University of Alberta

An Experimental Test of the Learning Errors and Formative Feedback (LEAFF) Model: Creating Positive Learning and Assessment Environments for Students

by

Maria Clara Bustos Gomez

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Education

in

Measurement, Evaluation and Cognition

Department of Educational Psychology

Fall 2013
Edmonton, Alberta

To my three beloved ones

ABSTRACT

The Learning Errors and Formative Feedback [LEAFF] model is a three-part learning and assessment framework that is designed to create safe learning and assessment environments for students. This paper reports an experimental test of the first part of the model, namely, a manipulation of an instructional variable expected to alter the *Instructional Climate* for students. Specifically, a Learning Errors Intervention [LEI] was embedded in a one-hour Elementary Statistics lecture and was treated as a between-group variable. One hundred and one undergraduate students were randomly assigned to either the LEI (n=51) condition or a control condition (n=50). Statistical analyses, conducted on cognitive and affective measures collected during and at the end of the lecture, supported predictions derived from the LEAFF model on the positive effect of the LEI on students' feelings of safety, trust in the instructor, and increased likelihood of indicating areas of confusion with the content.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Emotions and motivation play a significant role in student learning and assessment by influencing the interpretations students form about their classroom environments (Bandura, 1993; Brookhart, Walsh, & Zientarski, 2006; Church, Elliot & Gable, 2001; Pekrun, 2011; Pintrich et al., 1991; Schunk, 1991). For example, a student who feels bored may be unmotivated to pay attention to a lesson on fractions or a student who lacks trust in the teacher may be unwilling to risk expressing an innovative answer to a problem lest he or she is prepared for some ridicule. The crippling test anxiety some students experience is well documented in the literature (Cassady, 2010; Huberty, 2009; Sarason, 1980; Zeidner, 1998, 2007), origins of which are sometimes rooted in negative classroom experiences (Cizek & Burg, 2006). For example, Cizek and Burg (2006) found that student test anxiety is most strongly correlated with teacher-manifested anxiety in comparison to any other variable.

Although student learning and, especially, student assessment are often viewed and examined from a cognitive perspective (Mislevy, 2006), typically focusing on how content knowledge and skills are shaped and evaluated in students, greater study is required of the emotional and motivational variables that influence how to engage students to learn and use assessment feedback (Leighton, Chu & Seitz, 2012). For example, in the absence of student motivation, it is difficult to engage in the deliberate practice required to master complex concepts. Likewise, in the absence of interpersonal trust it is difficult to imagine students and teachers collaborating towards mutually constructive assessment goals.

Teachers may administer assessments, even formative ones, but students may wonder and question the intent of the assessment as well as the value of the feedback (Black & Williams, 1998; Sadler, 2006).

Among the many emotional variables to consider in the student-teacher interactive relationship, perhaps none is more important than trust (Seitz, Chu, Bustos & Leighton, 2012); in fact, when trust is identified in the educational research literature it is mainly presented and discussed as a keystone of effective schools and collaboration among adults. To achieve high-level learning and assessment outcomes, students need to feel safe in their classroom environments as learning involves moving from a state of not knowing to a state of knowing, expressing novel ideas based on recently acquired knowledge, making mistakes as newly learned skills are implemented, and developing the confidence and self-efficacy to accept formative feedback from teachers (Clark, 2011; Shute, 2008). Indeed, an association is often reported between schools that have made significant learning gains and the presence of trusting relationships within the administrative, teaching and learning communities (Sebring & Bryk, 2000; see also for a review Forsyth, Adams, & Hoy, 2011). In schools where trust and cooperation are claimed to exist, students have also reported feeling safer, sensing that teachers care about them, and experiencing greater academic achievement. However, the majority of the educational research on the relationship between trust and learning (Fosyth et al., 2011) is correlational and not experimental, thereby barring causal conclusions to be made. Further, almost no research has been conducted on the relationship between trust and assessment (however, see Hirschfeld & Brown, 2009; Chu, Guo & Leighton, 2013).

To address the goal of providing students with learning and assessment experiences that effectively meet their needs along both cognitive and affective dimensions, teaching practices must become more attuned to students emotions and motivations in learning experiences. For example, Kochanek (2005) suggests a developmental approach to school leaders who aim to build trust – finding ways to communicate a vision, promoting low-risk exchanges through small group interactions, using interactions to ease feelings of vulnerabilities, and then creating opportunities for higher-risk interactions that lead to positive outcomes. Through these repeated exchanges, Kochanek claims that teachers can build confidence in themselves and with others to develop more trusting relationships. Although Kochanek (2005) recommends specific strategies, these strategies are aimed primarily towards the development of trust among teachers and school administrative staff and not necessarily with students. In fact, there is a general lack of experimental, evidence-based recommendations in the research literature for creating trusting student-teacher relationships in the classroom and also for repairing environments already plagued by a lack of trust (Louis, 2008).

Recently, a theoretical model – the learning errors and formative feedback (LEAFF) – has been proposed to help identify the causal inter-dependence between cognitive and emotional variables in students' learning and assessment classroom contexts (Leighton et al., 2012). The objective of this study is to report the results of an experimental test of the LEAFF model, namely, the effect of discussing “learning errors” in a Statistics class (lecture) to help generate stronger feelings of safety among learners, enhance their trust in the instructor or teacher, and impact positively learning and achievement outcomes. The balance of this paper is divided into four parts. First, we provide a brief literature review focusing on the LEAFF model and the cognitive and affective variables predicted to influence students’ feelings of safety and trust in the classroom, and learning outcomes. Second, we describe the experimental design and methods used to investigate one of the principal claims of the LEAFF model; that is, that discussion of learning errors with students will lead to stronger, more positive feelings of safety and trust, and will enhance learning outcomes. Third, we present the results from the study and, fourth, we discuss and elaborate on the implications of the findings.

CHAPTER 2: LITERATURE REVIEW

LEARNING ERRORS AND FORMATIVE FEEDBACK [LEAFF]

This section describes the LEAFF model shown in Figure 1. For a comprehensive introduction and explanation of the model, the reader is referred to Leighton, Chu, and Seitz (2012). The LEAFF model is also described in relation to the construct of trust in Seitz, Chu, Bustos, and Leighton (2012). The LEAFF model was developed from a review and synthesis of the emotional, psychological, and social factors that underlie successful learning, instruction and assessment (e.g., Black & Wiliam, 2009, Boekaerts & Corno, 2005; Frijda & Mesquita, 1995; Kluger & DeNisi, 1996; Leighton, 2009; Leighton & Gierl, 2007, 2011; Shute, 2008).

In Seitz et al. (2012), key aspects of the model were highlighted in relation to the foundational role that trust plays in helping to create learning environments where students feel at ease showing vulnerability – that is, revealing their learning errors, mistakes or misconceptions as they aim to learn in meaningful ways. The present paper empirically tests the idea originally put forward in Leighton et al. (2011) and elaborated in Seitz et al. (2012); namely, that discussing learning errors can help create a safe instructional or learning environment for students. The LEAFF model recognizes the importance of the teacher's words and actions in creating a safe learning environment for students, one that influences the development of students' mental models of learning. There are three distinct parts in the LEAFF model: the instructional environment; students' mental models of learning, that is, the affective and cognitive aspects influenced by the classroom learning environment; and students' performance.

Part One: Instructional Climate/Environment

As shown in Figure 2, the first part of the LEAFF model is most pertinent to the present paper as it focuses on the instructional environment of the classroom. In this first part of the model, the teacher's words and actions create either a positive or a negative learning climate. Depending on the learning climate created, students are expected to experience different thoughts and emotions about and for learning. For example, in a safe learning environment, students are expected to feel secure taking risks, making mistakes as they learn, and showing receptivity for formative feedback. In an unsafe learning environment, students are expected to feel insecure taking risks, making mistakes as they learn, and showing resistance to formative feedback. In a safe learning environment, trusting relationships are observed between students and teachers.

In order to develop trusting student-teacher relationships, the LEAFF model indicates that teachers should use a variety of explicit verbal and physical communication strategies (Seitz et al., 2012). For example, during instruction, teachers can verbally *identify* the value of making errors as part of the learning process. Further, they can *explain* that it is through the understanding and correction of these errors, that student-learning gaps are addressed, leading to higher-order thinking and the advancement of learning. Teachers can *show* students common errors that students often make as they acquire knowledge and skills in a particular content domain. Teachers can also *demonstrate* to students that teacher-led attempts to correct errors are not always successful and repeated attempts are necessary for success. It is proposed that teachers explicitly discuss and show students the pedagogical value of errors, including as part of the discussion, how understanding the origins of errors offer opportunities for deeper learning.

The timing for engaging in discussions of learning errors should happen from the first encounter that a teacher has with students, and it is expected that starting these interactions early will create opportunities to develop students' positive feelings of safety in the classroom and trust in the teacher. Although it might seem unlikely for the *seed* of a trusting relationship to develop quickly between individuals, research in counselling psychology indicates otherwise. For example, the *therapeutic alliance*–that is, the positive or negative relationship created between therapist and client—can be established as soon as the first to third session (within three hours; see Horvath, 2000; Horvath & Symonds, 1991). A positive or negative therapeutic alliance can be quickly established as it is emotionally evaluated by the client seeking treatment as early as the first session (Horvath & Symonds, 1991) and predicts the success/effectiveness of the therapy (Horvath, 2000).

Discussing learning errors as a strategy for building safe learning environments is based on two facets of trust – openness and honesty (see Hoy & Tschannen-Moran, 1999; Tschannen-Moran, 2003). Having open and honest conversations about learning errors is expected to lead students to feel increased comfort in showing teachers what they do not know or what they are struggling to understand. As a result, students and teachers can feel more at ease showing vulnerability in a relationship where seeking and providing support, respectively, for learning is contingent on mutual understanding of the complexity of the learning process (see Burke et al., 2007; Mayer & Davis, 1999). The default in many classrooms is often to have few deliberate discussions of (a) the complexity of what it means to learn and (b) the necessity of errors for clarifying where students are at in the learning progression. Additional strategies that are expected to build trust are presented in the LEAFF model (see Seitz et al., 2012). In addition to talking about the pedagogical value of learning errors, teachers could also tackle topics on the role of effort, intelligence, and motivation in helping students gain a deeper understanding and comfort with the many factors that influence learning. These discussions would not be expected to take place only at the beginning of a lesson plan but during other key time points – such as when formative assessments are administered, during feedback related to students' performance on formative assessments, and also before and after summative assessments.

Although teachers generally agree that making mistakes is an expected part of learning and they are receptive to their students making mistakes, there is often no explicit discussion with students about the importance of mistakes in the learning process (see Leighton et al., 2012). In addition, teachers may not be cognizant of how to initiate these conversations and help to formally establish an atmosphere of openness, honesty, and safety for students (Leighton et al., 2012). Students' attitudes and beliefs need to be addressed directly and overtly in the classroom in order to positively influence their development (Boekaerts & Corno, 2005; Bransford et al., 2000; Lajoie, 2008), and the main person responsible for making this possible is the teacher (Bandura, 1993). For example, Boekaerts and Corno (2005) indicate that student participation in classroom activities is influenced by teacher practices. These practices can either bolster or emasculate students' thoughts and feelings about learning. In addition, studies by Schunk (1989), and Ashton and Webb (1986) showed that teacher-led instructional programs designed to enhance academic self-efficacy were an important factor in building children's beliefs in their intellectual capabilities.

Part Two: Mental Models of Learning

The second part of the LEAFF model, shown in Figure 3, focuses on the mental models of learning that students create in response to their learning environments. According to Johnson-Laird (2004), who has gathered significant evidence on the occurrence of mental models, individuals base their reasoning, problem solving and behaviours on the representations or "models" they generate of the world around them. It is expected that students base their attitudes, beliefs, and actions about learning, in part, upon the models they generate of what they observe and hear in their classroom environments (Leighton & Sternberg, 2012). As illustrated in the middle box of Figure 3, mental models of learning include two parts – a cognitive part related to the knowledge and skills learned in the instructional context, and an affective part related to the emotive, motivational and expressive climate of the instruction (Leighton & Sternberg, 2012). For example, on the cognitive front, a student who observes a teacher repeatedly assign math homework may add to his or her model of learning the directive that math requires practicing skills. In addition, on the affective front, a student who observes a teacher constantly praise another student with words such as "you always get the right answer" may supplement his or her model of learning with the directive that mistakes are to be avoided because immediate right answers are desirable for receiving recognition. Likewise, a student who sees a teacher dismiss or not take ownership of teacher-made instructional mistakes may extend his or her model of learning with the rule that mistakes are not only to be avoided but also denied when they occur. Frijda and Mesquita (1995) explain that almost all environmental events are judged to be emotionally relevant or irrelevant, positive or damaging to a particular purpose or concern.

A teacher who deliberately attempts to create a safe classroom environment by ensuring trusting student-teacher relationships can be expected to increase the emotional positivity of students' mental models and, thus, their cognitive receptivity for learning. In a safe classroom environment, students would be expected to make more mistakes, initially, as they learn content material because they would allow themselves to be more vulnerable in front of a teacher who is explicit about the value of learning errors. In a safe learning environment, students would also be expected to be more receptive to formative feedback about their performances, show more innovative problem-solving strategies and greater mastery orientations toward learning. Kluger and DeNisi (1996) and Shute (2008) indicate that in order for feedback to be formative, the student must trust the agent who delivers the information and be willing to accept and utilize the feedback. In contrast, students who emotionally evaluate the instructional context to be unsafe would be expected to act in opposite ways to students who judge the instructional context to be safe. That is, students would be expected to avoid making mistakes, as they would not want to be vulnerable in front of a teacher who is unclear or unopened about the role of errors in the learning process. Further, in a learning environment perceived as unsafe, students would be expected to develop performance orientations toward learning rather than mastery orientations, and be less willing to engage in innovative thinking and to develop problem solving skills. If students experience trepidation to show a teacher what they know or do not know, they may fail to believe that a teacher's feedback about their performance is accurate. In their meta-analysis of feedback on performance, Harris and Rosenthal (1985) showed that the climate created for feedback had a greater effect on student performance than the quantity of feedback provided.

Part Three: Learning and Achievement Performance

The third part of the LEAFF model, shown in Figure 4, focuses on student performance. Specifically, this third part indicates that students' performance will be more transparent or informative of their actual learning levels when students feel safe showing any lack of understanding on assignments and assessments. If students view their performance on assessments, especially formative assessments, as opportunities to show what they truly do not understand and receive help, they are more likely to hear and receive formative feedback as a tool for improving learning. Conversely, when students feel unsafe to show their lack of understanding on assignments and assessments, they are less likely to pay attention to and receive formative feedback as they may not believe the assignment or assessment measured what they were struggling to learn.

In addition, several predictions are described in Leighton et al. (2012) about students' performance when they develop cognitively-strong and emotionally-positive mental models of the learning environment: students will show greater comfort with the learning process and environment; students will show more creativity, risk-taking (ingenuity) and higher-order thinking as they encounter new material and as they learn the material; students will initially make more learning errors on formative assessments; students will demonstrate higher intrinsic motivation in the content area and express higher trust in the teacher; and students will make fewer errors on summative assessments. The opposite is expected of students who develop emotionally negative mental models of the learning environment, irrespectively of the cognitive strength of their models. The LEAFF model indicates that the most important *input variable to* student learning and assessment is the safety of the learning environment, where *trust* is a key part of developing and maintaining the health of this environment. Without a positive learning environment, the mental models students create will direct them to avoid showing what they truly know (and do not know) for fear of making mistakes. By not being transparent in their learning, students will not trust the accuracy of the feedback delivered to them based on their performance on assessments. The objective we therefore address in the next section – methods – is defining and operationalizing trust in the creation of safe learning environments. The specific research question guiding this study is: *What is the effect of discussing "learning* *errors" in a Statistics class (lecture), which has been designed to help generate stronger feelings of safety among learners and enhance their trust in the instructor or teacher?* It is hypothesized that the discussion of learning errors will influence students' mental models in positive ways by increasing their comfort in the class and trust in the instructor. In turn, it is expected that influencing mental models in positive ways will increase students' likelihood of indicating areas of confusion with the presented material.

CHAPTER 3: METHOD

Participants

For the purpose of the study, 101 students enrolled in a Fall 2012 course on *Technology for Teaching and Learning Tools* offered as part of a Bachelor of Education program at a large research-intensive University participated as part of a research option. Participants voluntarily registered for a one-hour research session conducted extra to their classes; and they did not receive any compensation for their participation in the study. Of the 101 participants, 75 self-reported their gender as female with a mean age of 23.6 years (SD = 4.45 years) and 26 students self-reported their gender as male with a mean age of 23 years (SD = 6.18 years). Most of the participating students self-reported their ethnicity as White or Caucasian (81.3%) with the next highest self-reported ethnicity being South Asian (10%) followed by Chinese and Filipino (5%).

Design and Materials

The present study involved a key between-subject independent variable, namely, an intervention employing a targeted discussion on learning errors within an Elementary Statistics class/lecture. Before describing in greater detail the instructional and experimental design presented in Figure 5, it essentially involved two conditions: (a) presence of a targeted discussion on learning errors within an Elementary Statistics class (Learning Errors Intervention or LEI) or (b) absence of a targeted discussion on learning errors within an Elementary Statistics class (no-LEI or control). Both conditions were designed to be exactly the same in all respects (content and procedure) except for the presence or absence of the discussion on learning errors. In addition, previous to any data collection, the researchers repeatedly rehearsed the lecture delivery and the administration of all measures in order to reduce any discrepancies in instruction between the two condition groups. These practice sessions occurred without participants until the delivery of the lectures were identical with the exception of the manipulated variable – discussion of learning errors.

A research assistant (author of this paper) served as the instructor to deliver the content of the 40-minute Elementary Statistics class (see Appendix A) and administered the different measures used during the study. To account for extraneous administration factors that could influence the interpretation of results, the one-hour LEI and control conditions were scheduled systematically on similar days throughout the morning and afternoon with equal frequency. Participants were randomly assigned to either the control (n=50) or the LEI conditions (n=51).

A set of either 15 or 18 PowerPoint presentation slides, depending on the condition, was used to deliver the lecture in the present study. After welcoming the participants to the study, all participants were provided with a package of information, which included (1) 15 PowerPoint slides (control) or 18 PowerPoint slides (LEI), (2) a 10-item multiple choice assessment, (3) nine scales or surveys (including a *Feeling Scale* [FS]), and (4) a background questionnaire. During the lecture, the discussion of learning errors – which was embedded in three extra PowerPoint slides in the LEI condition (see Figures 6, 7, and 8) – was administered to only those participants who had been randomly assigned to this condition.

During the lecture, participants in both conditions were asked approximately every 4 minutes (10 times during the lecture) to indicate how they were feeling using the *Feeling Scale* [*FS*] developed by Hardy and Rejeski (1989) and modified for use in the present study. The FS, shown in Appendix B, is a rating scale with anchor points ranging from -5 (Very Bad) to +5 (Very Good). The scale allows individuals to express their feelings of comfort in the class. In addition, participants were asked to identify any difficult content material directly on the PowerPoint slides (they were provided at the start of the research session) by writing “CM” on the slides. It was explained to students that the “CM” they wrote would be used to indicate any content material on the PowerPoint slides that was confusing to understand and had the potential to lead students to make mistakes during their understanding and performance. Participants were asked to place the CM as closely as possible to the content on the slide that was confusing and problematic and they could use multiple “CMs” on any single slide.

Following the 40-minute lecture, the research assistant administered a 10-item multiple-choice assessment to students, then requested to students to swap assessment papers with other students for marking as the assistant provided the keyed answers orally. Then, the assistant asked students to return the assessments to their owners. The assessment portion of the class lasted 10 minutes. Following the assessment, the assistant administered 8 surveys and a background questionnaire at the end of the lecture, all of which required 10 minutes to complete. The 8 surveys measured affective states and traits. Finally, the background questionnaire that included questions about previous academic performance was administered. The assessment and surveys served as dependent variables in the present study. In the following section, the design of the present study is described along with the manipulation of learning error, followed by a description of each of the dependent measures.

