand the possibilities for reciprocal causal linkages such as these as well as enable us to differentiate by gender in this path model. With an attribution analysis, students' beliefs about the causes of success and failure are the primary link to achievement-related behaviors (and clarify the vulnerability hypothesis).

The act of academic help seeking may have a number of different subjective meanings in the assessment of one's own ability. Some students may view help seeking as a public admission they have

failed in the past, and thus, they may avoid getting help in order to maintain esteem among peers. Seeking help, however, is also an act of effort, that is, the student is actively using available resources in order to increase the probability of success in the future. However, since failing when he or she has tried is a strong indicator of low ability, the student is risking an even greater loss of esteem if he or she should fail a second time (Ames and Lau, 1982).

One of the intriguing aspects of this study concerns the use of information about help sessions, (e.g., students were either informed that help sessions had been beneficial to students the previous semester or were simply told that help sessions were available). The analysis separated students by level of prior performance (on an examination) such that help-relevant and help-irrelevant attributions might be made. Help-relevant configurations involved the beliefs that one had adequate ability, that effort increased chances of success, and that one did not invest sufficient effort on the prior test; external factors such as luck were not seen to inhibit performance. Help-irrelevant configurations involved the beliefs that the instructor did not motivate the student, that questions were tricky, etc.; internal factors such as ability and effort were seen as unlikely to matter in subsequent assessments.

Gender differences were observed in two of these variables: low-performing females who made help-relevant attributions indicated a greater expected benefit from help sessions than either low-performing males or high-performing students of both genders. However, females were more likely than males to make help-irrelevant attributions (66% made task-

irrelevant attributions while males were split 50/50) such that Attribution by Sex effects were reported: 58% of low-performing males attended the help session compared to 43% of low-performing females, (i.e. the relevant vs. irrelevant attribution was an important predictor of whether or not a low-performing student attended or not, and males attended in higher numbers than females).

The extent to which such help-seeking behavior, including gender effects, on the part of college students generalizes to sixth grader students is arguable. However, it is perhaps useful to note that gender effects with regard to the diagnosis and treatment for Attention Deficit Hyperactivity Disorder (ADHD), which pertains to an alternate explanation for student noncompliance, are widely recognized for the age group under consideration.

Intervention Fidelity viz. Accurate Implementation

It might be useful to note that the Oxford English Dictionary traces the earliest use of *fidelity* (viz. fealty) to 1508, when its meaning pertained to matters of loyalty or oaths of allegiance. It is difficult to locate the first use the term with regard to intervention fidelity, though other uses of *fidelity* were found in the literature on therapeutic models, (e.g., aversion fidelity). The earliest use of the term in the educational psychology literature, as found by the present author, occurred in 1979, in an article by Penelope Peterson.

Given the earlier use of fidelity for therapeutic models, it's perhaps not surprising that we find that the National Institute of Health (NIH) currently sponsors a *Treatment Fidelity Workgroup*, embedded in its *Behavior Change Consortium*, which utilizes the following definition: "Treatment fidelity refers to the methodological strategies used to

monitor and enhance the reliability and validity of behavioral interventions. In a recent article, published on its behalf, we find, “It also refers to the methodological practices used to ensure that a research study reliably and validly tests a clinical intervention.”

Since many of the protocols it systemizes to ensure treatment fidelity are confined to clinical practice, (i.e., to health behavior practices) and might not generalize to other fields of study, (e.g. educational settings and practices), it seems useful to consider the implications that intervention fidelity might have viz. reliable and valid interpretations of a curriculum intervention. This is not intended to differentiate between treatment and intervention fidelity; its purpose is to reap commonalities that fit within a cost-analytic framework vis-à-vis a transactional approach (to the assessment of intervention efficacy).

