



# Can Group Instruction Facilitate Personalized Education? Revisiting Bloom's Learning for Mastery Model

## Citation

Rouhani, Parisa. 2017. Can Group Instruction Facilitate Personalized Education? Revisiting Bloom's Learning for Mastery Model. Qualifying Paper, Harvard Graduate School of Education.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33797239>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

**Can Group Instruction Facilitate Personalized Education?**

**Revisiting Bloom's *Learning For Mastery* Model**

Qualifying Paper

Submitted by

Parisa Rouhani

March 2016

## Table of Contents

---

Overview	4
Introduction	5
Goal of the Qualifying Paper	9
Part I: Learning for Mastery	11
Bloom's Model	11
<i>Theoretical foundation</i>	12
<i>Cognitive entry behaviors</i>	13
<i>Affective entry characteristics</i>	14
<i>Quality of instruction</i>	16
Analysis	19
<i>Meta-Analysis</i>	19
<i>Cognitive variables</i>	20
<i>Affective variables</i>	23
<i>Instructor variables</i>	24
Shift in Priorities	25
Subsequent Research	26
<i>Descriptive characteristics</i>	28
<i>General findings</i>	29
<i>Mastery level</i>	30
<i>Subject of instruction</i>	34
<i>Duration</i>	36
<i>Correctives</i>	38
<i>Summary of subsequent research</i>	41
Criticisms of Bloom's model	41
<i>Pace and rate of learning</i>	41
<i>Use of instructional time</i>	44
<i>Equity of educational opportunity</i>	46
Synthesis of Part I	46
Part II: Looking Forward	47
Design of the learning environment	48
Entry characteristics	51
Actionable feedback	53
Cooperative learning	55
Synthesis of Part II	57
Conclusion	58

References	62
Appendix	72
Table 1: Overview of studies in the analysis	72
Meta-Analysis by Kulik-Kulik, & Banger-Drowns	74
Differentiated Instruction	76
Keller's Personalized System of Instruction (PSI)	80
Brief Overview of Studies	83

## Overview

Group instruction has been a staple of the public education system since its inception. Schools were designed around teacher-paced approaches to classroom instruction, which allowed schools to accommodate an increasing number of students but also gave rise to achievement differences. Students whose learning does not conform to the pace of instruction struggle to keep up as the disparity between their understanding and that of their better-performing peers increases. In an effort to address achievement differences, a variety of mandates and incentives aimed at providing equitable learning experiences have been employed, but the phenomenon of achievement differences still persists (Hecht, Torgesen, Wagner, & Rashotte, 2001). Educational programs meant to customize curricula and personalize instruction have been implemented but these efforts do not utilize the structure of the existing system, often placing unreasonable expectations on teachers, making systematic implementation challenging and wide adoption unlikely. It is believed that the existing system is fundamentally flawed and must be completely overhauled, but attempting to do so brings with it countless risks and unintended consequences (Reville, 2015). Before investing effort and financial resources in completely redesigning schools, it is worthwhile to consider the extent to which the current structure of schools can nurture the learning of students. *In this paper I examine the feasibility of providing quality educational experience in which all students are supported to perform at high levels, within the group-based teacher-paced structure of the existing educational system. To*

achieve this goal, I first review Bloom's *Learning for Mastery* (LFM) model, a model of group instruction that showed promise in its time, and analyze nearly four decades of empirical studies on it. I then consider the applicability of Bloom's LFM model in today's classrooms and its effectiveness in light of advancements in research and technology. I propose four leverage points – design, entry characteristics, feedback, and cooperative learning – that, if appropriately implemented within the existing classroom structure, could create a personalized learning experience in which all students could be supported to perform at high levels within the context of group instruction.

### **Introduction**

Public education as we know it became prevalent in the 1800's and has changed relatively little since that time. Prior to the availability of public schooling, the common educational approaches involved private tutoring, or one-room school houses where children ranging in age were taught alongside one another under the supervision of an adult. In 1837, the secretary of education in Massachusetts, Horace Mann, began the "common schools" movement (adopted from the Prussian model), in which trained teachers provided group instruction through the class lecture method. The "common school" was widely accepted as a central force for societal well-being, spreading swiftly across the northern and mid-western states and was eventually adopted throughout the southern and far-western states. Attendance to elementary schools became compulsory across the nation by 1918 as public education became increasingly viewed as an effective

way to discipline children and turn them into judicious citizens (Peterson, 2010; Wagoner & Haarlow, 2002; “Only a Teacher,” n.d.; Education News, 2013). This approach to education, in which one teacher is mandated to pace a classroom of students through curricular material as a group, has been largely unchanged since its inception.