Independent variable: Learning errors intervention [LEI]. A classroom intervention involving a targeted discussion of learning errors within an Elementary Statistics lecture was administered to half of participants (51 students - LEI). The other half of participants (50 students - control) did not receive the intervention but did receive the Elementary Statistics lecture. Both lectures were *designed to be exactly the same except for the LEI component that occurred in the first 5 minutes of the lecture and then again in the last 5 minutes* of the lecture before the assessment was administered to students.

Within the first 5 minutes of the lecture after welcoming and thanking students for their participation in the study, the LEI was operationalized in the 2nd, 3rd, and 16th slides of the PowerPoint presentation shown in Figures 6, 7, and 8. The full set of power point slides for the intervention and control condition is included in Appendix A. The research assistant who acted as the instructor in both conditions delivered the scripted lecture associated with each slide to ensure that all students received the same lecture and the LEI specifically in exactly the same way. For example, the script for slide 2 was as follows:

What is our rationale for this study?

Well, as you might have experienced in past classes, learning is a rewarding experience but it also can be risky.

Learning takes us from a state of NOT KNOWING to a state of COMING TO KNOW, and this complex process involves several elements such as making mistakes.

Yes, making mistakes is part of learning. Actually, psychologists tell us that –in most cases- making mistakes help us learn.

Why is that?

Well, mistakes help our brain clearly separate what is correct and incorrect. In the process of learning, being able to identify mistakes, where they can happen and talking about them can help us learn better.

You may recall an experience when you were learning something and made a mistake (or more than one) and this helped you learn that knowledge or skill really well; for example, when you were learning to tight your shoes or to drive a bike or a car.

The script for slide 3 was the following:

As I said, this class is about Statistics, and learning statistics involves making mistakes. Why is that?

Well, in Statistics there are many formulas and concepts and it is very easy to get confused and make a mistake. So, it is very important to recognize the presence or potential to make mistakes. During the lecture, there will be two important procedures I need your help with. For the first one, you have been provided with copies of the PPT slides, and I need you to help me identify specific places *where someone can get confused and make a mistake, so please write "CM" every time you identify those places where there is potential for mistakes.*

The second procedure is rating your feelings of comfort. We have developed an 11-point rating scale to help us know how you are feeling during the class. The scale goes from -5 meaning that you are feeling really bad, up to +5 meaning that you are feeling really good. You will be asked to rate your feelings at 10 different moments during the lecture and I have included check marks on the PowerPoint slides and the rating number so nobody gets lost. We will be using the set of ratings, and each time we finish one rating, we will flip over the page to get ready for the next rate.

The two procedures described during the presentation of slide 3 – asking participants to indicate with a "CM" where in their copy of the slides they found the content material to be confusing and potential for making mistakes, and rating their feelings during the lecture – were dependent variables that were also included in the control condition but without the supporting context of the importance of learning errors. For example, in the control condition, students were invited to indicate any places on the PowerPoint slides that might be confusing. These two procedures are described in great detail in the next section of this paper (Dependent Measures).

The script for slide 16 – the last slide containing intervention information – reminded participants in the LEI condition about the reality of making mistakes while learning, the importance of recognizing confusing areas during learning in order to correct them to improve assessment performance:

> *Remember learning Statistics requires making mistakes.*
>
> *Now we want to find out how well we conveyed the material to you.*
>
> *The assessment you will complete will help us confirm all the places "where someone might get confused and make a mistake"*

After the 40-minute lecture, participants responded to a 10-item multiple-choice assessment based on the information presented during the lecture (see Appendix C). This assessment served as one of the dependent measures used to evaluate the LEI and is described in greater detail in the next section. After the assessment, which took participants about 5 minutes to complete, participants were invited to trade their assessments with another person in the class for peer grading. During peer grading, the instructor announced the keyed responses, while participants graded the assessments of those with which they had traded. Peer grading took about 5 minutes to complete. During the final 10 minutes, participants responded to 8 surveys, measuring affective dimensions and one background questionnaire. These surveys are discussed in the next section.

Dependent measures. There were two types of dependent variables included in the present study – affective and cognitive. The affective variables involved 8 instruments or subscales plus the *Feeling Scale* (FS). First, the 11-point *Feeling Scale* (FS) was adapted for the present study so that participants could continuously rate their feelings of comfort during the lecture (see Table 1). Our FS scale was adapted based on Hardy and Rejeski's (1989) FS scale, which provides a unidimensional measure of individuals' experience of comfort during a situation. The FS scale in the present study was used in the control and LEI conditions at 10 different times during the lecture. Shown in Table 1 is the temporal alignment or correspondence between (a) the lecture's PowerPoint slide numbers and (b) when the instructor requested participants to rate their feelings of comfort (rating occasion). This alignment can also noted in the PowerPoint slides shown in Appendix 1. Rating occasions are shown in the bottom right-hand corner of the slides. For example, in the control condition, the first request to have participants rate their feelings using the FS occurred during the 2nd slide when the FS was introduced and this first rating allowed students to practice responding with a rating. In the intervention condition (LEI), the first request to have participants rate their feelings using the FS occurred during the 4th slide when the FS was introduced and this first rating allowed students to practice responding; in the intervention condition, the first rating took place on the 4th slide because the LEI was introduced in the 2nd and 3rd slides.

In addition to the FS, eight subscales designed to measure affective states and/or traits, shown in Appendix D, were administered to participants after the lecture and assessment. Of the eight, three subscales -*Intrinsic Goal Orientation, Extrinsic Goal Orientation, and Self-efficacy for Learning and Performance*- were taken from the *Motivated Strategies for Learning Questionnaire (MSLQ*) (Pintrich et al., 1991) and used to measure participants' motivational orientations and their use of different learning strategies. Participants responded to the instruments using a 7-point scale, ranging from 1-'Not at all true of me,' to 7-'Very true of me.' Internal consistency values for the original measures are reported in parenthesis. The *Intrinsic Goal Orientation* scale (α = .78) consisted of four items designed to measure participants' perceptions of the reasons why they engage in learning tasks, including, challenge, curiosity, and mastery; for example: 'I prefer course material that arouses my curiosity, even if it is difficult to learn.' The *Extrinsic Goal Orientation* scale (α = .77) consisted of four items designed to complement the intrinsic goal orientation subscale. The Extrinsic Goal Orientation subscale measures the degree to which participants perceive their engagement in learning tasks to originate from external factors such as grades, rewards, performance, evaluation by others, and competition; for example: 'Getting a good grade is the most satisfying thing for me right now.' When completing both these scales, participants were given the instruction to *think about their classes in general* because the items generally reflected long-term affective dispositions. The *Self-efficacy for Learning and Performance* subscale (α = .97) consisted of eight items measuring participants' self-appraisal of their ability and confidence to master the knowledge and skills presented during the lecture. For example, 'I'm confident I can understand the basic concepts taught in this lecture.' When completing this subscale, participants were given the instruction to *think about your experience in this lecture today* because the items reflected short-term affective dispositions.

In addition to these three MSLQ subscales, another three subscales - *Mastery Goal Orientation, Performance Goal Orientation, and Performance-Avoid Goal Orientation*- from the *Patterns of Adaptive Learning Scales (PALS)* (Midgley, et al., 2000) were used to measure participants' states/traits related to goal orientation towards achievement. These subscales are normally designed to measure the relationship between a participants' learning environment and their motivation, affect, and behaviour. Participants responded to the instruments using a 5-point Likert-type scale ranging from 1 - "Not at all true", 3 - "Somewhat true", and 5 - "Very true." The *Mastery Goal Orientation* (Revised scale, α = .84) included five items designed to measure the extent to which participants' purpose or goal in achievement settings is focused on developing their competence, extend their mastery and understanding; for example: 'It's important to me that I thoroughly understand my class work.' The *Performance-Approach Goal Orientation* (Revised scale, α = .90) included five items designed to measure the extent to which participants' purpose or goal in achievement settings is focused on demonstrating their competence; for example: 'One of my goals is to show others that class work is easy for me.' The *Performance-Avoid Goal Orientation* (Revised scale, α = .86) included four items designed to measure the extent to which participants' purposes or goal in achievement settings is focused on avoiding the demonstration of incompetence; for example: 'It's important to me that I don't look stupid in my classes.' When completing these three subscales, participants were given the instruction to *think about their classes in general* because the items generally reflected long-term affective dispositions.

Two more subscales were included, the *Self-Efficacy for Self-Regulated Learning* (α = .84), developed by Bandura (2006) and the *Student Trust in Faculty Scale* (α = .90) developed by Forsyth, Adams and Hoy (2011). First, the *Self-Efficacy for Self-Regulated Learning* (Bandura, 2006) included 11 items designed to measure participants' perceived capability to use a variety of self-regulated behavioral strategies during learning; for example: 'How well can you finish homework assignments by deadlines?' Participants responded using a 7-point Likert-type scale, ranging from 1- 'Not well at all,' 3 = 'Not too well', 5 - 'Pretty well' and 7 = 'Very well.' When completing this subscale, participants were given the instruction to *think about how you normally behave in terms of your schoolwork* because the items reflected behaviours indicative of general affective dispositions. The *Student Trust in Faculty Scale* (α=.91; Forsyth, Adams & Hoy, 2011) was designed to measure participants' trust in an instructor. The 13 items were adapted for use in the present study to measure the extent to which participants taking part in the Elementary Statistics lecture perceived the instructor as open, honest, reliable, competent, and benevolent; for example: 'The instructor of this lecture is/appears always honest with me.' Participants responded using a 4-point scale, ranging from 1 – 'Strongly disagree' to 4 – 'Strongly agree.' When completing this subscale, participants were given the instruction to *think about your experience in this lecture today* because the items reflected short-term affective dispositions.

In addition to the affective measures, we also included three cognitive measures. First, during the lecture, participants in both the LEI and control conditions were provided with copies of the lecture slides and they were asked to identify with the label "CM" any content information on the slides that was difficult to understand; that is, any content information on the slides that was *confusing* and therefore could lead to *mistakes* during learning. Second, a 10-item multiple-choice assessment on the topic of Elementary Statistics was developed to assess participants' performance on the knowledge and skills presented during the lecture. Third, a background questionnaire was administered at the end of the study, which included questions about participants' Grade 12 performance in Mathematics and English Language Arts, their current university GPA and their prediction of a grade in a class similar to the Elementary Statistics lecture delivered for the study (see Appendix E).

CHAPTER 4: RESULTS

As mentioned previously, the objective of this study was to conduct an experimental test of the LEAFF model, namely, the effect of discussing "learning errors" (LEI; see Method section) in an Elementary Statistics lecture to determine whether this manipulation generated stronger feelings of safety among learners, trust in the instructor compared to a control condition, and enhance learning outcomes. At the end of the Elementary Statistics lecture, that is, after the LEI and control condition were administered, eight instruments were administered to measure participants' affective states/traits and three instruments were used to measure participants' cognitive states/traits. The findings of this study are organized into two main sections: first, findings reflecting how the LEI versus the control condition influenced participants' affective dispositions, and, second, participants' cognitive dispositions. All statistical analyses, including descriptive and inferential statistics, were conducted in IBM SPSS Statistics Version 20 (IBM Corporation, 2011).

Affective Measures

Pearson Product-Moment correlations. The Pearson Product-Moment correlation coefficients among the total scores of affective measures – *Intrinsic Goal Orientation (TIGO), Extrinsic Goal Orientation (TEGO), Self-efficacy for Learning and Performance (TSELP), Mastery Goal Orientation (TMGO), Performance Goal Orientation (TPGO), Performance-Avoid Goal Orientation (TPAGO), Self-Efficacy for Self-Regulated Learning (TSESRL), Student Trust in Faculty Scale (TSTIS), and Feeling Scale (FS)* – by condition are shown in Table

2. As can be seen in Table 2, the pattern of correlations across participants' responses to the scales differed by condition. For example, in the control condition, intrinsic goal orientation (TIGO) was positively associated with all other goal orientation measures (i.e., TEGO, TMGO, TPGO) except Performance Avoid Goal Orientation (TPAGO). This was not the case for LEI, where intrinsic goal orientation (TIGO) was only positively associated with mastery goal orientation (TMGO) as one would expect given the similarity of intrinsic and mastery goal orientations; as expected, intrinsic goal orientation was also not associated with Performance Avoid Goal Orientation (TPAGO) in the LEI condition.

In the control condition, extrinsic goal orientation (TEGO) was positively associated with mastery goal orientation (TMGO), performance goal orientation (TPGO) and performance-avoid goal orientation (TPAGO). This pattern was not fully observed in the LEI condition, where TEGO was positively associated only with TPGO and TPAGO as one would expect given the similarity of external and performance-based orientations. These findings suggest that participants taking part in the LEI condition experienced more coherent goal orientations than those taking part in the control condition, especially given the pattern of correlations observed between TIGO and other goal orientations.

Among the self-efficacy subscales, the self-efficacy for learning and performance (TSELP) was not associated with any of the subscales in the control condition. However, in the LEI condition, TSELP was positively associated with intrinsic goal orientation (TIGO), extrinsic goal orientation (TEGO), and performance-goal orientation (TPGO) as expected given that self-efficacy for learning and performance involves both mastery-based and performance-based aspects. In contrast, the self-efficacy for self-regulated learning (TSESRL) was positively associated with intrinsic goal orientation (TIGO) and mastery goal orientation (TMGO) in the control condition, but this finding was not fully observed in the LEI condition. In the LEI condition, TSESRL was positively associated only with TIGO. Again, these findings suggest that participants in the LEI experienced more coherent affective dispositions, especially given the results associated with TSELP. Further, in the control condition, the student trust in faculty scale (TSTIS) was not positively associated to any of the goal orientation or self-efficacy subscales, whereas in the LEI condition, TSTIS was positively associated to intrinsic goal orientation (TIGO) and self-efficacy for learning and performance (TSELP). Finally, in the control condition, the aggregated feeling scale (FS) ratings were positively associated with mastery goal orientation (TMGO), and student trust in faculty scale (TSTIS). In the LEI condition, aggregated feeling scale ratings were positively associated with self-efficacy for learning and performance (TSELP), and student trust in faculty scale (TSTIS).

Non-parametric Mann-Whitney U-test for independent samples. Internal consistency values were relatively high for all scales used in the present study. However, any value less than 1 indicates some measurement imprecision and can reduce the effect size of experimental manipulations. When a set of aggregated scores are summed and used as a single point estimate for an underlying construct, this total score can be ambiguous in terms of its construct representation. For this reason, items across all eight scales were tested individually using the Mann-Whitney U-test for independent samples (see Table 3). Although conducting so many individual tests runs the risk of inflating the type-1 error, our results indicated that only a single subscale deserved further scrutiny – the *Student Trust in Faculty Scale (TSTIS).* As shown in Table 3, there were no significant differences found for any of the subscales or patterns for individual items between the control and LEI conditions except for the *Student Trust in Faculty Scale (TSTIS)*. For the STIS, participants' responses to three items were found to be significantly higher for the LEI than for the control condition. In particular, item STIS5 "The instructor of this lecture really listens to students," $Z = -2.79$, $p < .01$ (Mean Rank Control = 43.7, LEI = 58.2), item STIS6 "The instructor of this lecture is/appears always honest with me," $Z = -2.01$, $p < .05$ (Mean Rank Control = 45.5, LEI = 56.4), and item STIS10 "The instructor of this lecture DOES NOT care about students," $Z = -1.97$, $p < .05$ (Mean rank Control = 45.9, LEI = 56.0). (As a note, STIS10 was reversed to be aligned with all other items in the STIS scale before conducting any analysis given that its original polarity was negative instead of positive). When only these three items are aggregated, the subscale internal consistency is $\alpha = .68$ and participants' responses across the control and LEI conditions shows a statistically significant difference, with participants in the LEI indicating greater trust in the instructor, $Z = -3.00$, $p < .001$ (Mean Rank Control = 42.3, LEI = 59.5). These results provide evidence and are consistent with empirical studies in counselling psychology, which indicate that the seeds for safety and trust between individuals are established early in human interactions (Horvath, 2000). Shown in Table 3 are the mean scores, standard deviations, and ranks for each one of the scales in each condition, followed by the non-parametric independent samples test between conditions.

Repeated-measures analysis of Feeling Scale [FS]. All affective subscales except for the Feeling Scale (FS) were administered once during the study, after the Elementary Statistics lecture, to both the control and LEI conditions (see Table 1). However, the FS was administered 10 times during the Elementary Statistics lecture. Because repeated trials of participants' feelings were collected during the lecture, a repeated-measures analysis of the ratings was conducted. The repeated-measures representation of participants' ratings, shown in Figure 9, provides graphical evidence of the impact that an instructor's presentation of content and assessment practices can have on students' feelings of comfort within the learning environment. For example, consider how ratings #3, 4, 5, and 6, which are aligned to the presentation of Mode, Mean, Median, and Range, respectively (see Table 1), show a consistent rise in value for the LEI but stay approximately the same for the control. However, in both the LEI and control condition, there is a sharp decline in rating #7, which is aligned to the presentation of variance and standard deviation. Afterwards, there is a sharp increase for the LEI in rating #8 with the presentation of the example for how to calculate the variance and standard deviation. However, the increase is not as dramatic for the control condition. Rating #9 rises for the control condition but drops slightly for the LEI, which is nonetheless higher than the control. It is important to note at this point that ratings #1 through 8 could be viewed, more purely, as reflecting participants' feelings of comfort with the content of the lecture than ratings #9 and 10. Rating #9 was taken after participants completed the summative assessment and so this rating could be viewed as reflecting not only the material presented in the lecture but also participants' feelings about how they performed on the assessment. Likewise, rating #10 was taken after participants received feedback in the form of a percentage grade on their assessment performance so this rating reflects participants' feelings about the feedback they received on their performance.

A repeated-measure analysis was conducted of the feeling ratings to determine the effect of the learning error manipulation on participants' ratings of comfort during the presentation of material during the lecture. Only the first 8 ratings were included in the repeated-measure analysis because the last two ratings were confounded with the assessment and assessment feedback. A repeated-measure analysis of the ratings revealed a significant main effect across the 8 rating opportunities, Huynh-Felt correction F (4.374, 433.003) = 32.49, $p < .001$, partial eta squared = .25, indicating that participants' feelings changed significantly during the lecture. Further, ratings were consistently higher for participants taking part in the LEI condition than in the control condition, F (1, 99) = 4.50, $p < .05$, partial eta squared = .04. There was no interaction between ratings and condition, a result that is also evidenced in Figure 9. Not surprisingly, a separate repeated-measures analysis of the last 2 ratings –ratings #9 and 10 – revealed no statistically significant difference in feeling nor any difference between the control and LEI conditions. In the next section, the results associated with the cognitive variables are presented.

Cognitive Measures

Repeated-measure analysis of CM labels [CM]. During the Elementary Statistics lecture, participants in both conditions, LEI and control, were asked to label with a "CM" any content information on the PowerPoint slides that they thought might be *confusing* and could lead to making *mistakes*. A frequency count of CMs was made for all slides across both conditions. As shown in Table 4, out of 16 slides, only 13 PowerPoint slides could be labelled with CMs in both the control and LEI conditions based on content. Of these 13 PowerPoint slides, only 11 were actually found to have CMs indicated and two had none (slides 4 and 11). Another three slides –the assessment, peer-grading, and survey slides (slides # 16, 17, and 18)– did not convey any content in either the control or LEI condition and so were unlikely to be labelled with CMs. These three slides served to transition to the next segment of the study and, in fact, no participant indicated a CM on any of these slides. For example, the assessment slide (#16) was used to signal the onset of the summative assessment and to request a feeling rating but not to deliver any Elementary Statistics content to participants.