Additional cross-disciplinary augmentation can be found in the use of stakeholder approach to program changes in organizations, where the emphasis is directed at communication, (i.e., transactions) among stakeholders viz. implementation). Laurie K. Lewis notes that stakeholder theory “provides a view of organizations – their internal and external relationships with individual and organizational “stakeholders – and provides both researchers and practitioners with a framework to assess those relationships.” This approach lends itself to both the cost-analytic techniques, as well as to the transactional approach used for this post hoc analysis. Lewis elaborates the form such assessment of relationships might take:

All scholars making use of the stakeholder perspective address issues of how organizational decision makers allocate stakes to stakeholder groups...These authors argue that organizational leaders

have a clear and immediate requirement to focus their resources on definitive stakeholders' needs (Lewis, 2007).

Moreover, Lewis has identified a problem with regard to communications studies (directed at processes of implementation of change in organizations), which is similar to the problem found with regard to the NIH attempts to structure treatment fidelity, (e.g., each is reductionist to the extent that empirical efforts address one or another component of change, such as implementation or efficacy, at the expense of a theoretical integration of how or why such components interact. Lewis argues that, "The time has come to launch a more theoretically holistic perspective geared toward accounting for communication processes and outcomes in change implementation that can be tested, altered, and developed."

Since this analysis directs attention to concerns with accuracy of intervention, (i.e., as opposed to generalization of its findings to similar students), many of the NIH recommendations concerning intervention fidelity are tangential. The focus on internal validity seems appropriate insofar as the stance of this secondary analysis is toward the GEM data as a pilot study.

Conclusion

The literature reviewed in this chapter has revealed a number of potential costs viz. transactions among stakeholders, some of which appear as risks while others emerge in the disparity between estimates and outcomes; each of which pertains to damage control with regard to significance testing, (e.g., effect size), and assessment of intervention efficacy

Chapter 3

ANALYSIS AND RESULTS

The research design is directed at the relations of the student compliance measures to global math attitudes utilizing data from the math attitudes questionnaire, (MATT); as well as to each other. Description of the characteristics of these measurements is included in the following presentation of the research hypotheses.

Hypotheses

The first working hypothesis is that students who complied with instructions by signing his/her name at the top of the MATT questionnaire demonstrated more favorable attitudes towards mathematics than did students who did not comply. The null version of this hypothesis is that there are no differences among students' math attitudes vis-à-vis whether or not they complied with instructions for survey completion, suggesting the possibility that teacher supervision adequately explains student compliance.

The second working hypothesis is that student compliance with instructions to report a single favorite subject is more common among students who identified themselves. Its null version is that there are no differences among students who did or did not identify a single

favorite subject vis-à-vis whether or not they identified themselves at the top of the questionnaire. Aside from teacher supervision of survey completion, the word ‘favorite’ was not interpreted to mean one and only one subject, (i.e., calling the appropriateness of the survey design for its intended subjects into question).

To test these hypotheses, student compliance has been operationally defined as a dichotomous variable for each form of compliance, (e.g., Identity Compliance [IDC] and Favorite Subject Compliance [FSC]); as such, each variable is treated at the ordinal level of measurement. The measure of a (global) student attitude towards mathematics is in the form of a response to a single item on the MATT questionnaire (MLIKE1), which asks students to select a response to the question, “In general, how much do you like math?” Available responses ranged from, *Not at all*, *A little*, *Some*, to *A lot*. Thus, MLIKE1 is also at the ordinal level of measurement.

Ordinal to ordinal comparisons generally require nonparametric statistical models, as in the case of the present analysis. Recommendations for additional parametric analyses, which might be used with the measure of student compliance, are included in the next chapter.

Data Analyses

Descriptive statistics for this cohort include a sample size of seventy nine students; forty seven students complied with instructions by identifying themselves and thirty two did not. Sixty six students complied with instructions to select a single, favorite subject and thirteen did not.

Chi square tests were conducted to test both hypotheses. Table 1 contains the row and column counts for the tabulation of MLIK1 and IDC, Figure 1 displays a histogram, and Table 2 reports the results of the test.

		MLIKE1				
IDC		1	2	3	4	Total
	Count	5	10	9	8	32
	Count	4	12	20	11	47
Total	Count	9	22	29	19	79

Table 3-1. Row and column counts for MLIK1 and IDC.