If group instruction is the defining characteristic of the educational system, its intolerable failure has arguably been persistent achievement differences. Students in the same classroom, receiving the same instruction from the same teacher, reach drastically different learning outcomes. Students who do not acquire a solid foundation of the fundamentals like the number line (Friso-van den Bos, Kroesbergen, Van luit, Xenidou-Dervou, Jonkman, Van der Schoot, & Van Lieshout, 2015), vocabulary (Ehri, & Rosenthal, 2010; Ehri, & Rosenthal, 2007), and reading early on have a hard time catching up (Chatterji, 2006). Small initial differences in performance persist and eventually become increasingly evident achievement disparities over time. This phenomenon has plagued public education since its inception, and factors contributing to these differences have been widely studied (Eccles, Wigfield, Harold, et al., 1993; Hecht, Torgesen, Wagner, & Rashotte, 2001).

One common explanation for why students differ in performance is the argument that there are innate differences in intelligence: some students are inherently smarter than others (Plomin, McClearn, Smith, et al., 1994). Proponents of this view believe that levels of achievement derive from student’s

innate capacities – fixed characteristics – that faithfully hold across contexts and cause some students to grasp challenging concepts better and more quickly than others. The other explanation for why students differ in outcomes is based on the environmental upbringing, be it the parenting approach a person was exposed to or the living conditions of poverty or affluence, that dictate academic performance (Goldberg, Prause, Lucas-Thomson, et al., 2008). Advocates of this view believe that environmental factors shape the student's habits and attitudes about learning, and these behaviors are reflected in student achievement levels. While this nature-and-nurture-debate suggests polarized views on what contributes to academic achievement, these views coincided in diminishing the importance of schooling in development and learning. Schools were not expected to nurture the minds of students and develop their intellect, but rather to provide technical information and ascertain the degree to which they acquire the material presented (Peterson, 2010) to rank them accordingly.

Benjamin Bloom had a different perspective: he believed that small learning experiences within the school environment could have a profound influence on student achievement, and that when provided appropriate learning conditions, all students can attain high levels of performance (Bloom, 1968). To prove this, he directed a seminal study that showed that under tutoring conditions student performance was 2-sigma (two standard deviations) better than students undergoing traditional large-group instruction (Anania, 1981; Burke, 1983). Bloom was able to establish that learning conditions – such as teaching



procedures, and feedback and corrective techniques – have a tremendous impact on student achievement. However, he also acknowledged that individual tutoring was not a viable solution for school-based learning. The education system was designed to around group-based classroom instruction, and any model of instruction that would be adopted would have to exist within the constraints of the existing design. With this in mind, Bloom directed his energies toward developing methods of group-based instruction that could produce learning outcomes similar to those observed in individual tutoring conditions (Bloom, 1984).

Advocates of school reform, such as School Reform Initiative, Mission School Transformation, Annenberg Institute for School Reform, and Federation for Community Schools, call for a transformation of the educational system, claiming that its existing structure is outdated and broken. Charter schools are emerging to provide alternative approaches to the conventional classroom structures that define the traditional public school system (National Alliance for Public Charter Schools, 2014). Instructional programs like differentiated instruction or personalized systems of instruction are adopted and adapted in an effort to provide students a more individualized educational experience (see Appendix for summaries of these methods, including the strengths and shortcomings of these programs). Technology is introduced as a way to offload teacher group instruction so as to free up classroom time for teachers to engage directly and discretely with struggling students (Brame, 2013; Zucker, 2009). While such approaches have merit and deserve to be studied in their own right,

they all make the fundamental assumption that the existing system is irredeemably flawed because group instruction is incapable of supporting the needs of individual learners.

Educational leaders believe personalization of education is a necessity, but there is a lack of clarity around what that actually means or how it would be systematically implemented (Cavanagh, 2014). Many believe it requires designing a new educational system (Reville, 2015; Bonk, Lee, Reeves, & Reynolds, 2015), but it is important to note that such large-scale change brings with it risks and a host of unintended consequences<sup>1</sup>. To attain optimal outcomes, in which students receive an equitable experience and reach high levels of performance, it is important to ask whether the fundamental structure of the educational system needs to be changed. Perhaps there are elements within the existing system that, if carefully adjusted, can allow us to reach our ultimate goal.