As can be seen in Table 4, participants in the control condition indicated fewer CMs across the 13 slides than those in the LEI condition; in particular, control participants identified a total of 33 CMs across the 13 slides, whereas LEI participants identified a total of 160. However, independent of condition, participants indicated most CMs (i.e., CM9 to CM12) in relation to variability content shown in slides 11 to 15 (LEI = 63.1% of 160, Control = 54.5% of 33), followed by CMs (i.e., CM 5 to CM8) in relation to central tendency content shown in slides 7 to 10 (LEI = 26.3% of 160, Control = 30.3% of 33). Figure 10 shows a graphical representation of the CM occurrences across slides for the control and LEI conditions, and the tendency for participants in the LEI condition to indicate more CMs than participants in the control condition.

A non-parametric k-related samples Friedman test was used to evaluate whether there were any statistically significant differences in the distribution of CMs across PowerPoint slides within condition. A non-parametric test was used because the CMs represented frequency counts and were not normally distributed. The Friedman test revealed a statistically significant difference of CM occurrences across slides within both the control and LEI condition, Control χ^2 (12, 50) = 31.00, $p < .01$; LEI - χ^2 (12, 51) = 245.91, $p < .001$. Further, a non-parametric 2-independent samples Mann-Whitney U test was used to evaluate any differences in the occurrence of CMs between the control and LEI conditions, which revealed a statistically significant difference between conditions, Control Mean Rank = 32.58; LEI Mean Rank = 69.06, $Z = -6.53$, $p < .001$.

Non-parametric Mann-Whitney U-test for independent samples on summative assessment and cognitive background measures. As mentioned previously, a 10-item summative assessment was administered at the end of the Elementary Statistics lecture to evaluate participants' mastery and performance on the content covered during the lecture. The assessment, shown in Appendix C, focused largely on interpretation and application of methods associated with central tendency and variability. Most items were not difficult as evidenced by item p-values ranging between .78 and .89 calculated separately for both groups (see Table 5). A non-parametric independent samples Mann-Whitney U-test was conducted to evaluate any differences in summative test performance and no differences were found between the conditions. Similarly, an analysis per item did not reveal any significant differences between conditions (see Table 6).

In addition to the summative assessment, participants were asked in the background survey administered after the assessment (see Appendix E) to predict their expected grade if they were to enrol in a class that was similar to the Elementary Statistics lecture presented during the study. Shown in Table 7 is the distribution of expected grades per condition. An independent samples Mann-Whitney U-test was conducted to evaluate any differences in expected grade and no significant differences were found (see Table 8). Finally, participants were also requested in the background questionnaire to provide information related to their previous achievement in High School, Grade 12 Mathematics and Language Arts, and to provide their current university GPA. An independent samples Mann-Whitney U-test was again conducted to evaluate any differences in Grade 12 Mathematics, Language Arts, and current GPA and no significant differences were found between conditions (see also Table 8).

Spearman Rho correlations between cognitive and affective variables. In addition to evaluating the statistical frequency of CMs across PowerPoint slides for each condition, we calculated the correlations between the cognitive measures and the affective measures. Specifically, as shown in Table 9, the cognitive measures included grade in Grade 12 Math, grade in Grade 12 English, current GPA, expected grade in a course similar to the lecture delivered in the present study, the summative assessment administered at the end of the Elementary Statistics lecture, and the total number of CMs indicated on the PowerPoint slides (TCMs). Also shown in Table 9, the affective measures included all the affective subscales (i.e., TIGO, TEGO, TMGO, TPGO, TPAGO, TSELP, TSESRL, TSTIS, and TFS).

One of the noteworthy findings that can be seen in Table 9 is that total scores on the *Feeling Scale* (TFS) were correlated with a select number of cognitive measures, but only for the LEI and not for the control condition. For example, total scores for the FS were correlated significantly and positively with participants' *expected grade* in a lecture similar to the one delivered in the present study (r = .517) and with their *summative assessment* score (r = .520). However, other correlations were significantly negative. For example, total scores for the FS were negatively correlated with total number of CMs (r = -.547). The Figures 11 and 12 show the graphical correspondence between the 10 PowerPoint slides where a feeling rating was requested and the occurrence of CMs for the control and LEI, respectively. Inspection of Figures 11 and 12 suggests that there was a correspondence between participants indicating that material was confusing and indicating less-than-good feelings on the FS about the material being presented in the lecture. Another noteworthy finding shown in Table 9 is that occurrence of CMs was negatively and significantly correlated with total scores for the Self-

Efficacy for Learning and Performance (TSELP) in the LEI (r = -.447) but not for the control condition.

In addition, in both the LEI and control conditions, participants' *expected grade* in a lecture similar to the one delivered in the present study was positively and significantly correlated with total scores on the SELP, (r = .678 for control; r = .814 for LEI). In the control condition, participants' expected grade was also positively and significantly correlated with total scores on the Performance-Avoid Goal Orientation (TPAGO) scale (r = .369). Interestingly, only in the control condition, current GPA was positively and significantly correlated with total scores on the Intrinsic Goal Orientation (TIGO; r = .307), Extrinsic Goal Orientation (TEGO; r = .333), Performance Goal Orientation (TPGO; r = .323), Performance Avoid Goal Orientation (TPAGO; r = .308), and Self-Efficacy Regulated Learning (TSERL; r = .386) scales. The grades for Grade 12 English were not correlated with any of the affective measures in both the control and LEI conditions. However, the grades in Grade 12 Mathematics showed more positive and significant correlations in the control condition than in the LEI condition. Specifically, in the control condition the grades Grade 12 math were positively and significantly correlated with total scores on TPGO (r = .312), TPAGO (r = .357) and TSELP (r = .557), but in the LEI the grades in Grade 12 math were only correlated with total scores on TSELP (r = .419).

CHAPTER 5: DISCUSSION AND CONCLUSION

This section is organized in five parts. It begins with a summary of the present investigation and method, followed by a discussion of the findings in the study in relation to both the hypothesis tested and previous studies. The importance of this study for research and practice is also elaborated, followed by the limitations. In the last part, recommendations for future research are presented.

Summary

To enhance learning and assessment experiences in classrooms and in other learning settings, instructional designs need to consider learners' mental models as they function as 'frames' in guiding students to behave in particular ways. Mental models include not only cognitive variables, such as knowledge and skills, but also affective variables, such as feelings of safety, comfort, and trust. The nature of learners' mental models depends on many variables such as parental influences, peer interactions, and classroom experiences. The LEAFF model, the theoretical framework used in the present study, indicates that the classroom learning environment created by the teacher can have a significant influence on students' mental models. For example, if a teacher can create trusting relationships with students by talking about the value of learning errors in classroom discussions and increase students' comfort and safety in the learning process, students' mental models can be expected to support their specific behaviours such as revealing what they do not understand. In turn, students increase their opportunities to obtain feedback and learn because they are involved in helping to identify areas of challenge for teachers.

In the study reported in the present paper, the first part of the LEAFF model was tested. The first part of the LEAFF model is focused on the *Instructional Climate/Environment* created by the teacher. In the present study, an intervention was developed to determine if discussing the value of learning errors, within an Elementary Statistics lecture, and explaining the expectations of such errors for learning, increased students' feelings of safety or comfort in the learning environment, trust in the instructor, and likelihood of indicating areas of confusion with the content of the lecture. The *learning errors intervention* (LEI) used in the present study was operationalized as a between-subject variable. In contrast to the control group, the LEI involved an embedded targeted discussion of learning errors within an Elementary Statistics lecture.

Based on the LEAFF model, we hypothesized that the discussion of learning errors would influence students' mental models in positive ways by increasing their comfort in the class and trust in the instructor. In turn, we expected that influencing mental models in positive ways would increase students' likelihood of indicating areas of confusion with the presented material. In the study, participants were randomly assigned to either the LEI or control condition. In each condition, participants received a 40-minute Elementary Statistics lecture, rated their feelings 10 times during the length of the lecture, answered a 10-item multiple choice assessment and completed a series of surveys (subscales) at the end of the lecture designed to measure affective and cognitive states and traits. Descriptive and inferential statistics were conducted on the data and results revealed preliminary support for the predictions derived from the LEAFF model.

Discussion of Results

The present discussion elaborates on the findings in four areas: (a) participants' feeling ratings during the Elementary Statistics lecture, (b) participants' reports of trust in the instructor, (c) participants' identification of CMs to illustrate content that they found confusing and possibly leading to mistakes, including their assessment performance, and (d) participants' motivation and beliefs towards learning and performance.

Participants' feeling ratings. In the LEI and control conditions, participants were asked to rate their feelings of comfort 10 times during the lecture (see Appendix B). We hypothesized that if the LEI produced the results expected, namely, an increase in students' overall sense of safety during the lecture, they would indicate higher feelings of comfort during the lecture compared to the control group. The results provided support for our hypothesized expectations. First, as expected, we found that the feeling scale was sensitive to fluctuations in participants' feelings of comfort with the content of the lecture. For example, participants within each condition reported higher feelings of comfort with easier material (e.g., the mode) than with more difficult material (e.g., variance). The FS scale was particularly informative in measuring the temporal and dynamic interplay between content and participants' feelings of comfort. Second, and as expected, we found that consistently throughout the lecture, participants in the LEI condition indicated significantly higher feelings of comfort compared to participants in the control condition.

The fact that participants in the LEI condition consistently reported more positive feelings than participants in the control condition provides empirical support for the intervention delivered based on the LEAFF model. Although the intervention was of short duration and could be critiqued as insufficient for changing entrenched beliefs about mathematics or statistics, the evidence indicates that even a brief intervention can momentarily influence students' mental models of the learning situation and induce higher ratings of comfort compared to a control group. Our findings can be situated within a broader research literature on the role that emotions play in academic performance and achievement, and particularly in the areas of Mathematics and Statistics (Frenzel, Pekrun, & Goetz, 2007; Kleine, Goetz, Pekrun, & Hall, 2005; Wei & Mei, 2005; Zeidner, 1999). Although past research indicates that math anxiety during learning and performing can originate from negative prior experiences in math, including low math self-efficacy and lack of applicability of the content in students' real life context, our findings indicate that feelings may be positively influenced to increase openness to learning. The trends of the FS ratings under both LEI and control conditions showed similarities in how feelings fluctuated according to the content presented, with ratings decreasing after an overview of the lecture content, increasing slightly with the presentation of central tendency measures and decreasing sharply when variance was discussed. Although, the LEI condition showed the same trends in ratings as in the control condition, overall ratings were consistently higher in the LEI, indicating that participants' feelings of comfort throughout the class can be significantly increased even with the presentation of challenging content.

Participants' trust in the instructor. Based on results associated with three specific items in the Student Trust in the Faculty Scale (STIS; Forsyth et al., 2011), we found that participants taking part in the LEI indicated higher levels of trust in the instructor than participants in the control condition. Unlike participants in the control condition, participants in the LEI indicated viewing the instructor as being more benevolent, open, and reliable. This result again provides preliminary support for the LEAFF model. As the model indicates, an instructor's discussion of the value of learning errors is a way to begin to create a safe classroom environment where learners can view the teacher as an open and honest individual who can freely talk about mistakes, and where learners feel safe talking about material they do not understand. Although we did not find all items in the STIS to differentiate between LEI and control participants, three items appeared to be sufficiently sensitive to detect differences between the conditions in their trust of the instructor.

The statistically significant differences found between the control and LEI conditions for the Students Trust in Faculty Scale's (STIS) items suggest that the instructor-led discussion of learning errors conveyed a sense of benevolence, openness and reliability in students. Participants in the LEI condition found in the instructor more consideration for their needs, willingness to support their learning interests, transparency in the intentions of the teaching and assessment practices implemented in the class, and potential to establish a teacher-student relationship in which they could depend on the instructor for help. As predicted from the LEAFF model, then, the discussion of learning errors appears to have the potential to foster students' positive mental models. In turn, such positive mental models could facilitate participants' feelings of safety with the instructor in particular and with the learning environment in general so as to risk identifying what they do not understand –such as confusing material that could lead to mistakes-.

Participants' identification of CMs. During the Elementary Statistics lecture, participants were asked to indicate with a "CM" those locations on the PowerPoint slides where content was confusing and could lead to mistakes. We expected, based on the LEAFF model, that participants taking part in the LEI would be more likely to indicate CMs on the PowerPoint slides than participants taking part in the control condition. We expected more CMs from LEI participants than control participants because the LEAFF model indicates that stronger feelings of safety in the learning environment should lead to more positive mental models, which, in turn should lead learners to behave in particular ways such as taking risks to indicate material they do not understand. Our results revealed that LEI participants indicated CMs more often than control participants. In other words, the LEI appears to have facilitated participants' comfort with the learning environment and trust in the instructor (discussed previously) to such an extent that they openly expressed their doubts about the Elementary Statistics' content. This result provides preliminary evidence of the effect of the learning errors intervention – namely, the instructor's discussion of the value of learning errors on students' learning of complex material.

Although we do not have evidence to indicate that participants in the LEI condition were necessarily more cognizant than control participants of places on the PowerPoint slides where errors could be made, we do infer from their greater use of the label CM that they were more at ease indicating where errors could be made. It is possible that control participants were just as cognizant as LEI participants of places on the PowerPoint slides where confusion could arise and mistakes could be made but, unlike LEI participants, hesitated to indicate these locations due to discomfort with the learning environment.

The importance of having participants feel safe to identify errors should not be underestimated – learners who can recognize areas where errors can occur may be able to avoid attributing confusion and mistakes to their own cognitive limitations, and to experience greater openness to receive and act on an instructor's formative feedback. Previous research studies (e.g., Frese & Altman, 1989; Lorenzet et al., 2005) have come to similar conclusions; for example, Frese and Altman (1989) in the context of training learners to develop software found evidence to support the pedagogical value of having learners identify common sets of errors to help reduce the likelihood that they would attribute errors to uncontrollable, internal causes. In addition, these studies revealed that the use of errors during training was associated to superior performance and academic self-efficacy. Learners' identification of CMs can help teachers foresee areas requiring special attention during learning and assessment, while at the same time guiding students to grasp content and skills in those specific areas.

Participants' motivation and beliefs towards learning and performance. Research on self-efficacy indicates that there is a positive correlation between the beliefs learners have about their academic self-efficacy and, their goal orientation and performance (Bandura, 1991); for example, stronger perceived self-efficacy is associated with setting higher goals and making firm commitments to achieve goals. Also, stronger, more positive beliefs about academic self-efficacy are associated with more effective self-regulation strategies for and during learning, which, in turn, can help facilitate performance. In the present study, a series of scales in addition to the STIS, were administered to participants after the Elementary Statistics lecture to measure self-efficacy and goal orientations (see Table 2). One of the study findings was that participants, independent of the condition, did not show statistically significant differences in their mean responses on these measures, which included self-efficacy for self-regulated learning (SESRL), self-efficacy for learning and performance (SELP), intrinsic goal orientation (IGO), extrinsic goal orientation (EGO), mastery goal orientation (MGO), performance goal orientation (PGO), and performance-avoid goal orientation (PAGO). Given this result, one might conclude that the learning errors intervention did not influence participants' reported self-efficacy and goal orientations. However, there were some differences between the control and LEI conditions in the pattern of correlations among these measures as shown in Table 2. In particular, for the control group, subscales measuring distinct types of goal orientations were strongly associated – for example, intrinsic goal orientation and extrinsic goal orientation (r = .410). This was not the case for LEI participants. Likewise, in the control condition, performance-goal orientation was moderately correlated with intrinsic and mastery goal orientations (*rs* = .352 and .298). Again, this was not case for LEI participants. The patterns within the conditions suggested that participants responded “more purely” in terms of their goal orientations in the LEI than in the control condition.

However, the most notable finding was the pattern of correlations found for the total score for self-efficacy for learning and performance (TSELP) and other subscales. In the control condition responses on the SELP scale were uncorrelated with any of the other subscales, but in the LEI condition it was strongly correlated with intrinsic goal orientation, extrinsic goal orientation and performance goal orientation as one would expect given that SELP involves not only intrinsic goals but also performance-based goals. The instructions for the SELP requested participants “to think about their experience in the lecture”, and the items asked them to consider their likelihood of success in a class similar to the lecture they were being delivered; for example, “I believe I could receive an excellent grade on an exam in this lecture” and “I’m certain I can understand the most difficult material presented in this lecture.” That the SELP scale was strongly and positively correlated with goal orientation measures in the LEI but not in the control conditions could be interpreted as signalling the effectiveness of the learning errors intervention for galvanizing participants’ clarity in their goal orientations.

Nevertheless, one concern worth mentioning with interventions focused on errors in classroom discussions is that students can find themselves experiencing negative emotions because the topic of errors can be anxiety provoking (see Frese & Altman, 1989). Further, negative emotions such as anxiety could lower students' academic self-efficacy and diminish self-appraisals of competence (Kluger & DeNisi, 1996). Although we did not find evidence that the learning errors intervention caused participants' self-efficacy to diminish –that is, there were no significant differences between the LEI and control conditions on SESRL or SELP – we did find evidence, at least for LEI participants, that responses on the SELP scale were negatively correlated with frequency of CMs (see Table 9). In other words, participants who indicated more CMs on the PowerPoint slides were also those who reported lower self-efficacy for learning and performance. Thus, one interpretation of this finding is that the learning errors intervention may in fact help those students, who need the most guidance as evidenced by their low self-efficacy scores, to find the courage to indicate using CMs what they do not understand about the content material. Over time, their self-efficacy for learning and performance may improve as they find the courage to indicate what they do not understand, obtain feedback, and gain confidence with their knowledge and skills.

Studies in the careful integration of errors during instruction, which include framing errors as a natural part of the learning process, show that negative affective reactions can be avoided and used to help students achieve higher levels of self-efficacy and performance (e.g., Frese & Altman, 1989; Lorenzet et al., 2005). In this way, errors are clearly framed as not occurring due to deficiencies in cognitive ability but as expected behaviours in light of building complex knowledge. For example, Lorenzet et al. (2005) used guided-errors to help students acquire software presentation knowledge and skills, leading to improved faster performance and enhanced self-efficacy in comparison to students who were not trained with guided-errors.

Importance of this Study

The conclusions presented in this study provide support for the first part of the LEAFF Model and add to the growing body of literature on trust in the context of educational organizations (see Forsyht, Adams & Hoy, 2011). The findings of the current study provide specific empirical evidence to support the role of trust in building feelings of safety during learning and assessment in classrooms.

The classroom intervention presented and tested in this study suggests that including discussion of learning errors, early in the class and at specific moments such as prior to assessment, promotes higher feelings of comfort in participants receiving instruction in a mathematical content. In addition, this type of brief intervention can encourage students to indicate content material that appears confusing and could lead to mistakes. Finally, although this type of intervention requires further study, initial results are promising for helping to create opportunities for higher-risk pedagogical interactions between students and teacher regarding content complexity. Furthermore, a classroom environment inhabited by trusting relationships is expected to set up appropriate conditions for effective formative feedback experiences. Students who are not afraid of showing what they know and what they do not know are creating mental models that are more transparent and that are more likely to benefit from formative feedback than those students who are afraid to verbalize their sources of confusion. Consequently, students accepting the vulnerability that arises during learning are expected to ask or use feedback to build stronger cognitive structures, demonstrate higher performance, and have more positive motivation and beliefs towards learning.

Although continued research must be conducted of the LEAFF Model, the present study suggests that the inclusion of a learning errors intervention in everyday teaching and assessment practices is possible and can have positive results on students' affective and cognitive states. Moreover, a lesson learned from this testing of the LEAFF Model is that even a short and simple classroom intervention operationalizing safety and trust has the potential to enhance students' mental models and the relationship between students and teachers.