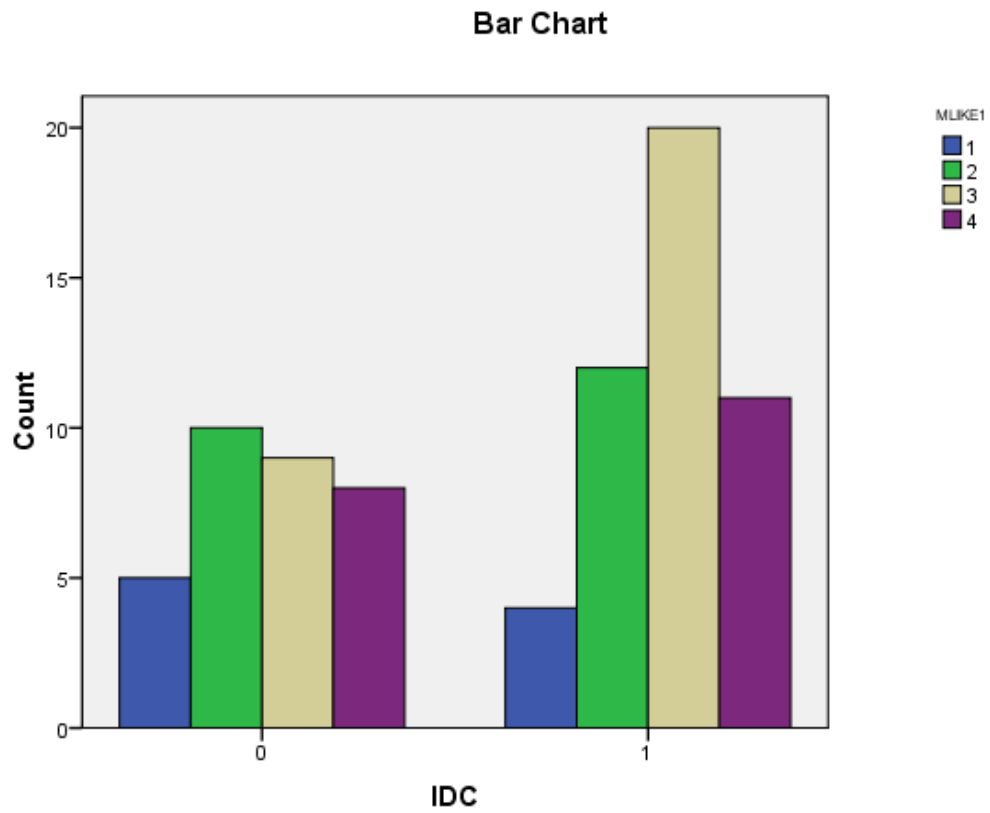


Figure 3-1. Histogram of MLIKE1 and IDC.

	Value	df	Probability
Pearson Chi-Square	2.169 ^a	3	.543
N of Valid Cases	79		

Table 3-2. Chi Square tests of the relation between MLIKE1 and IDC.

Turning to the second working hypothesis that student compliance with instructions to report a single favorite subject is more common among students who identified

themselves, the population size remains at seventy nine students, (i.e., there is no missing data). Table 3 contains the row and column counts for IDC and FSC, Figure 2 a histogram, and Table 4 displays the results of the chi square test.

			FSC		
			0	1	Total
IDC	0	Count	8	24	32
	1	Count	5	42	47
	Total	Count	13	66	79

Table 3-3. Row and column counts for IDC and FSC.