### **Goal of the Qualifying Paper**

The purpose of this qualifying paper is to examine whether it is possible to provide an individualized, equitable educational experience in which all students achieve the learning goals, while preserving the core structure of the school system: group-instruction. The research question I will be examining through the literature is:

*Are there methods of group instruction that can be implemented within the existing educational system to produce large gains for all students within a classroom?*

---

<sup>1</sup> An example of an unintended consequence can be found in the school choice movement, which was intended to increase educational equity, but instead elevated confusion and uncertainty so much that it led families to abandon the public school system (Gardner, 2010).

Specifically, I am going to look at whether Bloom's Learning for Mastery (LFM) model, or some variant of it, can produce the significant gains demonstrated by individualized tutoring conditions<sup>2</sup>. This is a promising model to examine as it was designed for situations in which instructional time with students is relatively fixed, and the classroom structure is restricted (Block & Burns, 1976). The answer to this question has practical consequences that matter for where we direct our efforts, how we spend research money, what kinds of technologies we invest in, the way we train teachers, and what kinds of pedagogies we use to support, nurture, and educate all students.

This qualifying paper is organized in two parts: the first is an analysis of Bloom's Learning for Mastery model, and the second is an examination of elements rooted in modern technologies and research that can be incorporated in his model. I will begin part one by providing an overview of Benjamin Bloom and his Learning for Mastery model. I will review a notable meta-analysis (Guskey & Piggott, 1988) that evaluates early research on his model, and then discuss subsequent research conducted after publication of the meta-analysis. I will describe criticisms of Bloom's model and conclude with a synthesis of both the pros and cons of the model. I will begin part two of the qualifying paper with an overview of four key aspects of Bloom's model that could be improved upon. I will then describe the research underpinning each of the four recommendations,

---

<sup>2</sup> The significant gains referred to here are 2-sigma: Anania (1981) and Burke (1983) found that under individual tutoring conditions, student performance was 2-sigma better than students undergoing traditional large-group instruction.

and discuss advancements in technology and frameworks developed since Bloom's model that could be incorporated to produce more systematic gains. I will conclude part two by addressing outcomes that could reasonably be anticipated as a result of the recommendations.

### **Part I: Learning for Mastery**

In this section I am going to introduce Bloom's Learning for Mastery model; I will describe its key tenants and their purported influence on academic achievement. I will then discuss the findings of a meta-analysis that evaluated empirical research on Bloom's model within the first two decades of its implementation. This will be followed by a brief discussion about the change in educational climate since the model's popularity, and I will then analyze recent research that has evaluated Bloom's model. My analysis will discuss findings from 14 papers examining Bloom's Learning for Mastery model, specifically examining mastery levels, subject areas, study durations, and correctives utilized. I will then explain major criticisms of Bloom's model, and conclude with a synthesis of the strengths and weaknesses of the model.

#### **Bloom's Model**

Many models of instruction offer suggestions for iterative improvements to the learning experience or even methodical frameworks for improving learning outcomes, but they do not account for the realities and limitations of classrooms within the existing school structure, and they fail to make concrete predictions about outcomes that can be systematically tested. Bloom, on the other hand,

offered a program of instruction based on a theory of learning with specific predictions about student achievement, affect, and speed of learning that can be implemented within existing classroom constraints.

Benjamin Bloom was a renowned teacher, scholar, and researcher in the field of education. He served on the staff of the Board of Examinations at the University of Chicago, where he received his Ph.D. in Education, and then he became the university examiner while holding an appointment as an instructor in the Department of Education, and serving as educational adviser to numerous nations, including the governments of Israel and India (Eisner, 2000). His theories on talent development and his research around mastery learning had a tremendous impact on every level of education throughout the world (Guskey, 2001).

**Theoretical foundation.** One of the great influences on Bloom's mastery learning theory was John B. Carroll, who challenged traditional beliefs about student aptitude and rate of learning (Bloom, 1976; Guskey, 2001). Student aptitude was generally viewed as the level to which a student could learn a subject of instruction – for example, that those with high aptitudes were considered good learners because they could learn complexities of a subject whereas those with low aptitude, or poor learners, could only learn its basic elements. Carroll proposed that with the appropriate amount of time and instructional quality, all students should be able to attain the criterion of achievement. He did not, however, discuss the issues around how to incorporate these provisions within the constraints of a classroom setting (Carroll & Spearritt, 1967).