Limitations

As mentioned throughout the present document, discussing the necessity of learning errors with students is a promising strategy to help students create positive mental models of the learning environment. In this investigation, data were collected from 101 undergraduate students during a one-hour lecture and results indicated that discussion of learning errors led to specific positive results. It is important to note that these results can be generalized to populations with similar characteristics to the ones in the present study, and also replicating the same experimental and instructional design. However, future studies need to be conducted to explore the ecological validity (generalizabilty) of these findings to other populations and conditions. Nevertheless, the results are encouraging and provide a motivation to include the *learning errors intervention* within instructional units of longer duration, in other academic levels and with other subject matters.

Moreover, the instructional unit implemented in this study was in Elementary Statistics, a content that has inherent negative emotions linked to it, given its dependence on mathematical concepts. Although the learning errors intervention showed favourable effects on participants' positive feelings and trust during the class, it will be necessary to develop future studies to determine whether these relationships extend to content other than mathematics.

Because the present study was focused on investigating the effects of a brief learning error intervention, there was no plan or opportunity to measure either long-term cognitive gains or changes in affective states as consequence of the intervention. In future studies, it will be important to measure and identify possible gains in performance and changes in cognitive and affective states or traits as consequence of a learning errors intervention.

Recommendations for Future Research

Future studies are encouraged to address the limitations outlined in this paper. In this study, a short classroom-based intervention showed the potential to promote students' feeling of comfort and trust in their learning environment, and risk indicating areas of confusion with the content material. Once a safe learning environment is cultivated and students are more apt to create positive mental models, more specific investigations can be undertaken such as exploring the specific psychological interactions between students and teachers regarding subject-matter content, and particularly those interactions regarding correction of errors, to help illuminate how feedback can be provided to increase students' learning and performance.

BIBLIOGRAPHY

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement.* White Plains, NY: Longman.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117-148.

Bandura, A.(1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*, 199-231.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school: Expanded Edition.* Washington, D.C., USA: National Academy Press.

Brown, D. F. (2005). The Significance of Congruent Communication in Effective Classroom Management. *Clearing House: A Journal of Educational Strategies, Issues and Ideas, 79*(1).

Burke, C. S., Sims, D. E., Lazzara, E. H. & Salas, E. (2007). Trust in leadership: A multi-level review and integration. *The Leadership Quarterly, 18*, 606-632.

Cassady, J. C. (2010). Test anxiety: Contemporary theories and implications for learning. In J.C. Cassady (Ed.), *Anxiety in schools. The causes, consequences, and solutions for academic anxieties* (pp. 7-26). New York, NY: Peter Lang.

Cizek, G. J., & Burg, S. S. (2006). *Addressing test anxiety in a high-stakes environment: Strategies for classrooms and schools.* Thousand Oaks, CA: Corwin Press.

Chu, M., Guo,Q., & Leighton, J. P. (2013). *Students' Interpersonal Trust and Attitudes Towards Standardized Tests: Exploring Affective Variables Related to Student Assessment.* Manuscript submitted for publication.

Church, M. A., Elliot, A. J., & Gable, S. L. (2001). Perceptions of Classroom Environment, Achievement Goals and Achievement Outcomes. *Journal of Educational Psychology*, *93*(1), 43-54.

Dweck, C. S., & Elliot, A. J. (Eds.). (2005). *Handbook of competence and motivation*. New York, NY: The Guilford Press.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725-747.

Frenzel, A. C., Pekrun, R. & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), 478–493.

Forsyth, P. B., Adams, C. M., & Hoy, W. K. (2011). *Collective Trust: Why Schools Can't Improve Without It.* New York, NY: Columbia University Teachers College Press.

Frijda, N. H., & Mesquita, B. (1995). The social roles and functions of emotions. In S. Kitayama & H.R. Markus (Eds.), *Emotion and culture* (pp. 51–88). Washington, DC, USA: American Psychological Association.

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin, 97*, 363-386.

Horvath, A. O. (2000). The therapeutic relationship: From transference to alliance. *Psychotherapy in Practice, 56*(2), 163-173.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counselling Psychology, 38*, 139-149.

Hoy, W., & Tschannen-Moran, M. (1999). Five faces of trust: An empirical confirmation in urban elementary schools. *Journal of School Leadership*, *9*(3), 184-208.

Hoy, W., & Tschannen-Moran, M. (2003). The conceptualization and measurement of faculty trust in schools: The omnibus T-scale. In W. Hoy & C. Miskel (Eds.), *Studies in leading and organizing schools* (pp.181-207). Greenwich, CT: Information Age Publishing.

Huberty, T. J. (2009). Test and performance anxiety. *Principal Leadership, 10* (1), 12-16.

IBM Corporation. (Released 2011). *IBM SPSS Statistics for Windows, Version 20.0.* Armonk, NY: IBM Corp.

Johnson-Laird, P. N. (1983). Mental models. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (2004). Reasoning with Mental Models. *International Encyclopedia of the Social and Behavioral Sciences*. 12821-12824

Johnson-Laird, P. N. (2004). The history of mental models. In K. Manktelow (Ed.), *Psychology of reasoning: Theoretical and historical perspectives* (pp. 179-212). London: Psychology Press.

Kochanek, J. R. (2005). *Building trust for better schools: research-based practices.* Thousand Oaks, CA: Corwin Press.

Kilpatrick, J., Swafford, J., & Findell, B. (Ed.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC, USA: National Academies Press.

Kleine, M. Goetz, T., Pekrun, R., & Hall, N. (2005). The structure of students' emotions experienced during a mathematical achievement test. *ZDM, 37*(3), 221-225.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119,* 254-284.

Lajoie, S. P. (2008). Metacognition, self regulation, and self-regulated learning: A rose by any other name? *Educational Psychology Review, 20,* 469-475.

Leighton, J. P. (2009). Can you learn how to learn for life? Components from expert learning research. *Journal of Applied Research on Learning, 2* (Special issue), 1-14.

Leighton, J. P., Chu, M-W., & Seitz, P. (2012). Cognitive Diagnostic assessment and the learning errors and formative feedback (LEAFF) model. In R. Lissitz (Ed.), *Informing the Practice of Teaching Using Formative and*

Interim Assessment: A Systems Approach (pp. 183-207). Charlotte, NC: Information Age Publishing.

Leighton, J. P., & Gierl, M. J. (Eds.). (2007). *Cognitive diagnostic assessment for education. Theory and applications*. Cambridge, MA: Cambridge University Press.

Leighton, J. P. & Gierl, M. J. (2011). *The learning sciences in educational assessment.* Cambridge, MA: Cambridge University Press.

Leighton, J. P. & Sternberg, R. J. (2012). Reasoning and problem solving. In I.B. Weiner, A. F. Healy, & R W. Proctor (Eds.), *Handbook of psychology* (2nd ed.), *Vol. 4 Experimental psychology* (pp. 631-659). New York, NY: Wiley.

Lorenzet, S. J., Salas, E., & Tannenbaum, S. I. (2005). Benefiting from mistakes: The impact of guided errors on learning, performance, and self-efficacy. *Human Resource Development Quarterly, 16*, 301-322.

Louis, K. S. (2007). Trust and improvement in schools. *Journal of Educational Change. 6*(1), 1-24.

Louis, K. S. (2008). Creating and sustaining professional communities. In R. Coles & A. Blankenstein (Eds.), *Sustaining learning communities* (pp. 41-57). Thousand Oaks, CA: Sage.

Mayer, R. C., & Davis, J. H. (1999). The effect of performance appraisal system on trust for management: a field quasi-experiment. *Journal of Applied Psychology*, *84*, 123−136.

McGregor, H.A., & Elliot, A. J. (2005). The shame of failure: examining the link between fear of failure and shame. *Personality and Social Psychology Bulletin*, *31*, 218-231.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., Gheen, M., Kaplan, A. Kumar, R., Middleton, M. J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the patters of adaptive learning scales*. University of Michigan.

Misztal, B. A. (1996). *Trust in modern societies: The search for the bases of social order*. Cambridge, UK: Polity Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Oxford, UK: Prentice-Hall.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review, 103,* 241-262.

Osborne, J.W., Tilman, D., & Holland, A. (2010). Stereotype threat and anxiety for disadvantaged minorities and woman. In J. C. Cassady (Ed.), *Anxiety in schools. The causes, consequences, and solutions for academic anxieties* (pp. 7-26). New York, NY: Peter Lang.

Pekrun, R. (2011). Emotions as drivers of learning and cognitive development. In R. A. Calvo & S. K. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies, Vol 3. Explorations in the Learning Sciences, Instructional Systems and Performance Technologies* (pp 23-39). New York, NY: Springer.

Pintrich, P. R., Smith, D. A.F., Garcia, T., & McKeachie, W. J. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ). Ann Arbor, MI: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Priest, A., & Roach, P. (1991). Learning from errors. *Cognitive Systems, 3,* 79-102.

Reason, J. T. (1990). *Human error.* Cambridge, England: Cambridge University Press.

Rejeski, W. J., Gauvin, L., Hobson, M. L., & Norris, J. L. (1995). Effects of baseline responses, in-task feelings, and duration of activity on exercise-induced feeling states in women. *Health Psychology*, *14*(4), 350-359.

Sarason, L. G. (1980). *Test anxiety: Theory, research, and applications.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Schank, R. C. (1986). *Explanation patterns: Understanding mechanically and creatively*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. In C. Ames & R. Ames (Eds.), *Research on motivation in education, Vol. 3, Goals and cognitions* (pp. 13-44). San Diego, CA: Academic.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207-231.

Sebring, P. B., & Bryk, A. S. (2000). School leadership and the bottom line in Chicago. *Phi Delta Kappan, 81*(6), 440-443.

Seitz, P., Chu, M., Bustos, M. & Leighton, J. (2012, May). Creating safe classroom environments using the learning errors and formative feedback (LEAFF) model. In C. Poth (Chair), *Conceptual and Technical Advances to Enhance Teaching and Learning Environments*. Symposium conducted at the Canadian Society for the Study of Education CSSE Conference, Waterloo, Ontario Canada.

Shute, V. (2008). Focus on formative assessment. *Review of Educational Research, 78*, 153-189.

Stigler, J. W., & Stevenson, H. W. (1991). How Asian teachers polish each lesson to perfection. *American Educator, 15*, 12-20.

Thorndike, E. L. (1913). *Educational psychology: The psychology of learning (Vol.* 2). New York, NY: Columbia University Teachers College Press.

Tschannen-Moran, M. (2004). *Trust matters: Leadership for successful schools*. San Francisco, CA: Jossey-Bass.

Tschannen-Moran, M., & Hoy, W. (1998). Trust in schools: a conceptual and empirical analysis. *Journal of Educational Administration, 36*(4), 334-352.

Tschannen-Moran, M., & W. Hoy. (2000). A multidisciplinary analysis of the nature, meaning, and measurement of trust. *Review of Educational Research, 70*(4), 547-593.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *The Journal of Mathematical Behavior, 3*, 3-71.

VanLehn, K. (1987). *Student modelling. (Tech. Report PCG-4).* Pittsburgh, PA: Departments of Psychology and Computer Science, Carnegie-Mellon University.

Wei, P., & Mei, T. (2005). Students' perceptions on factors of statistics anxiety and instructional strategies. *Journal of Instructional Psychology, 32*(3), 205-214.

Willis, J. (2010). *Learning to love math: Teaching strategies that change student attitudes and get results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: some interesting parallels. *British Journal of Educational Psychology, 61* (3), 319–328.

Zeidner, M. (1998). *Test anxiety: The state of the art.* New York, NY: Plenum Press.

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P.A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 165-184). San Diego, CA, Academic Press.

Table 1

Alignment of Control and LEI PowerPoint Slides with Feeling Scale Rating Occasions

PowerPoint Slide	Control	Rating Occasions	Intervention	Rating Occasion
1	Title/Introduction	0	Title/Introduction	0
2	--	--	Learning Errors Intervention	0
3	--	--	Learning Errors Intervention Cont'd	0
4	Feeling Scale introduction	1	Feeling Scale introduction	1
5	Statistics Framework	0	Statistics Framework	0
6	Data Set	2	Data Set	2
7	Central Tendency	0	Central Tendency	0
8	Mode/Definition/Formula	3	Mode/Definition/Formula	3
9	Mean/Definition/Formula	4	Mean/Definition/Formula	4
10	Median/Definition/Formula	5	Median/Definition/Formula	5
11	Variability	0	Variability	0
12	Range/Definition/Formula	6	Range/Definition/Formula	6
13	Variance/Definition/Formula	7	Variance/Definition/Formula	7
14	Example Calculation	0	Example Calculation	0
15	Example Calculation Cont'd	8	Example Calculation Cont'd	8
16	Assessment*	9	Assessment/Learning Errors Intervention*	9
17	Peer grading**	10	Peer grading**	10
18	Surveys	0	Surveys	0
TOTAL		10		10

* Rating occurred after participants complete the assessment but before grading/feedback
** Rating occurred after participants receive assessment grading/feedback

Table 2

All Affective Measures: Pearson Moment Correlations

Condition	**Subscale**	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**
Control	1. TIGO	-	.410**	.552***	.352*	.053	.344*	.144	-.116	.056
	2. TEGO	-	-	.413**	.547***	.366**	.185	.095	-.046	.050
	3. TMGO	-	-	-	.298*	-.058	.301*	-.117	-.045	.290*
	4. TPGO	-	-	-	-	.650***	.103	.236	.122	.084
	5. TPAGO	-	-	-	-	-	-.039	.194	.115	.098
	6. TSESRL	-	-	-	-	-	-	-.065	.114	.044
	7. TSELP	-	-	-	-	-	-	-	.153	.323*
	8. TSTIS	-	-	-	-	-	-	-	-	.436**
	9. TFS	-	-	-	-	-	-	-	-	-
LEI	1. TIGO	-	.158	.518***	.194	.038	.350*	.394**	.298*	.278*
	2. TEGO	-	-	.102	.538***	.371**	.004	.334*	.075	.165
	3. TMGO	-	-	-	.016	.205	.146	.040	.253	-.027
	4. TPGO	-	-	-	-	.586***	.114	.396**	.197	.236
	5. TPAGO	-	-	-	-	-	-.177	.166	.145	.024
	6. TSESRL	-	-	-	-	-	-	.023	.153	.031
	7. TSELP	-	-	-	-	-	-	-	.330*	.627***
	8. TSTIS	-	-	-	-	-	-	-	-	.474***
	9. TFS	-	-	-	-	-	-	-	-	-

*** $p < 0.001$ level (2-tailed).
** $p < 0.01$ level (2-tailed).
*$p < 0.05$ level (2-tailed).

Table 3

Affective Measures: Non-parametric Mann-Whitney U-test for independent samples, and analyses per item

Subscale	Internal Consistency (α)	Item	Condition	Mean	SD	Mean Rank	Mann-Whitney U-test	
							Z	P
Intrinsic Goal Orientation (IGO)	.78	IGO1	Control	5.62	1.398	49.98	-.363	.716
			LEI	5.61	1.168	52.00		
		IGO2	Control	5.30	1.581	51.68	-.239	.811
			LEI	5.45	1.376	50.33		
		IGO3	Control	5.18	1.848	48.83	-.754	.451
			LEI	5.06	1.678	53.13		
		IGO4	Control	5.02	1.558	48.33	-.923	.356
			LEI	5.18	1.682	53.62		
		TIGO	Control	17.60	4.000	49.34	-.566	.572
			LEI	18.08	4.516	52.63		
Extrinsic Goal Orientation (EGO)	.77	EGO1	Control	4.54	1.092	52.16	-.408	.683
			LEI	4.63	1.148	49.86		
		EGO2	Control	5.30	1.182	50.33	-.235	.815
			LEI	5.16	1.488	51.66		
		EGO3	Control	4.70	1.446	52.73	-.600	.548
			LEI	4.94	1.462	49.30		
		EGO4	Control	3.06	1.449	49.08	-.665	.506
			LEI	3.35	1.610	52.88		
		TEGO	Control	21.12	4.897	50.76	-.082	.935
			LEI	21.29	4.649	51.24		

Mastery Goal Orientation (MGO)	.84	MGO1	Control	3.80	.756	51.70	-.257	.797
			LEI	3.75	.845	50.31		
		MGO2	Control	3.84	.912	51.16	-.058	.954
			LEI	3.84	.758	50.84		
		MGO3	Control	3.64	.985	50.74	-.093	.926
			LEI	3.67	.841	51.25		
		MGO4	Control	3.94	.843	46.77	-1.563	.118
			LEI	4.20	.749	55.15		
		MGO5	Control	4.10	.909	50.20	-.295	.768
			LEI	4.20	.749	51.78		
		TMGO	Control	19.32	3.617	50.15	-.290	.772
			LEI	19.65	2.911	51.83		
Performance Goal Orientation (PGO)	.89	PGO1	Control	3.14	1.246	51.44	-.154	.878
			LEI	3.10	1.136	50.57		
		PGO2	Control	3.12	1.154	51.50	-.175	.861
			LEI	3.06	1.139	50.51		
		PG03	Control	2.12	.918	48.19	-1.018	.308
			LEI	2.31	.969	53.75		
		PG04	Control	2.50	1.233	51.04	-.014	.989
			LEI	2.51	1.206	50.96		
		PGO5	Control	2.44	1.264	50.27	-.256	.798
			LEI	2.47	1.120	51.72		
		TPGO	Control	13.32	4.880	50.91	-.031	.976
			LEI	13.45	4.764	51.09		
Performance-Avoid Goal Orientation	.86	PAGO1	Control	3.58	1.108	48.58	-.857	.391
			LEI	3.75	1.055	53.37		
		PAGO2	Control	2.74	1.275	51.88	-.308	.758
			LEI	2.67	1.160	50.14		

Performance -Avoid Goal Orientation cont'		PAGO3	Control	3.00	1.278	47.26	-1.317	.188
			LEI	3.33	1.033	54.67		
		PAGO4	Control	2.78	1.200	50.74	-.091	.927
			LEI	2.80	1.132	51.25		
		TPAGO	Control	12.10	4.220	49.43	-.536	.592
			LEI	12.55	3.557	52.54		
Self-Efficacy for Self-Regulated Learning (SESRL)	.84	SESRL1	Control	5.84	1.315	50.44	-.200	.841
			LEI	5.84	1.332	51.55		
		SESRL2	Control	3.92	1.576	51.72	-.249	.803
			LEI	3.82	1.519	50.29		
		SESRL3	Control	4.56	1.373	49.85	-.400	.689
			LEI	4.61	1.429	52.13		
		SESRL4	Control	4.96	1.690	45.37	-1.963	.050
			LEI	5.57	1.446	56.52		
		SESRL5	Control	3.86	1.852	45.76	-1.802	.072
			LEI	4.53	1.837	56.14		
		SESRL6	Control	4.64	1.687	48.91	-.722	.470
			LEI	4.92	1.547	53.05		
		SESRL7	Control	4.90	1.705	49.73	-.440	.660
			LEI	5.10	1.526	52.25		
		SESRL8	Control	4.76	1.098	51.33	-.116	.908
			LEI	4.71	1.316	50.68		
		SESRL9	Control	4.70	1.810	50.95	-.017	.986
			LEI	4.73	1.779	51.05		
		SESRL10	Control	4.32	1.647	52.39	-.481	.630
			LEI	4.14	1.744	49.64		
		SESRL11	Control	3.86	1.830	49.11	-.651	.515
			LEI	4.10	1.803	52.85		
		TSESRL	Control	50.32	10.630	48.07		