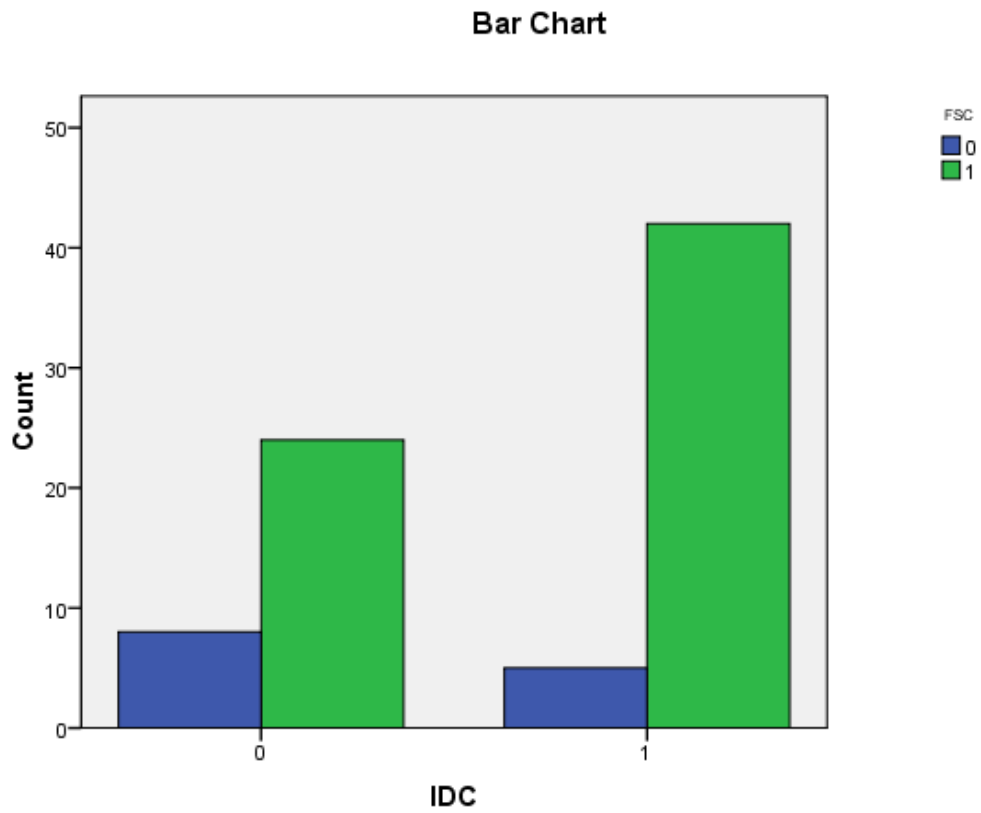


Figure 3-2. Histogram of FSC and IDC.

	Value	df	Probability
Pearson Chi-Square	2.856 ^a	1	.124
N of Valid Cases	79		

Table 3-4. Chi square test of the association between FSC and IDC.

Conclusion

In the next chapter, discussion of these results is accompanied by recommendations for additional analyses.

Chapter 4

DISCUSSION OF RESULTS

As a group, the students involved in this study had demonstrated a statistically significant improvement ($p=0.01$) in pre- vs. post-intervention math achievement scores, (e.g., the P.S.S.A. tests) at the end of the GEM curriculum intervention. The effect size extrapolated from the matched-pairs t tests was reported at 0.409, (i.e., it was small to medium in magnitude). This secondary analysis sought to differentiate these students according to proxies for engagement, interest, and/or motivation with regard to learning mathematics.

The first hypothesis is that students who complied with instructions by signing their name at the top of the MATT questionnaire demonstrated more favorable attitudes towards mathematics than did students who did not comply. Pearson's Chi Square test provides insufficient evidence ($p= 0.543$) to reject the null version of this hypothesis, (i.e., that there are no differences among students' math attitudes vis-à-vis whether or not they complied with instructions for survey completion). The second hypothesis is that student compliance with instructions to report a single favorite subject would be more common among students who identified themselves. Pearson's Chi Square test provides insufficient evidence ($p= 0.124$) to reject the null version of this hypothesis, (i.e., that there are no differences among students who did or did not identify a single favorite subject vis-à-vis whether or not they identified themselves at the top of the questionnaire).

Failure to reject the first hypothesis suggests that teacher supervision might adequately explain student compliance with regard to identity compliance; that is, had teachers scrutinized the MATT upon completion, this source of error might well have been diminished if not eliminated. In other words, to the extent that compliance was at issue with regard to accurate implementation of the research design, evidence suggests that it might have involved noncompliance to protocols on the part of teachers rather than students.