Bloom was inspired by Carroll's ideas and extended Carroll's theoretical premises to develop a model of group-based classroom instruction that incorporated ideas of Carleton W. Washburne (1920), Henry C. Morrison (1926), Jerome Burner (1966), Burrhus F. Skinner (1954), and Patrick Suppes (1966), among others (Bloom, 1968; Hymel, & Dyck, 1993). He was interested in developing a strategy that would allow for the individuality of learners to be considered in the process of teaching (Bloom, 1968). Bloom's theory consisted of three tenants – cognitive entry behaviors, affective entry characteristics, and quality of instruction – whose interactions together would account for anywhere from 80% to as much as 90% of the variation in school achievement.

**Cognitive entry behaviors.** Cognitive Entry Behaviors (CEB) are comprised of the accumulated types of prerequisite knowledge, skills, and competencies that the learner has previously acquired and that are essential to the learning of the new task (Bloom, 1978, pg. 32). Instructors often assume that all students enter a course with a similar set of background knowledge that can be drawn upon for the tasks ahead, but such assumptions are unwarranted. Bloom theorizes that differences in achievement are not exclusively related to aptitude or ability, but rather, that the prior learning experiences specific to each student have resulted in differing amounts of relevant background knowledge, making the learning task differentially difficult for students who do not possess the assumed knowledge that is necessary. Cognitive entry behaviors refer to the prerequisite knowledge that a student must have in order to understand the content being

taught. Theoretically, without acquiring the relevant cognitive entry behaviors, it would be impossible to meet the criterion of achievement; therefore his model emphasizes their acquisition as an essential part of the learning process.

To determine the importance of cognitive entry behaviors, Bloom examined the relationship between a learner's prior history with a subject and the immediate learning of that subject. Tests of aptitude were used as proxy measures for CEB from longitudinal studies of achievement<sup>3</sup>; specific subtests of aptitude measures that were most closely aligned with the cognitive entry behaviors necessary for subsequent learning in a subject area were used, and examined against performance as measured by exam achievement scores at the end of the course. Aptitude tests correlated with later achievement in the vicinity of  $+.50$  to  $+.70$ , suggesting that cognitive entry behaviors account for between one-quarter to one-half of the variation in achievement (Bloom, 1976).

**Affective entry characteristics.** Affective Entry Characteristics (AEC) refer to the combination of interests, attitudes, and self-views that a student possesses when encountering a new explicit learning opportunity. Whether the content is approached with eagerness, agitation, or anxiety influences the student's emotional preparedness and self-confidence, which has an effect on how much effort the student will expend on learning (Bloom, 1978). Affective entry characteristics include the residual emotional history from previous learning

---

<sup>3</sup> Bloom indicates that Bracht & Hopkins (1972) and Payne (1963) are the studies in which certain relevant aptitude subtests were used as a proxy for cognitive entry behaviors.

activities either in specific subjects or in the general school environment, which influence the student's confidence and interest in the material, and ultimately affects how a student performs on that learning task. Bloom posits that a student's affect emerges from self-perception about prior history of achievement over a series of learning tasks. Positive experiences of success and accomplishment have positive effects on the student's affect and encourage further learning whereas a history of negative experiences and failure have a negative impact on the student's motivation and desire to continue. Although these repeated experiences reinforce the students' attitudes toward learning, Bloom argues that affective entry characteristics can actually be altered.

To determine the predictive value of affective entry characteristic, Bloom examined studies<sup>4</sup> on achievement (determined by grade point average or a composite of achievement tests) and students' academic self-concept (defined as the perception of how the student is doing compared to peers), which was used as a proxy for affective entry characteristics<sup>5</sup>. The studies showed correlations of about +.50 for students in grade 5 and above, suggesting that AEC accounts for as much as 25% of variation in school achievement (Bloom, 1986, pg. 95).

A student's learning history is shaped by cognitive and affective experiences, so there is overlap between the predictive effects of cognitive and

---

<sup>4</sup> Bloom specifies that when considering affective entry characteristics, he examined studies by Kifer (1973), Dolan (1974), Malpass (1953), and Crosswhite (1972).