			LEI	52.06	11.122	53.87	-.996	.319
Self-Efficacy for Learning and Performance (SELP)	.97	SELP1	Control	4.36	1.747	53.17	-.745	.456
			LEI	4.02	2.159	48.87		
		SELP2	Control	4.34	1.955	54.22	-1.106	.269
			LEI	3.88	2.197	47.84		
		SELP3	Control	5.42	1.727	51.86	-.304	.762
			LEI	5.27	1.856	50.16		
		SELP4	Control	4.02	1.900	54.27	-1.123	.261
			LEI	3.63	2.209	47.79		
		SELP5	Control	4.44	1.897	52.95	-.671	.502
			LEI	4.14	2.040	49.09		
		SELP6	Control	4.26	1.688	52.03	-.354	.723
			LEI	4.10	2.042	49.99		
		SELP7	Control	4.74	1.882	51.72	-.248	.804
			LEI	4.65	1.896	50.29		
		SELP8	Control	4.62	1.850	52.72	-.592	.554
			LEI	4.35	2.057	49.31		
		TSELP	Control	36.20	13.189	53.28	-.775	.438
			LEI	34.04	14.874	48.76		
Student Trust in Faculty (STIS)	.90	STIS1	Control	3.24	.591	47.72	-1.272	.203
			LEI	3.37	.631	54.22		
		STIS2	Control	3.32	.587	50.40	-.228	.819
			LEI	3.31	.707	51.59		
		STIS3	Control	3.10	.647	47.13	-1.481	.139
			LEI	3.27	.666	54.79		
		STIS4	Control	3.24	.591	46.76	-1.634	.102
			LEI	3.43	.575	55.16		
		STIS5	Control	2.88	.799	43.66	-2.795	.005**
			LEI	3.31	.616	58.20		

Student Trust in Faculty cont' (STIS)		STIS6	Control	3.30	.580	45.51	-2.095	.036*
			LEI	3.53	.612	56.38		
		STIS7	Control	3.04	.669	49.04	-.735	.462
			LEI	3.14	.693	52.92		
		STIS8	Control	3.11	.633	48.22	-1.032	.302
			LEI	3.22	.783	53.73		
		STIS9	Control	2.81	.748	48.64	-.872	.383
			LEI	2.90	.878	53.31		
		STISR10[a]	Control	3.39	.680	45.94	-1.972	.049*
			LEI	3.59	.726	55.96		
		STIS11	Control	3.34	.479	50.82	-.070	.944
			LEI	3.29	.701	51.18		
		STIS12	Control	2.78	.648	49.15	-.700	.484
			LEI	2.88	.739	52.81		
		STIS13	Control	3.08	.634	45.90	-1.944	.052
			LEI	3.29	.729	56.00		
		TSTIS	Control	40.63	5.286	45.46	-1.884	.060
			LEI	42.55	6.607	56.43		
Feeling Scale (FS)	.95	FS1	Control	6.34	1.780	48.80	-.763	.446
			LEI	6.67	2.269	53.16		
		FS2	Control	5.90	1.930	47.07	-1.351	.177
			LEI	6.43	2.402	54.85		
		FS3	Control	6.24	1.944	45.32	-1.952	.051
			LEI	6.98	2.319	56.57		
		FS4	Control	6.24	1.901	44.69	-2.171	.030*
			LEI	7.02	2.276	57.19		
		FS5	Control	6.32	2.045	43.84	-2.460	.014*
			LEI	7.25	2.180	58.02		
		FS6	Control	6.08	2.146	41.94	-3.111	.002**

		LEI	7.33	2.215	59.88		
	FS7	Control	4.56	2.251	46.99	-1.375	.169
		LEI	5.24	2.680	54.93		
	FS8	Control	4.96	2.312	42.67	-2.849	.004**
		LEI	6.27	2.736	59.17		
Feeling Scale cont'	FS9	Control	5.38	2.465	46.71	-1.471	.141
		LEI	6.02	2.753	55.21		
	FS10	Control	5.92	2.456	47.87	-1.072	.284
		LEI	6.27	3.232	54.07		
	TFS	Control	57.94	17.067	43.59	-2.517	.012*
		LEI	65.49	21.463	58.26		

**p< 0.01 level (2-tailed).
*p< 0.05 level (2-tailed).
a. This negative item was reversed before conducting statistical analysis.

Table 4

CM: Location, frequency and correspondence with Feeling Scale

CM Labels OCCASION (SLIDE#)	CONTENT	FEELING RATING	FREQUENCIES			
			Maximum # of CMs		Total CM	
			Control	LEI	Control	LEI
CM1 (1)	Introduction	-	0	1	0	1
-(2, 3)	Learning Errors*	-	0	2	-	3
CM2 (4)	Feeling Scale Intro	F1	-	-	-	-
CM3 (5)	Statistics Framework	-	1	2	3	8
CM4 (6)	Data set	F2	1	1	2	8
CM5 (7)	Central Tendency	-	1	1	2	1
CM6 (8)	Mode	F3	1	1	3	8
CM7 (9)	Mean	F4	1	2	2	24
CM8 (10)	Median	F5	1	2	3	9
CM9 (11)	Variability	-	-	-	-	-
CM10 (12)	Range	F6	1	1	3	5
CM11 (13)	Variance	F7	1	3	2	49
CM12 (14)	Example Calculation	-	1	2	8	20
CM13 (15)	Example Cont'd	F8	1	1	5	27
- (16)	Assessment**	F9	-	-	-	-
- (17)	Peer grading**	F10	-	-	-	-
- (18)	Surveys**	-	-	-	-	-
Total					**33**	**163**

*Slide 2/3 only available to LEI and not to control condition and therefore were not included in the control versus LEI CM analysis
** Slides 16-18 were not included in the CM analysis because these slides did not deliver any content that could have been confusing related to the Elementary Statistics lecture.

Table 5

Summative Assessment: Items Difficulty

Item	Condition	Frequency of Correct response	Phi		D (difficulty level)	
			Percent	Value	p	Description
1.	Control	40	80.0	.005	.961	Easy
	LEI	41	80.4			Easy
2.	Control	32	64.0	.028	.778	Moderate
	LEI	34	66.7			Moderate
3.	Control	35	70.0	-.175	.078	Moderate
	LEI	27	52.9			Difficult
4.	Control	43	86.0	.099	.321	Easy
	LEI	47	92.2			Very Easy
5.	Control	39	78.0	.005	.958	Moderate
	LEI	40	78.4			Moderate
6.	Control	40	80.0	-.088	.379	Easy
	LEI	37	72.5			Moderate
7.	Control	43	86.0	-.166	.096	Easy
	LEI	37	72.5			Moderate
8.	Control	14	28.0	.016	.875	Difficult
	LEI	15	29.4			Difficult
9.	Control	36	72	.006	.951	Moderate
	LEI	37	72.5			Moderate
10.	Control	40	80.0	.143	.150	Easy
	LEI	46	90.2			Very Easy

D (% correct): Very difficult (<20), Difficult ([20, 60]), Moderate ([61,80)), Easy ([80, 90)), Very Easy ([90,100]).

Table 6

Summative Assessment: Non-parametric Mann-Whitney U-test for independent samples and analyses per item

Item	Condition	Mean Rank	*Mann-Whitney U-test*	
			Z	p
1.	Control	50.90	-.049	.961
	LEI	51.10		
2.	Control	50.32	-.280	.779
	LEI	51.67		
3.	Control	55.35	-1.752	.080
	LEI	46.74		
4.	Control	49.43	-.988	.323
	LEI	52.54		
5.	Control	50.89	-.052	.958
	LEI	51.11		
6.	Control	52.90	-.875	.381
	LEI	49.14		
7.	Control	54.43	-1.657	.097
	LEI	47.64		
8.	Control	50.64	-.156	.876
	LEI	51.35		
9.	Control	50.86	-.061	.951
	LEI	51.14		
10.	Control	48.40	-1.434	.152
	LEI	53.55		
Total Score	Control	51.99	-.340	.734
	LEI	50.03		

$p<.05$

Table 7

Expected Grade: Distribution between Conditions

Condition	Letter grade	Grade Point Value	Frequency	Percent	Valid Percent	Cumulative Percent
Control	C-	1.7	4	8.0	8.0	8.0
	C	2	4	8.0	8.0	16.0
	C+	2.3	4	8.0	8.0	24.0
	B-	2.7	6	12.0	12.0	36.0
	B	3	14	28.0	28.0	64.0
	B+	3.3	6	12.0	12.0	76.0
	A-	3.7	8	16.0	16.0	92.0
	A/A+	4	4	8.0	8.0	100.0
LEI	C-	1.7	5	9.8	9.8	9.8
	C	2	3	5.9	5.9	15.7
	C+	2.3	9	17.6	17.6	33.3
	B-	2.7	6	11.8	11.8	45.1
	B	3	5	9.8	9.8	54.9
	B+	3.3	11	21.6	21.6	76.5
	A-	3.7	9	17.6	17.6	94.1
	A/A+	4	3	5.9	5.9	100.0

According to http://www.registrarsoffice.ualberta.ca/My-Personal-Records/Grades/Grading-SystemExplained.aspx, performance at the undergraduate levels is measured as Poor: 1.3 (D+), Satisfactory: 1.7-2.3 (C-, C, C+), Good: 2.7-3.3 (B+, B or B-), or Excellent: 3.7-4.0 (A-, A, A+).

Table 8

Background Cognitive Measures: Non-parametric Mann-Whitney U-test for independent samples

Variable	Condition	*M* (*SD*)	Mean Rank	Mann-Whitney U-test	
				Z	*p*
Expected Grade	Control	2.95 (.664)	51.70	-.240	.810
	LEI	2.90 (.697)	50.31		
GPA	Control	3.05 (.52)	50.18	-.284	.777
	LEI	3.07 (.49)	51.80		
Grade 12 performance in Mathematics	Control	78.74 (12.07)	54.62	-1.233	.218
	LEI	75.59 (12.65)	47.45		
Grade 12 performance in English Language Arts	Control	81.06 (8.01)	52.05	-.359	.720
	LEI	79.80 (11.02)	49.97		

$p < .05$

Table 9

Spearman Rho correlations between cognitive and affective variables

Condition	Scale	Grade 12 Math	Grade 12 English	GPA	Expected Grade	Summative Assessment	Total CM Labels
Control	**TIGO**	.250	.162	.307*	.101	.110	-.018
	TEGO	.222	.091	.333*	.320*	.185	.057
	TMGO	-.063	-.195	.191	-.037	.112	.041
	TPGO	.312*	.043	.323*	.334*	.128	.039
	TPAGO	.357*	-.022	.308*	.369**	.256	.201
	TSERL	.106	.180	.386*	-.005	-.378**	-.093
	TSELP	.557***	-.010	.125	.678***	.551***	.108
	TSTIS	.073	-.221	-.143	.324*	.126	.081
	TFS	.145	-.236	-.032	.275	.211	.020
Intervention	**TIGO**	.254	.234	.217	.247	.115	-.135
	TEGO	.218	-.102	-.135	.108	-.037	-.108
	TMGO	.003	.068	.273	-.022	-.122	-.034
	TPGO	.162	-.008	-.047	.331*	.099	-.148
	TPAGO	.053	-.147	-.230	.061	-.039	-.045
	TSERL	.008	.059	.268	.155	-.094	.046
	TSELP	.419**	.018	.018	.814***	.583***	-.447**
	TSTIS	-.023	.076	.030	.245	.312*	-.242
	TFS	.249	-.208	.004	.517***	.520***	-.547***

*** p< 0.001 level (2-tailed).
** p< 0.01 level (2-tailed).
*p< 0.05 level (2-tailed).

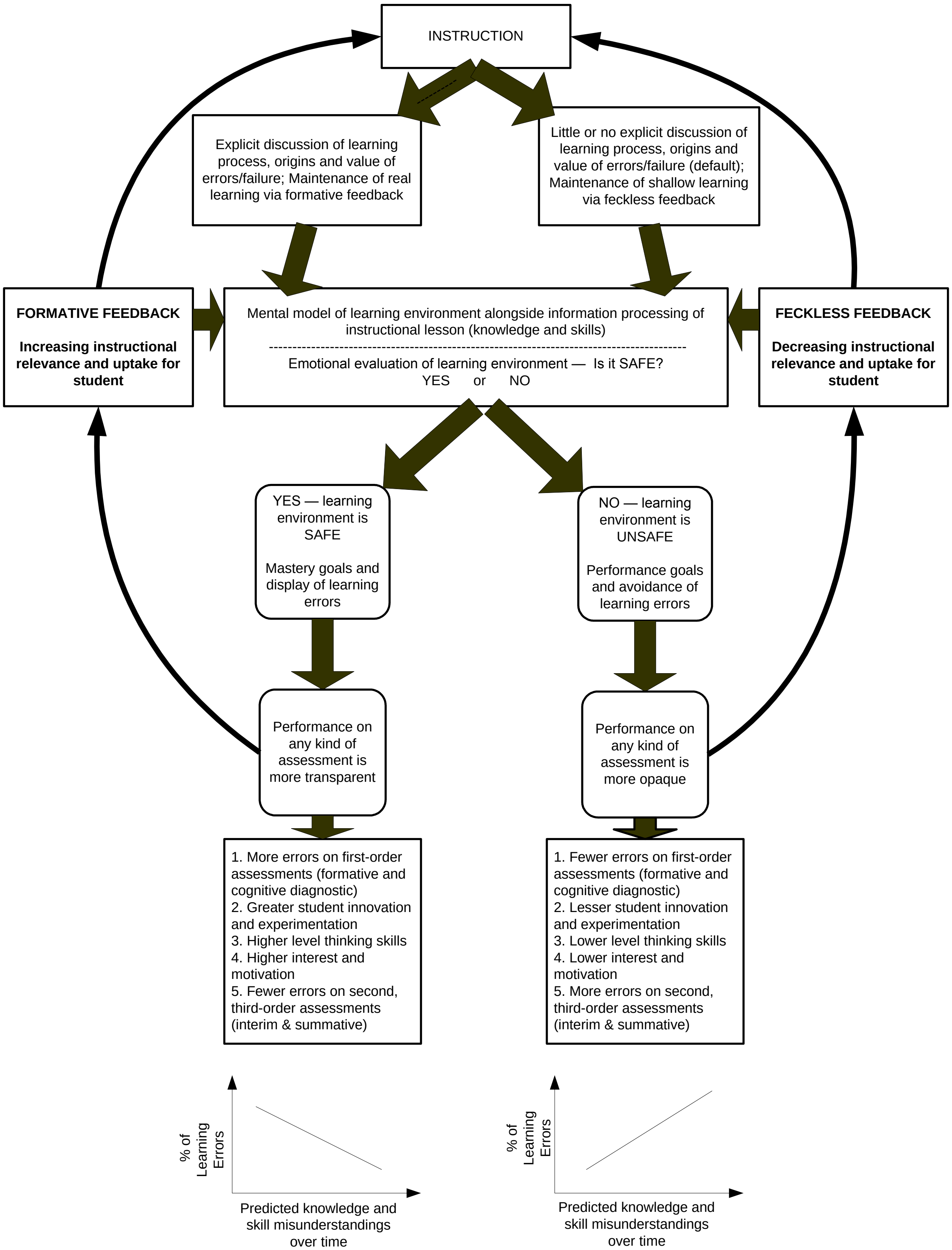


Figure 1. LEAFF Model (whole).

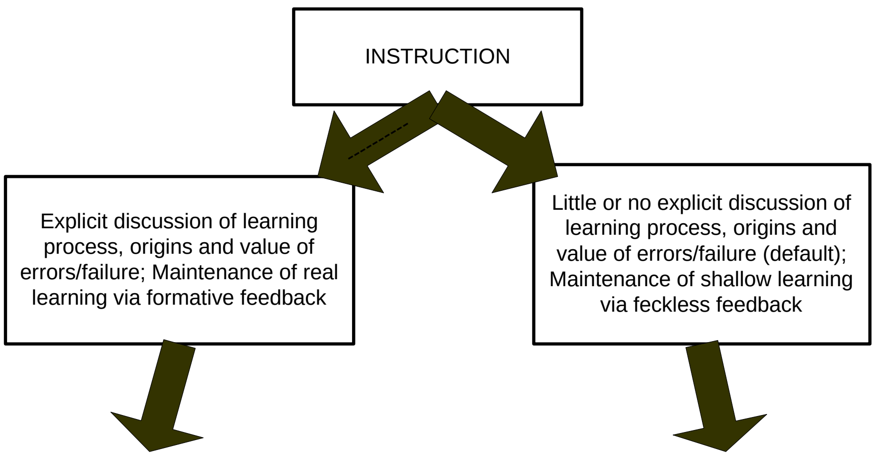


Figure 2. LEAFF Model Part One: Instructional Climate/Environment.