Though this analysis of the MATT is concerned with accuracy of implementation, two threats to external validity bear mention. The cancellation of classes due to flooding in the aftermath of Hurricane Ivan necessitated a delayed administration of the MATT, (i.e. history interacted with administration of the assessment), and, the MATT was administered at the beginning of the following school week, during homeroom period, as opposed to during math class. It is possible these effects might have interacted. For example, students might have been atypically rambunctious or inattentive after returning to school following such a disaster. Moreover, no information was available concerning when the MATT was administered, across the four separate grade six classrooms, among other tasks that occurred during that homeroom period.

Rather than test the second form of compliance against students' global math attitudes, (i.e., no correlation would be expected with the exception of students selecting mathematics as his/her favorite subject), it seemed reasonable to assess the extent to which a pattern of noncompliance might be detected. Failure to reject the null form of the second hypothesis suggests that such a pattern is not discernible.

The lack of a pattern regarding instances of noncompliance is interesting, suggesting that student inattentiveness to detail rather than reading comprehension might explain this

disparity. This in no way diminishes the plausibility of the alternate explanation that at least some of the students might not have interpreted favorite to preclude more than a single answer.

Recommendations

Teacher training with regard to administration of the MATT did not sufficiently emphasize the importance of student identification. Regardless of the extent of teacher commitment to the GEM curriculum intervention, or to managing a variety of administrative tasks during homeroom period, it is apparent that greater diligence in detecting unsigned questionnaires could have helped with this source of missing data. Of course, given the failure to reject the null hypothesis (that identification did not predict global math attitudes very well); the pertinent question is, ‘Why does it matter?’.

Students exposed to the GEM curriculum intervention completed the MATT questionnaire at the end of grade six, (i.e., the end of project year one), and at the end of grade seven, with the conclusion of the intervention. While not pertinent to the present analysis, comparisons of students’ pre- and post-intervention math attitudes are made problematic by this lack of (initial) identification. This problem is exacerbated by the attrition rate of students in the Midtown School District, (e.g., fifty eight of the original seventy nine students were still enrolled at the time of project completion).

Given this attrition rate, which was largely attributable to movement to and from nearby school districts, the importance of complete data gives rise to an additional recommendation. Investigators need to make provisions that a student enrolled in grade six

during the course of the school year also be administered the MATT questionnaire, that is, prior to immersion in the curriculum intervention, which likely differed from the grade six mathematics curriculum in her former school. Furthermore, efforts should be made to administer the MATT questionnaire at the beginning of grade seven to students new to the school district, in the interest of detecting patterns of attitude change with regard to mathematics.

The MATT questionnaire had been designed in such a way as to allow maximum flexibility for students to report his/her favorite class by inclusion of a write-in response, if the appropriate response was not among those on the checklist. This option may have exacerbated problems with interpretation of *favorite* to mean a single response meaning; it surely resulted in less-than-useful data for the analysis of intervention efficacy. For example, the modal write-in response was physical education (aside from whether it was accompanied by one of the academic subjects on the checklist or not).

A forced choice checklist that included all sixth grade subjects is a compromise worth considering. For instance, problems occurring due to write-in responses that included reading or spelling, regardless of whether or not such responses accompanied a check on English as her favorite subject, could have been avoided. Ultimately, this decision must be made according to the importance assigned to knowing whether or not a student's favorite subject changed from grade six to grade seven, (i.e., the MATT questionnaire had been designed for students in grades six through eight and there is some variability in actual classes).

Conclusion

The limitations of nonparametric statistical techniques are at issue with regard to additional research using the GEM data. A second measure of math-related attitudes and beliefs is available in the form of a summated rating based on sixteen items in the MATT. This scale, which arguably falls somewhere between an ordinal and interval level of measurement, might be more suited to the analysis of the relation of math attitude to math achievement, (e.g., the PSSA is an interval level of measurement).

Since the purpose of the GEM curriculum intervention was improved math attitudes and achievement, it is important to consider the additional insights that might be obtained by including the ordinal measures of compliance alongside of the interval measures of math achievement. The pre-intervention PSSA (at the end of grade five) might provide valuable information concerning students' prior math or reading achievement vis-à-vis compliance with instructions.