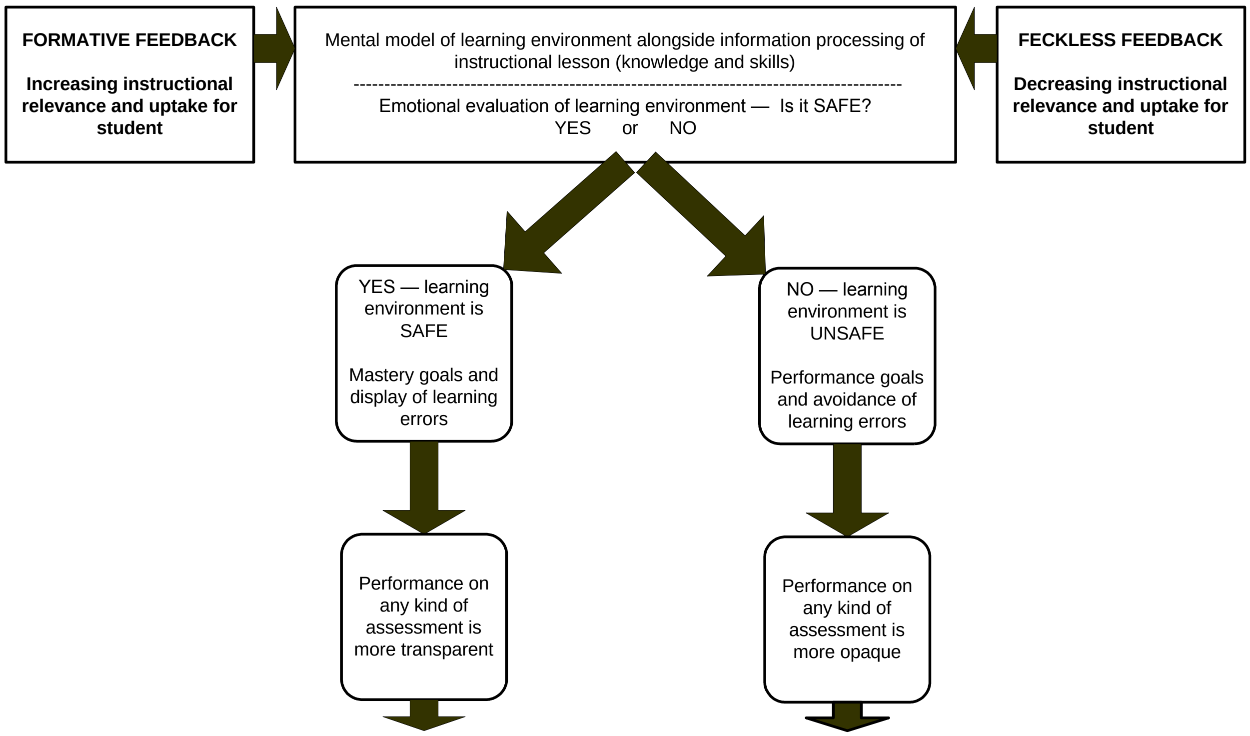


Figure 3. *LEAFF Model Part Two: Mental Models of Learning.*

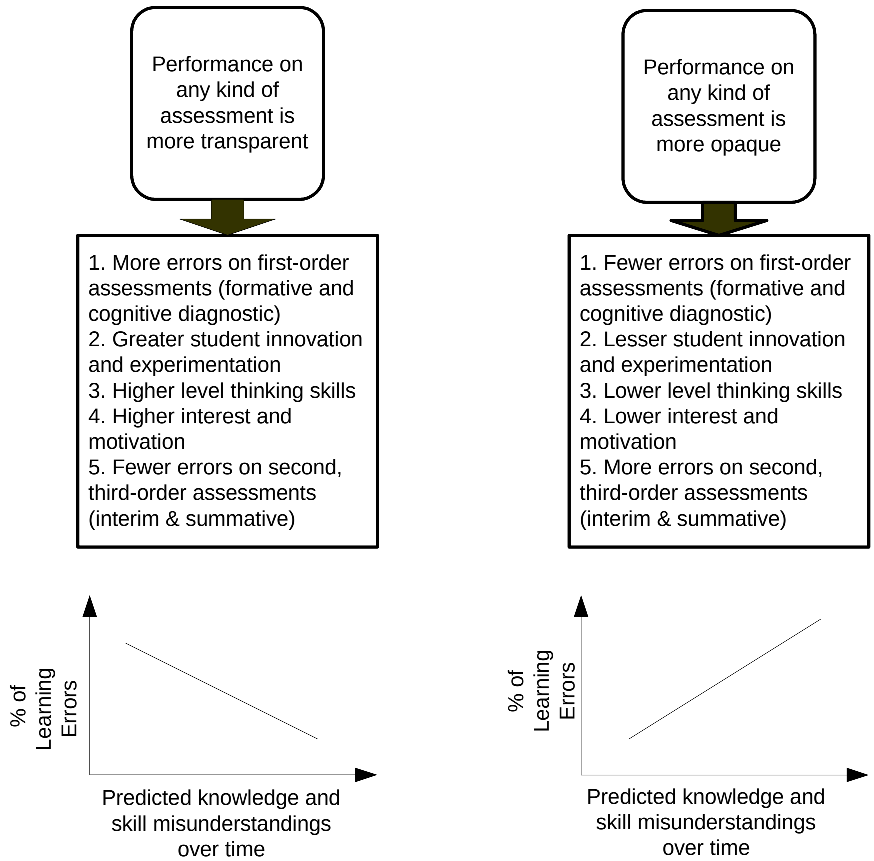


Figure 4. *LEAFF Model Part Three: Learning and Achievement Performance.*

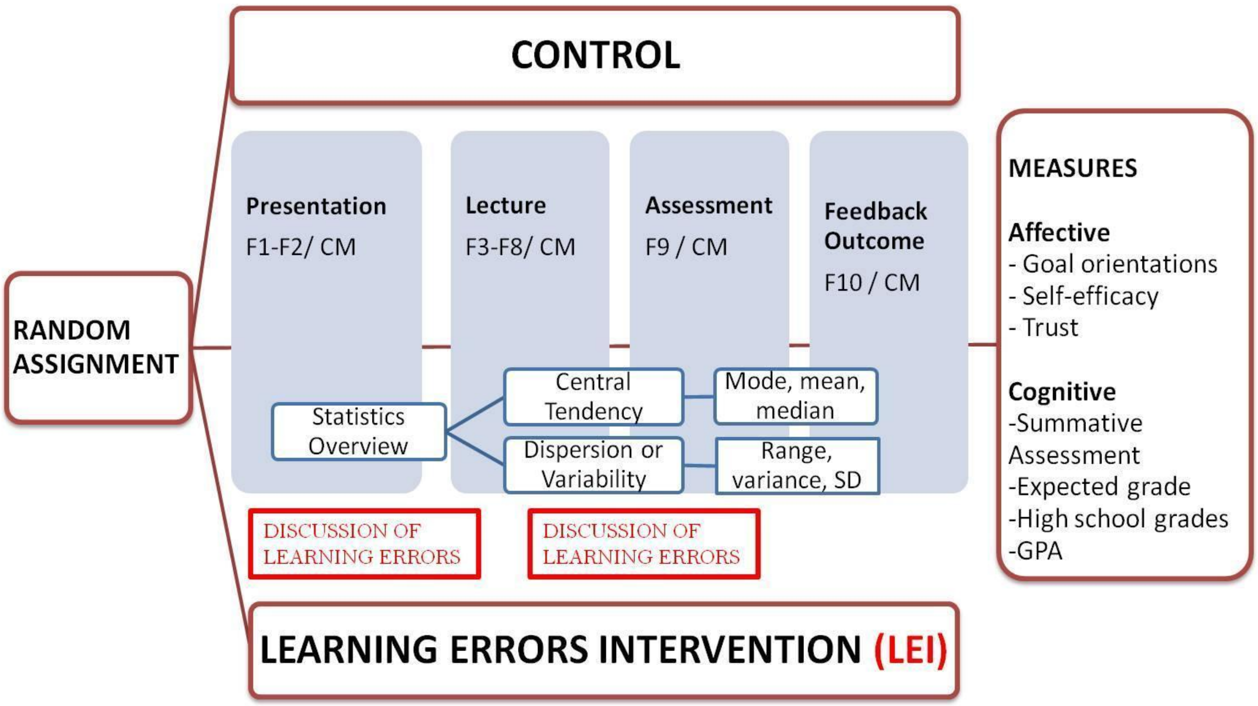


Figure 5. Instructional and Experimental Design Used in the Study.

What is our rationale?

- Learning is rewarding but it can be risky
- Learning = **not knowing** and then **coming to know**
- Psychologists tells us that, in most cases, we **need to make mistakes to learn** Why?
  - Mistakes help the brain clearly separate what is correct and incorrect
  - When we identify mistakes and where they can happen, talking about them can help us learn
  - *Remember a time when you made a mistake and it helped you learn something really well...*

Centre for Research in Applied Measurement and Evaluation 2

Figure 6. 2nd Slide LEI.

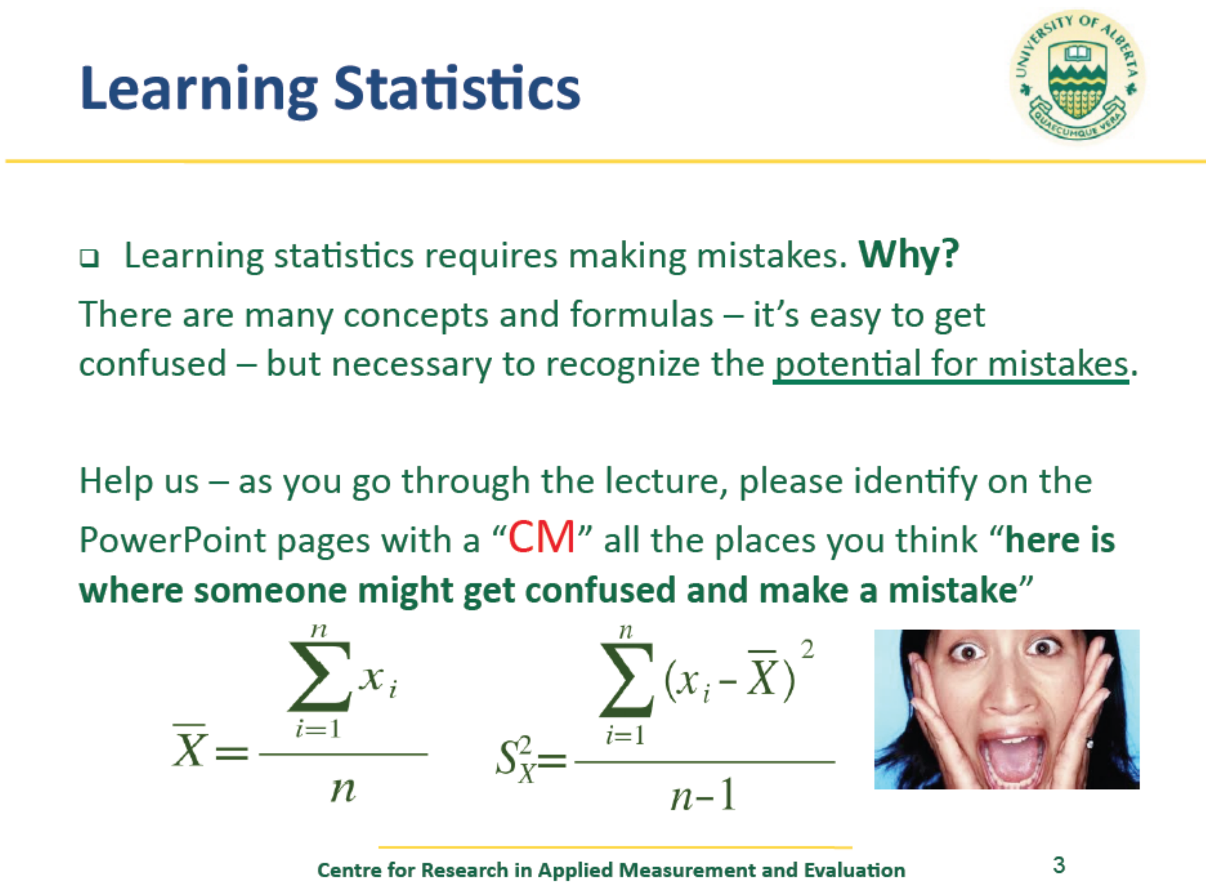


Figure 7. 3rd Slide LEI.

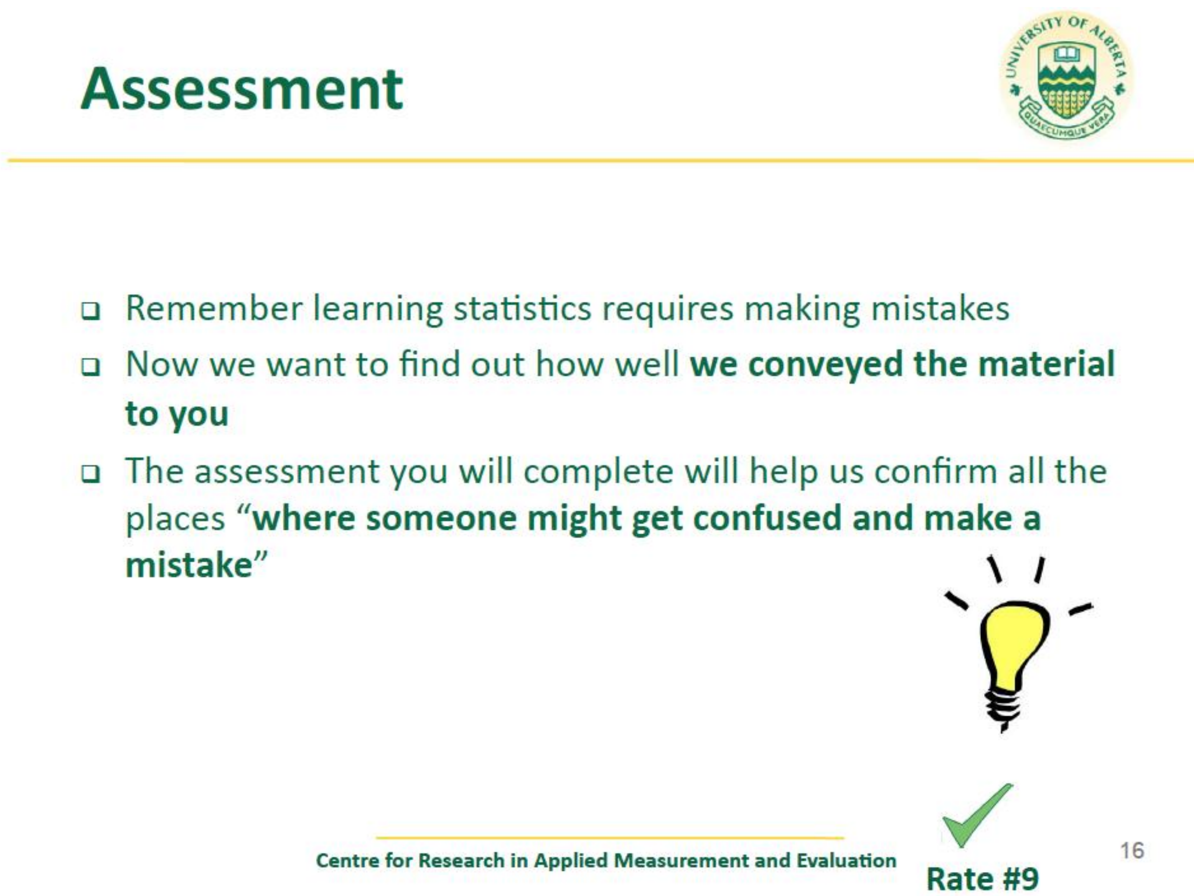


Figure 8. *16th Slide LEI.*

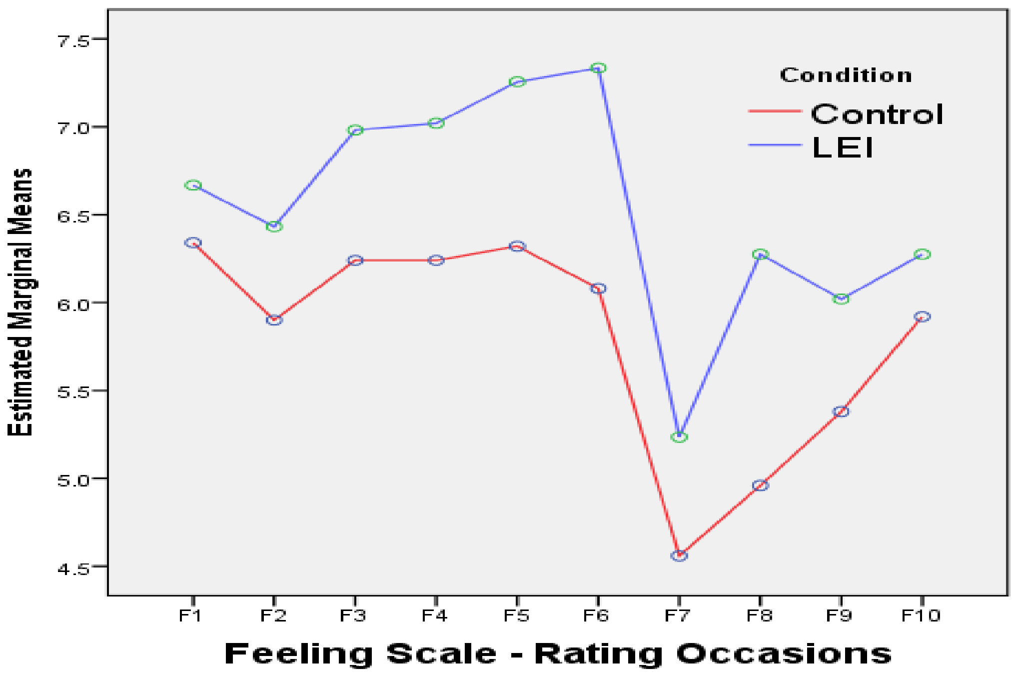


Figure 9. Feeling Scale Trends Across Conditions.

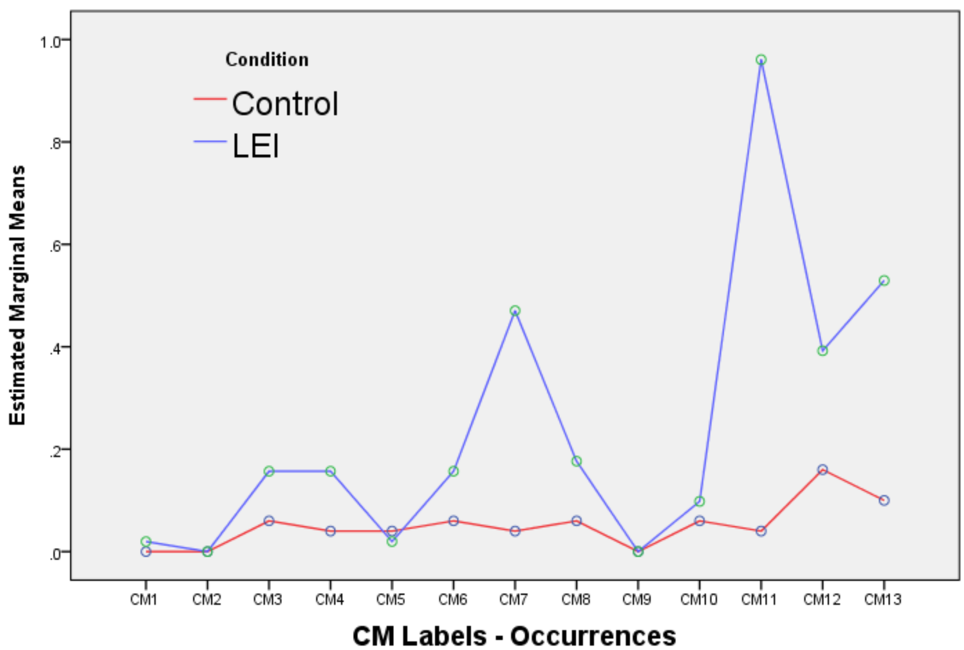


Figure 10. CM Trends Across Conditions (graph from the repeated measures analysis).

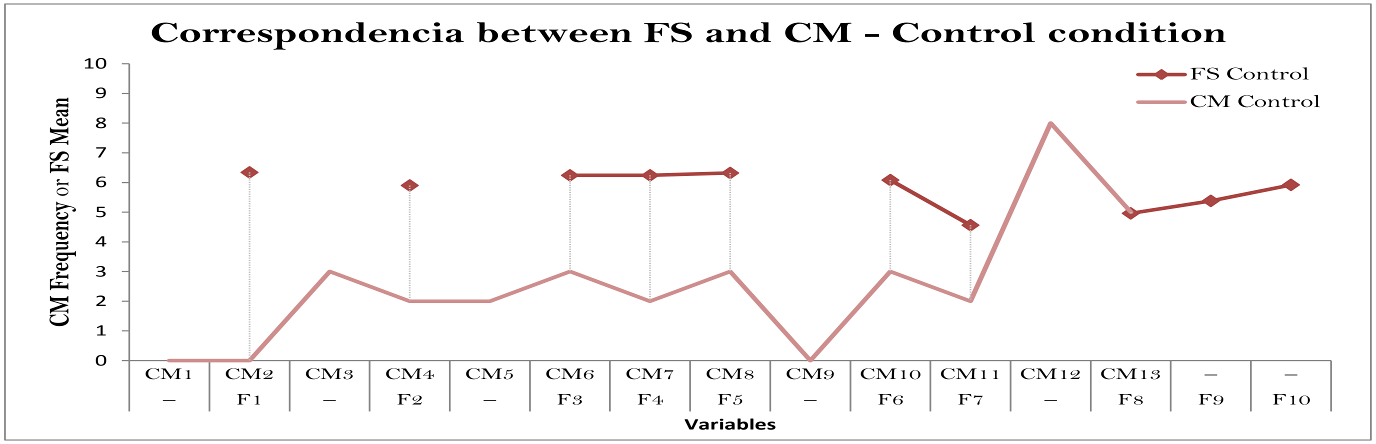


Figure 11. Feeling Scale and CM Labels in the Control Condition.

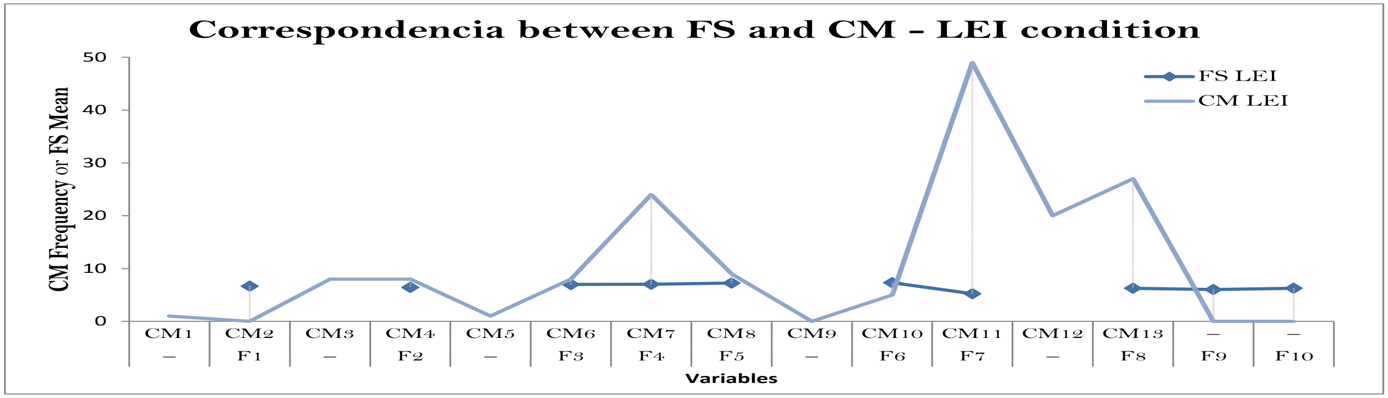
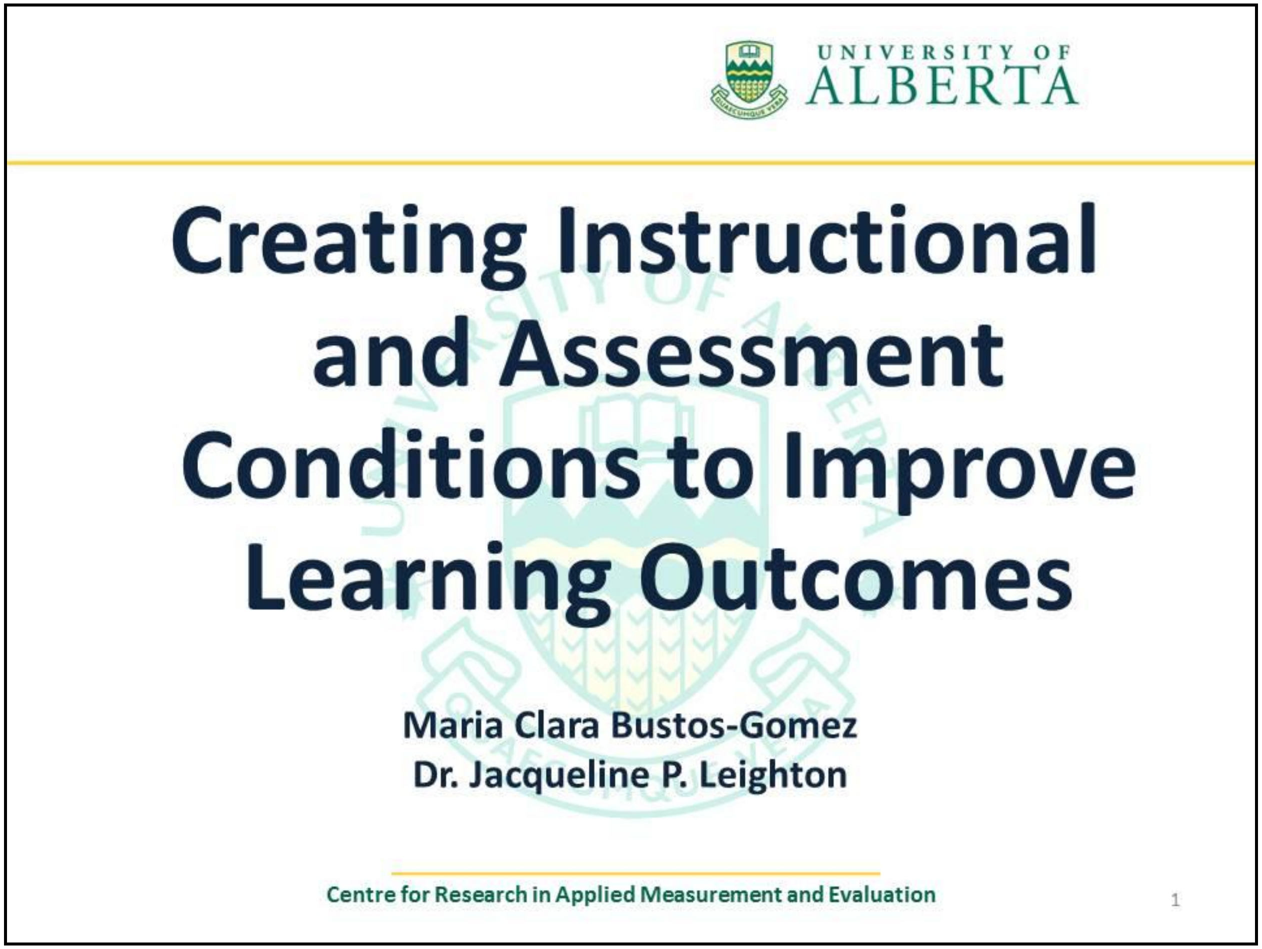


Figure 12. Feeling scale and CM labels in the Learning Errors Intervention (LEI) condition

Appendix A: Elementary Statistics Class PowerPoint Slides



What is our rationale?

- Learning is rewarding but it can be risky
- Learning = **not knowing** and then **coming to know**
- Psychologists tells us that, in most cases, we **need to make mistakes to learn** Why?
  - Mistakes help the brain clearly separate what is correct and incorrect
  - When we identify mistakes and where they can happen, talking about them can help us learn
  - *Remember a time when you made a mistake and it helped you learn something really well…*

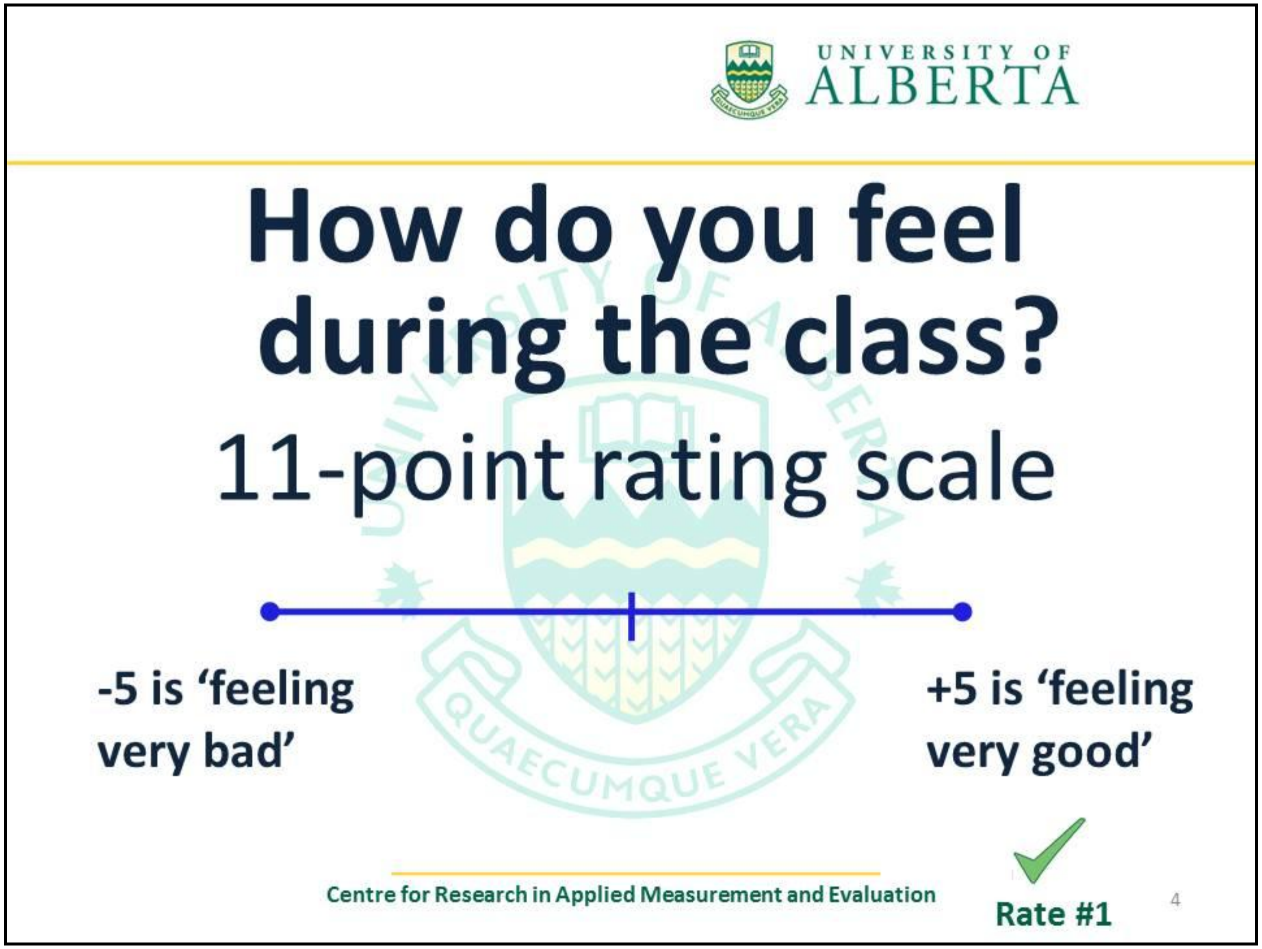
Learning Statistics

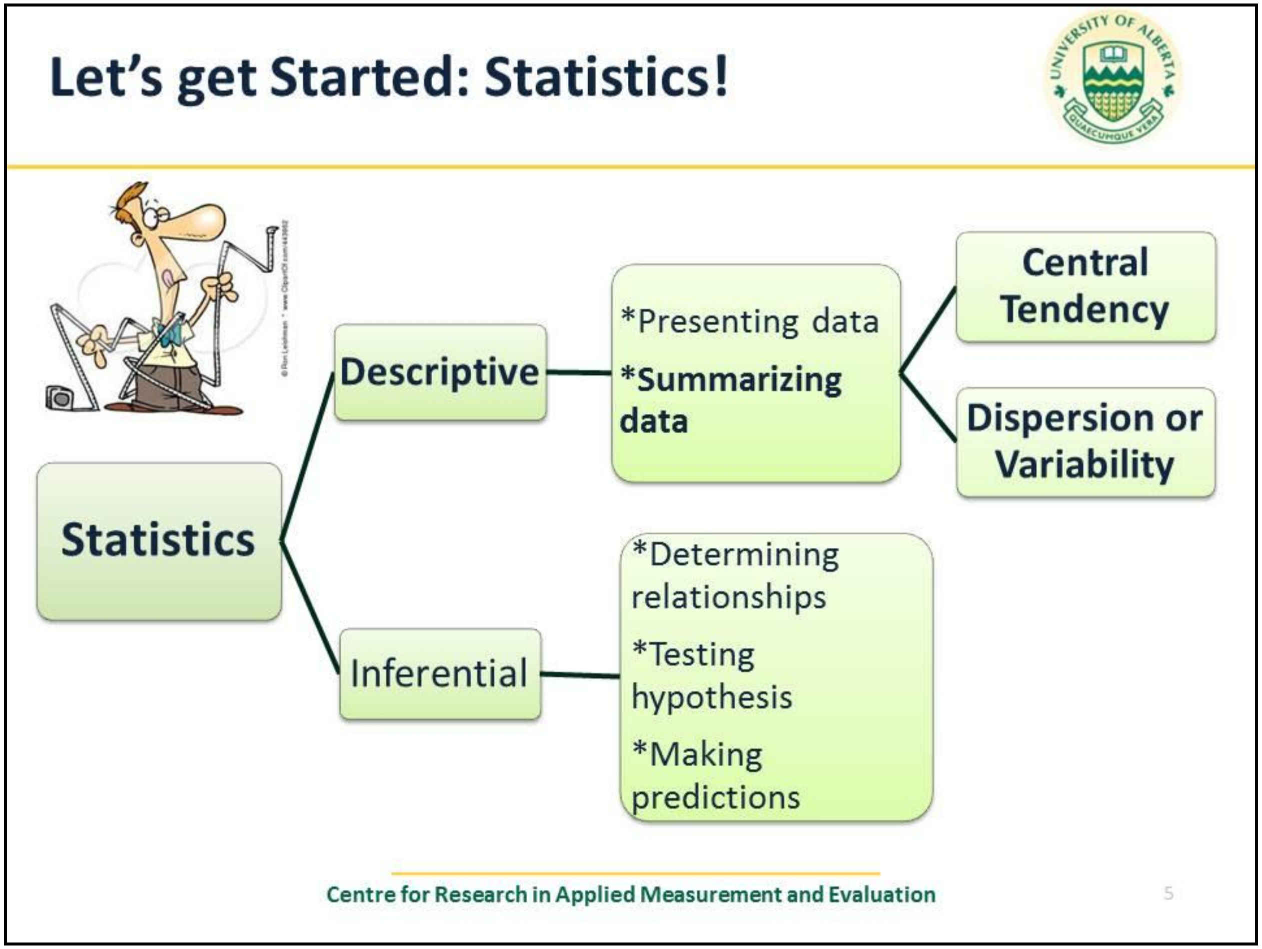
- Learning statistics requires making mistakes. **Why?**

There are many concepts and formulas – it's easy to get confused – but necessary to recognize the potential for mistakes.

Help us – as you go through the lecture, please identify on the PowerPoint pages with a "CM" all the places you think "**here is where someone might get confused and make a mistake**"

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^2}{n-1}$$

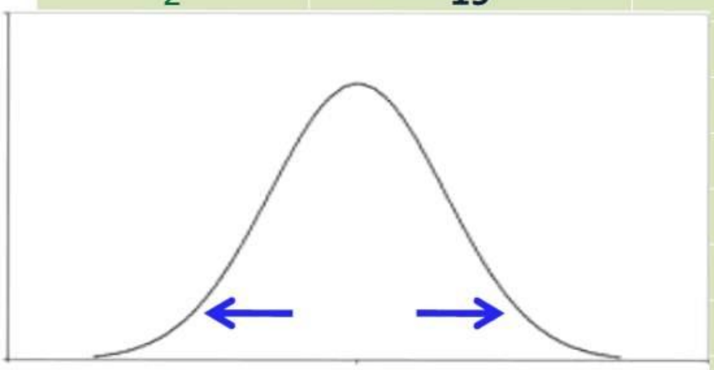

UNIVERSITY OF
ALBERTA
How do you feel during the class?
11-point rating scale
-5 is 'feeling very bad'
+5 is 'feeling very good'
Centre for Research in Applied Measurement and Evaluation
Rate #1
4


Let's get Started: Statistics!
UNIVERSITY OF ALBERTA
QUAECUMQUE VERA
Statistics
Descriptive
*Presenting data
*Summarizing data
Central Tendency
Dispersion or Variability
Inferential
*Determining relationships
*Testing hypothesis
*Making predictions
Centre for Research in Applied Measurement and Evaluation
5

Data set

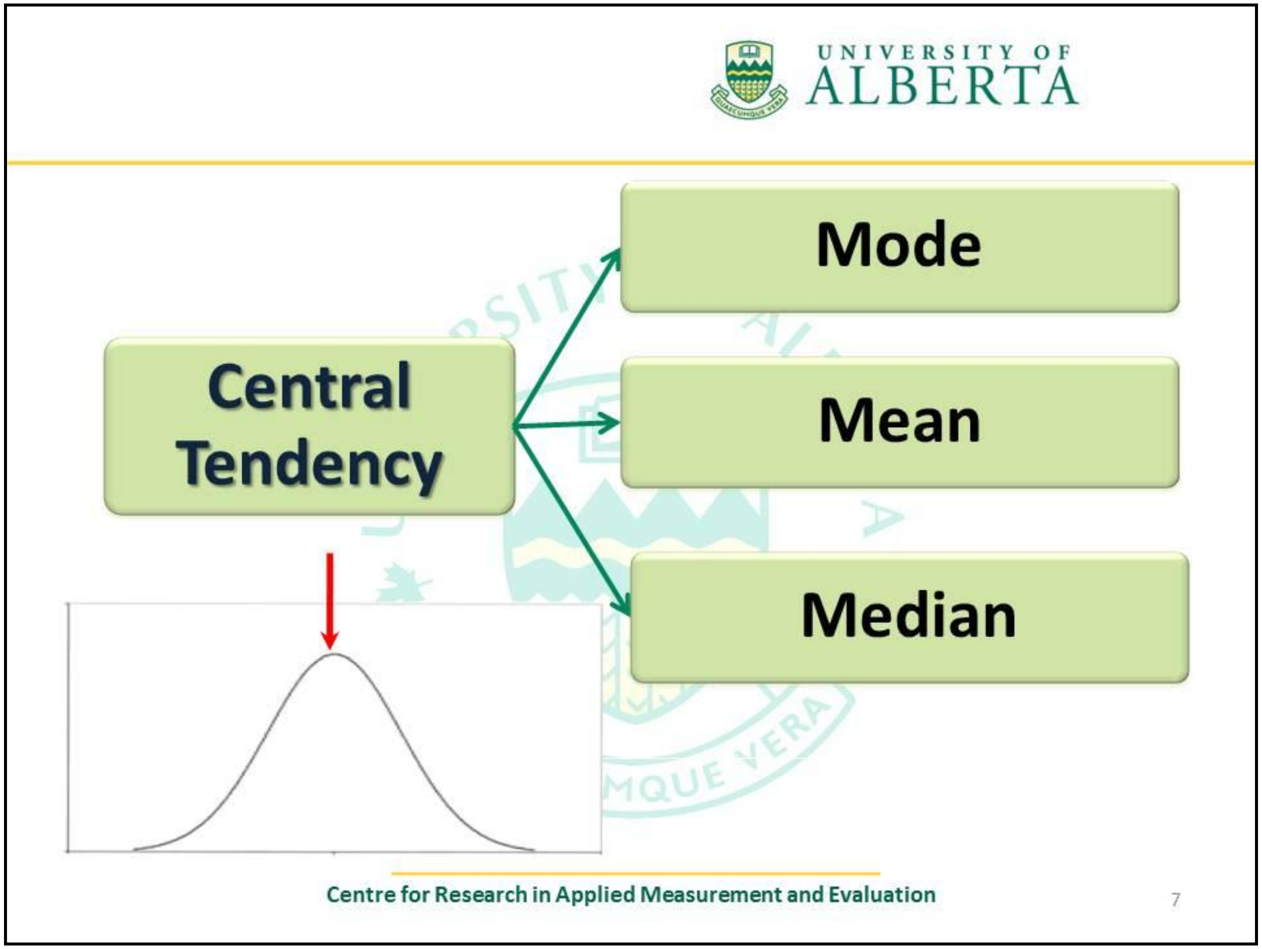
UNIVERSITY OF ALBERTA

Individual Number	VARIABLES		
	Age	Weight (Kg)	Height(Cm)
1	21	75	170
2	19	48	165
		60	160
		50	155
		55	165
		45	156
		68	160
		52	164
9	19	50	155
10	21	60	



Centre for Research in Applied Measurement and Evaluation

Rate #2


UNIVERSITY OF ALBERTA
Central Tendency
Mode
Mean
Median
Centre for Research in Applied Measurement and Evaluation
7

Mode

UNIVERSITY OF ALBERTA

N	Age
1	**21**
2	**19**
3	**20**
4	**21**
5	**19**
6	**21**
7	**20**
8	**21**
9	**19**
10	**21**

DEFINITION-NOTATION

- The **most frequently** occurring value for a variable.

FORMULA/PROCEDURE-EXPLANATION

- Frequency.
- Value(s) with the maximum frequency.

Rate #3

Mean

UNIVERSITY OF ALBERTA

N	Age
1	21
2	19
3	20
4	21
5	19
6	21
7	20
8	21
9	19
10	21

DEFINITION-NOTATION $\overline{X}$

- The value obtained from dividing in equal parts the total sum of the values that a variable takes.