Accurate implementation, even in the face of changes in the research design during the course of an intervention, depends on a consensus among investigators, administrators, teachers, and students. Such a consensus is more difficult when participants are not in physical proximity, as was the case with the GEM intervention. It is clear that more frequent contact on the part of investigators could have addressed issues that occurred during the course of implementation, such as teacher reassignments and attrition. Insofar as the GEM project is regarded as a pilot study, there is considerable room for improved efficacy with regard to its assessments.

Stakeholder theory might be useful in tracking (some of) these pitfalls, and especially concerning the development and maintenance of a consensus, (e.g., the afore-mentioned dimensions of power, legitimacy, and urgency for stakeholder theory). Whether a transactional approach to classroom compliance can augment such structural concerns remains a matter of future research.

REFERENCES

- Ames, R. & Sing Lau. (1982). An attributional analysis of student help-seeking in academic settings. *Journal of Educational Psychology*, 74(3), 414-423.
- Bellig, A. J. et al. (2004). Enhancing treatment fidelity in health behavior change Studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology*, 23(5), 452-456.
- Blazina, C. & Watkins. C. E. (1996). Masculine gender role conflict: Effects on college men's psychological well-being, chemical substance usage, and attitudes towards help-seeking. *Journal of Counseling Psychology*, 43(4), 461-465.
- Butler, R. & Neuman, O. (1995). Effects of task and ego achievement goals on help-seeking behaviors and attitudes. *Journal of Educational Psychology*, 87(2), 261-271.
- Cothran, D.J. & Ennis, C.D. (1997). Students' and teachers perceptions of conflict and power. *Testing and Teaching Education*, 13(5), 541-553.
- Daubman, K. & Lehman, T.C. (1993). The effects of receiving help: Gender differences in motivation and performance. *Sex Roles*, 28(11/12), 693-707.
- Feather, N. T. & Simon, J. G. (1971). Attribution of responsibility and valence of outcome in relation to initial confidence and success and failure of self and other. *Journal of Personality and Social Psychology*, 18(2), 173-188.
- Graham, S. & Barker, G. P. (1990). The down side of help: An attributional-developmental analysis of helping behavior as a low-ability cue. *Journal of Educational Psychology*, 82(1), 7-14.
- Hetherington, J. (2000). Role of theory and experimental design in multivariate analysis and mathematical modeling. In *Handbook of Multivariate Statistics and Mathematical Modeling*. Academic Press: Harcourt.
- Levanthal, H. & Friedman, M.A. (2004). Does establishing treatment fidelity help in understanding treatment efficacy?: Comment on Bellig et al. *Health Psychology*, 23(5), 452-456.
- Lewis, L. K. (2007). An organizational stakeholder model of change implementation communication. *Communication Theory*, 17, 176-204.
- Good, G.E., Dell, D.M., & Mintz, L.B. (1989). Male role and gender role conflict: Relations to help seeking in men. *Journal of Counseling Psychology*, 36, 295-300.
- Newman, R. S. (1991). Goals and self-regulated learning: What motivates children to seek academic help? *Advances in Motivation and Achievement: A Research Annual*, 6,
- Newman, R. S., Murray, B. & Lussier. C. (2001). Confrontation with aggressive peers at school: students' reluctance to seek help from the teacher. *Journal of Educational Psychology*, 93(2), 398-410.

- Newman, R.S. (1990). Children's help seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology*, 82(1), 71-80.
- Newman, R. S. & Goldin, L. (1990). Children's reluctance to seek help with schoolwork. *Journal of Educational Psychology*, 82(1), 92-100.
- Noboru Komiya, N., Good, G. E. & Sherrod, N. (2000). Emotional openness as a predictor of college students' attitudes toward seeking psychological help. *Journal of Counseling Psychology*, 47(1), 138-143.
- Peterson, P. (1979) Aptitude X treatment Interaction effects of teacher structuring and Student Participation in college instruction *Journal of Educational Psychology*, Vol. 77, No. 4, 521-523
- Popham, J. W. (1988). Educational Evaluation, 2nd ed. Englewood Cliffs: Prentice Hall
- Ryan, A. M. & Pintrich, P.R. (1997). Should I ask for help?: The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of Educational Psychology*, 89(2), 329-341.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71(1), 3-25.
- .