FORMULA/PROCEDURE-EXPLANATION

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Rate #4

Median

UNIVERSITY OF ALBERTA

N	Age
1	**21**
2	**19**
3	**20**
4	**21**
5	**19**
6	**21**
7	**20**
8	**21**
9	**19**
10	**21**

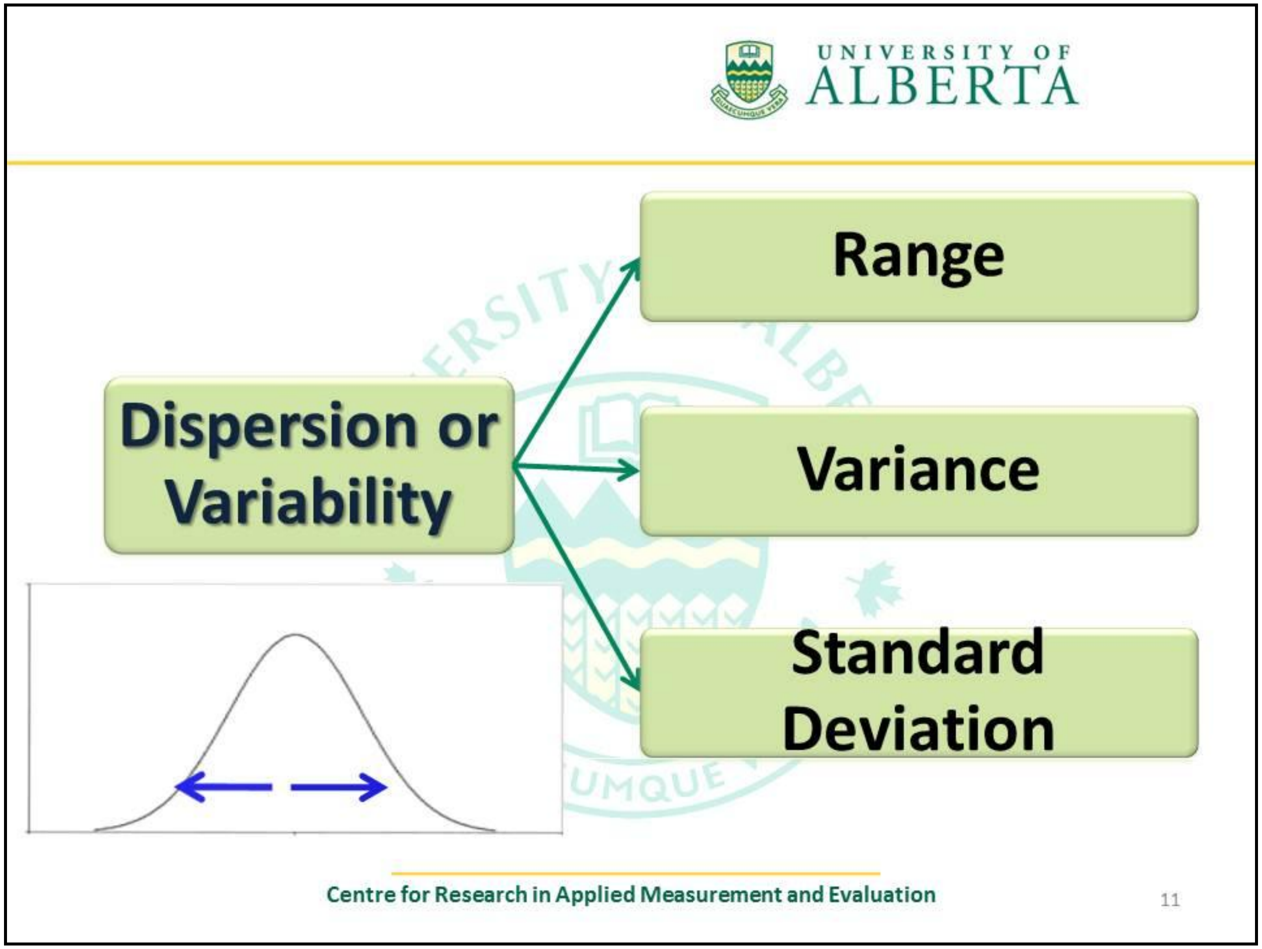
DEFINITION-NOTATION $\widetilde{X}$

- The value that 'lies' in the "middle" of the ordered distribution of a variable.

FORMULA/PROCEDURE-EXPLANATION

- Order the values.
- Find the middle value.

Rate #5


UNIVERSITY OF
ALBERTA
Range
Dispersion or
Variability
Variance
Standard
Deviation
Centre for Research in Applied Measurement and Evaluation
11

Range

UNIVERSITY OF ALBERTA

N	Weight (Kg)
1	75
2	48
3	60
4	50
5	55
6	45
7	68
8	52
9	50
10	60

DEFINITION-NOTATION

- The maximum distance that exists among the values of a variable.

FORMULA/PROCEDURE-EXPLANATION

- **Range** = $X_{max} - X_{min}$

Rate #6

Variance

UNIVERSITY OF ALBERTA

N	Height (cm)
1	170
2	165
3	160
4	155
5	165
6	156
7	160
8	164
9	155
10	150

DEFINITION-NOTATION S_X^2

- The mean square of the differences observed between each value of a variable and the mean.

FORMULA/PROCEDURE-EXPLANATION

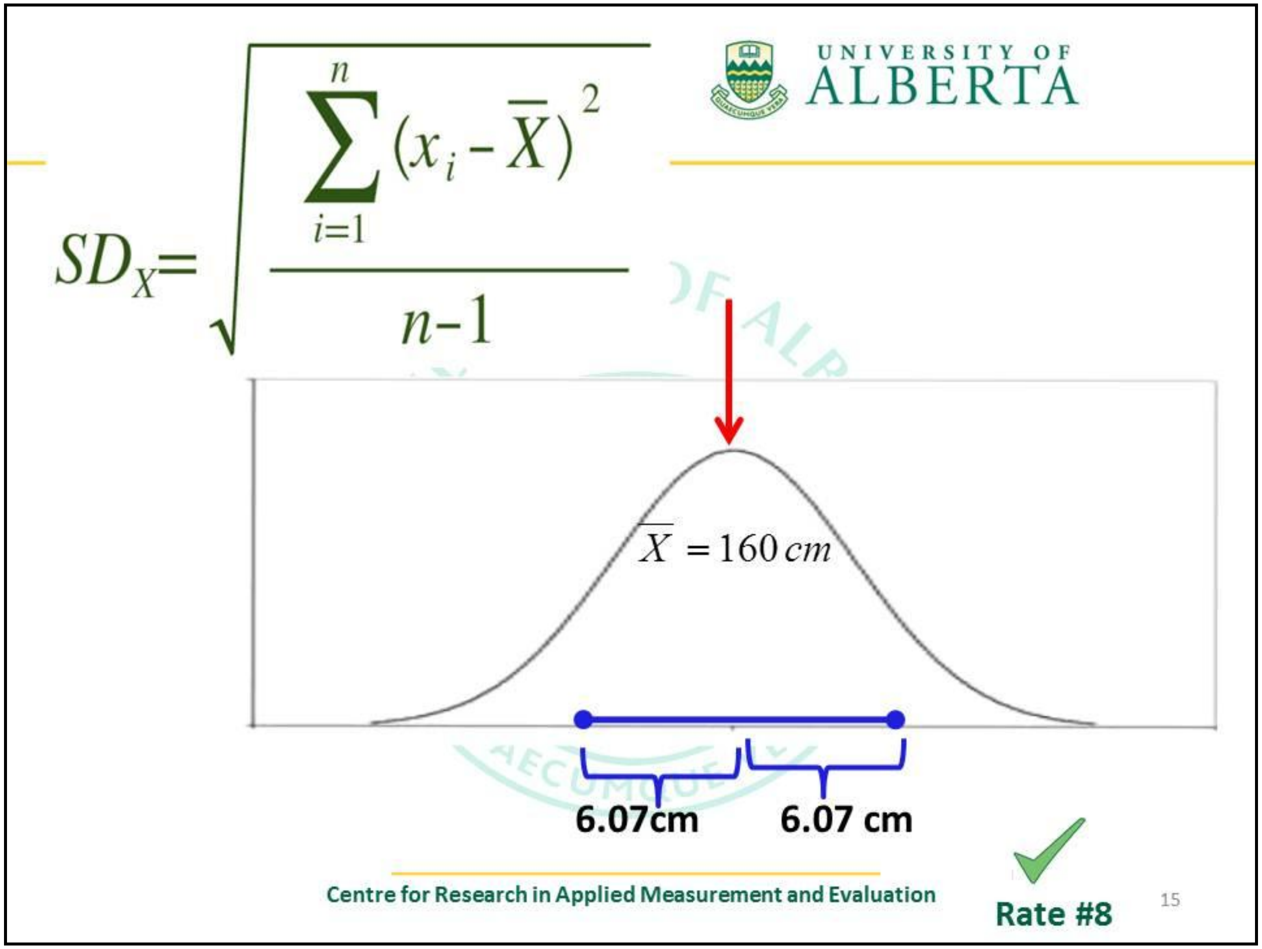
$$S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^2}{n-1}$$

Rate #7

What is the variance of the heights among the 10 individuals?

UNIVERSITY OF ALBERTA

N	Height (cm)	Deviations (cm)	Squares (cm^2)
1	170	10	100
2	165	5	25
3	160	0	0
4	155	-5	25
5	165	5	25
6	156	-4	16
7	160	0	0
1	164	4	16
9	155	-5	25
10	150	-10	100
Total	**1600**		**332**


UNIVERSITY OF ALBERTA
$SD_X = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n-1}}$
$\bar{X} = 160\,cm$
6.07cm
6.07 cm
Centre for Research in Applied Measurement and Evaluation
Rate #8
15

Assessment

- Remember learning statistics requires making mistakes
- Now we want to find out how well **we conveyed the material to you**
- The assessment you will complete will help us confirm all the places "**where someone might get confused and make a mistake**"

Rate #9

UNIVERSITY OF ALBERTA

Peer grading

Centre for Research in Applied Measurement and Evaluation

Rate #10

UNIVERSITY OF ALBERTA

Surveys

Thank you for your participation!

Appendix B: Feeling Scale [FS]

Feeling scale

As part of this study, we are interested in knowing how you are feeling at different moments during the class. Please rate your overall feeling of comfort each time the instructor ask you to do so. You will be ask to rate your feeling at a total of 10 different moments.

Feeling scale rate #1

I FEEL:

-5	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**
Very bad		Bad		Fairly bad	Neutral	Fairly good		Good		Very good

Note: Participants were given a set of 10 feeling rates, each on a single page.

Appendix C: Summative Assessment

For each of the following items, please select a single best answer in response to each question:

1. What is the central tendency method used to summarize the most common value(s) of a variable?
   a. Mean
   b. Mode
   c. Median

2. The blood types of 25 donors are given below.

AB	B	A	O	B
O	B	O	A	O
B	O	B	B	B
A	O	AB	AB	O
A	B	AB	O	A

What is the most appropriate central tendency method to summarize the data?
   a. Mean
   b. Mode
   c. Median

3. What is the most appropriate dispersion method to summarize the variable ‘Gender’?
   a. Range
   b. Variance
   c. Standard deviation
   d. None

4. A large standard deviation value indicates that:
   a. A given variable varies a lot.
   b. A given variable varies a little.
   c. A given variable does not vary.

5. A variable takes the values 1, 5, 1, 6, 1, 6, and 8. What is the median of the variable?
   a. 1
   b. 4
   c. 5
   d. 6

Use the following source to respond questions 6. and 7.

The table below shows the price of four cell phones:

Phone	Price
SJ-119	$49
SJ-220	$79
RV-375	$89
RV-400	$99

6. What is the mean price of the cell phones?
    a. $50
    b. $79
    c. $84

7. What is the median price of the cell phones?
    a. $50
    b. $79
    c. $84

8. There are two different set of blocks labelled X and Y. The lengths of the blocks in set X are 10, 20, 30, 40, 50, and 60 cm; the lengths of the blocks in set Y are 10, 10, 10, 60, 60, and 60 cm. Which set of blocks shows greater variability?
    a. X
    b. Y
    c. Both are equal

9. What is the relationship between the standard deviation and the variance of a variable?
    a. The standard deviation is obtained by squaring of the variance.
    b. The variance is obtained by taking the square root of the standard deviation.
    c. The standard deviation is obtained by taking the square root of the variance.

10. A package of chips says “mean weight of 28 grams”. If the standard deviation of the weight is 3 grams, which of the following values is less likely to be the weight for a package of chips randomly selected?
    a. 24 grams
    b. 26 grams
    c. 28 grams
    d. 30 grams

Appendix D: Surveys Booklet

Thank you very much for participating in this study on Creating Instructional and Assessment Conditions to Improve Learning Outcomes

Please ensure you've completed the consent form.

Please write your student code # below:

____

In the next 9 pages, you will see 9 short surveys. Please read the instructions and complete them. If you have any questions, please ask the researcher.

Survey #1

Intrinsic Goal Orientation

Using the scale below and **<u>thinking about your classes in general</u>**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

	Not at all true of me 1	**2**	**3**	**4**	**5**	**6**	**All true of me 7**
1. I prefer course material that really challenges me so I can learn new things.							
2. I prefer course material that arouses my curiosity, even if it is difficult to learn.							
3. The most satisfying thing for me in classes is trying to understand the content as thoroughly as possible.							
4. When I have the opportunity, I choose course assignments that I can learn from even if they don't guarantee a good grade.							

Survey #2

Extrinsic Goal Orientation

Using the scale below and **<u>thinking about your classes in general</u>**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

	Not at all true of me 1	**2**	**3**	**4**	**5**	**6**	**All true of me 7**
1. Getting a good grade is the most satisfying thing for me right now.							
2. The most important thing for me right now is improving my overall grade point average, so my main concern in my classes is getting a good grade.							
3. If I can, I want to get better grades in my classes than most of the other students.							
4. I want to do well in my classes because it is important to show my ability to my family, friends, employer, or others.							

Survey #3

Mastery Goal Orientation

Using the scale below and **thinking about your classes in general**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

1. It's important to me that I learn a lot of new concepts this year.

1	2	3	4	5
NOT AT ALL TRUE		**SOMEWHAT TRUE**		**VERY TRUE**

2. One of my goals in my classes is to learn as much as I can.

1	2	3	4	5
NOT AT ALL TRUE		**SOMEWHAT TRUE**		**VERY TRUE**

3. One of my goals is to master a lot of new skills this year.

1	2	3	4	5
NOT AT ALL TRUE		**SOMEWHAT TRUE**		**VERY TRUE**

4. It's important to me that I thoroughly understand my class work.

1	2	3	4	5
NOT AT ALL TRUE		**SOMEWHAT TRUE**		**VERY TRUE**

5. It's important to me that I improve my skills this year.

1	2	3	4	5
NOT AT ALL TRUE		**SOMEWHAT TRUE**		**VERY TRUE**

Survey #4

Performance Goal Orientation

Using the scale below and **thinking about your classes in general**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

1. It's important to me that other students in my classes think I am good at my class work.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

2. One of my goals is to show others that I'm good at my class work.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

3. One of my goals is to show others that class work is easy for me.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

4. One of my goals is to look smart in comparison to the other students in my classes.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

5. It's important to me that I look smart compared to others in my classes.

1	2	3	4	5
NOT AT ALL TRUE		SOMEWHAT TRUE		VERY TRUE

Survey #5

Performance-Avoid Goal Orientation

Using the scale below and **thinking about your classes in general**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

1. It's important to me that I don't look stupid in my classes.

1 2 3 4 5

NOT AT ALL TRUE SOMEWHAT TRUE VERY TRUE

2. One of my goals is to keep others from thinking I'm not smart in my classes.

1 2 3 4 5

NOT AT ALL TRUE SOMEWHAT TRUE VERY TRUE

3. It's important to me that my instructors don't think that I know less that others in my classes.

1 2 3 4 5

NOT AT ALL TRUE SOMEWHAT TRUE VERY TRUE

4. One of my goals in classes is to avoid looking like I have trouble doing the work.

1 2 3 4 5

NOT AT ALL TRUE SOMEWHAT TRUE VERY TRUE

Survey #6

Self-efficacy for self-regulated learning
Using the scale below and **thinking about how you normally behave in terms of your schoolwork,** please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

How well can you:	**Not well at all 1**	**2**	**3**	**4**	**5**	**6**	**Very well 7**
1. Finish homework assignments by deadlines?							
2. Study when there are other interesting things to do?							
3. Concentrate on school subjects?							
4. Take class notes of class instruction?							
5. Use the library to get information for class assignments?							
6. Plan your schoolwork?							
7. Organize your schoolwork?							
8. Remember information presented in class and textbooks?							
9. Arrange a place to study without distractions?							
10. Motivate yourself to do schoolwork?							
11. Participate in class discussions?							

Survey #7

Self-efficacy for learning and performance

Using the scale below and **thinking about your experience in this lecture today**, please rate the following items. Please answer all items, even if you are not sure. Please select only a single rating for each item.

	Not at all true of me 1	**2**	**3**	**4**	**5**	**6**	**All true of me 7**
1. I believe I could receive an excellent grade on an exam in this lecture.							
2. I'm certain I can understand the most difficult material presented in this lecture.							
3. I'm confident I can understand the basic concepts taught in this lecture.							
4. I'm confident I can understand the most complex material presented by the instructor in this lecture.							
5. I'm confident I can do an excellent job on the assignments and tests in this lecture.							
6. I would expect to do well in this lecture and future ones.							
7. I'm certain I can master the skills being taught in this lecture.							
8. Considering the difficulty of this lecture, the instructor, and my skills, I think I will do well in this lecture.							

Survey #8

Student Trust in Instructor Scale

Using the scale below and **thinking about your experience in this lecture**, please indicate how much you agree or disagree with each of the following statements. Please answer all items, even if you are not sure. Please select only a single rating for each item.

	Strongly Disagree	**Disagree**	**Agree**	**Strongly Agree**
1. The instructor of this lecture is/appears always ready to help.	1	2	3	4
2. The instructor of this lecture is/appears easy to talk to.	1	2	3	4
3. Students are/appear well cared for in this lecture.	1	2	3	4
4. The instructor of this lecture always does what he/she is supposed to.	1	2	3	4
5. The instructor of this lecture really listens to students.	1	2	3	4
6. The instructor of this lecture is/appears always honest with me.	1	2	3	4
7. The instructor of this lecture does a terrific job.	1	2	3	4
8. The instructor of this lecture is/appears good at teaching.	1	2	3	4
9. The instructor of this lecture has high expectations for all students.	1	2	3	4
10. The instructor of this lecture DOES NOT care about students.	1	2	3	4
11. Students at this lecture can believe what the instructor tells them.	1	2	3	4
12. Students learn a lot from the instructor in this lecture.	1	2	3	4
13. Students at this lecture can depend on the instructor for help.	1	2	3	4

Appendix E: Background Questionnaire

1. Please indicate your gender:

   ☐ Male ☐ Female ☐ I prefer not to respond

2. Please indicate your birth date: _______ (month)/_______ (day)/________(year)

3. Please indicate one or more of the following groups to which you self-identify in terms of ethnicity:
   - ☐ White
   - ☐ Chinese
   - ☐ South Asian (e.g., East Indian, Pakistani, Sri Lankan etc.)
   - ☐ Black
   - ☐ Filipino
   - ☐ Latin American
   - ☐ Southeast Asian (e.g., Cambodian, Indonesian, Laotian, Vietnamese, etc)
   - ☐ Arab
   - ☐ West Asian (e.g., Afghan, Iranian, etc.)
   - ☐ Japanese
   - ☐ Korean
   - ☐ Other ________________________

4. Please provide your final grade that you remember from Grade 11/12 Math in percentage: ____________

5. Please provide your final grade that you remember from Grade 11/12 English in percentage: ____________

6. In what program are you in? ____

7. What is your overall grade up to now at university courses? Mark one option.
   - ☐ A+/A
   - ☐ A-
   - ☐ B+
   - ☐ B
   - ☐ B-
   - ☐ C+
   - ☐ C
   - ☐ C- or lower

8. What is a potential grade you think could receive from this class?
   - ☐ A+/A
   - ☐ A-
   - ☐ B+
   - ☐ B
   - ☐ B-
   - ☐ C+
   - ☐ C
   - ☐ C- or lower