<sup>5</sup> Academic self-concept was used as an index of affective entry characteristics for prediction purposes because it was found that combining two or three key components that make up affective entry characteristics (subject affect, school affect, and academic-self concept) resulted in the same composite relation with school achievement as academic self-concept alone. For diagnostic purposes, however, subject-related affect, school-related affect, and academic self-concept are all to be considered when determining approaches to increase a student's positive affective entry characteristics (Bloom, 1978, pg. 97).



affective entry characteristics. While consideration to affect is an independent component of the model, the contribution of affective entry characteristics is thought to be relatively little over what is already contributed by cognitive entry characteristics. Multi-year longitudinal studies of achievement, which measured a combination of affective and cognitive entry characteristics found that the general estimate of both combined is  $+0.80$ , suggesting, CEB and AEC together account for about 65% of variance on achievement measures (Bloom, 1987, pg. 169).

**Quality of instruction.** To identify what constitutes quality of instruction, Bloom considered the ways in which practices used by good tutors could be applied to group settings such that instruction is effective for all individuals in the class. Bloom defined *Quality of Instruction* (QI) as the cues, reinforcement, feedback, and correctives that a student is given during the course of a learning task. Cues involve communication from the teacher or instructional material of what is to be learned and how the learner should proceed. Reinforcement and feedback provide the learner with information about how the learning is progressing, either through recognition of a correct assertion, or elucidation of an incorrect one. Correctives are alternative instructional materials that explain specific aspects of the content in different ways to elaborate and elucidate the initial group instruction (Bloom, 1978; Block & Burns, 1976).

Good quality of instruction provides clarity, offers appropriate support and reinforcements, and guides the student through the learning process. Emphasis is placed on the teaching (for example, the way teaching approaches are used, not

necessarily the qualities of the teacher), and the classroom environment of learning (for example, the learning atmosphere, not the physical characteristics of the classroom) so that students can be provided with appropriate learning conditions to reach high levels of achievement. The environment of learning is a product of the quality of instruction.

Bloom theorized that a lack of the necessary cognitive entry behaviors at the start of a task cannot be overcome by quality of instruction (unless the instruction is focused specifically on remedying the deficiencies in prerequisite knowledge), but that good quality of instruction can overcome the negative affective characteristics that a student may possess at the start of the task. The quality of instruction would help facilitate a high level of achievement and contribute to positive learning experiences that influence affective entry characteristics for subsequent learning tasks.

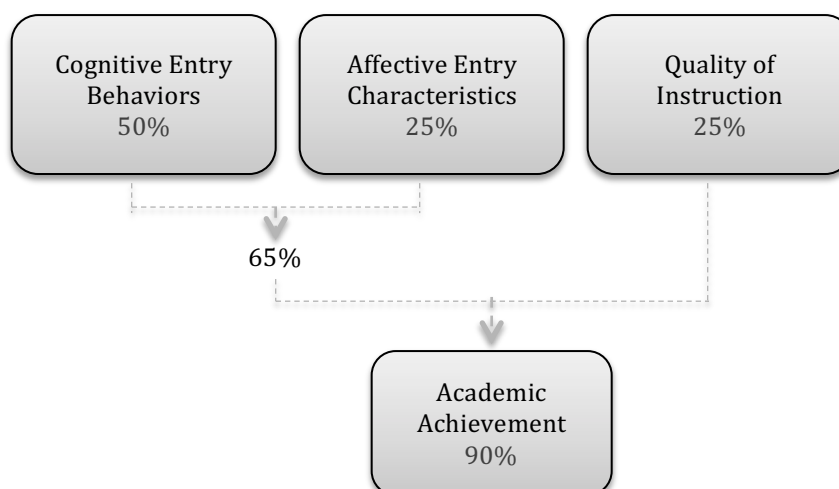
Quality of instruction has the greatest direct effect on student engagement and ultimately what is learned within a unit of study. Bloom's estimate of the effects of quality of instruction is through the examination of obvious and subtle participation of students in the learning process. Several studies<sup>6</sup> showed an average of correlation of  $+0.50$  to  $+0.60$  between student participation and achievement (measured by final formative exam or summative exams scores); this

---

<sup>6</sup> Block (1970), Anderson (1973), and Ozcelik (1974) were identified as the studies that produced correlations of  $.50$  to  $.60$  between student participation and achievement.

suggests that quality of instruction can account for at least 25% of variance on achievement measures (Bloom, 1987, pg133, pg187).

Bloom's model predicts that a learner's cognitive and affective history together account for about 65% of variance on achievement measures, and when combined with quality of instruction, is theorized to account for 90% of variance of achievement (Bloom, 1987, pg. 169). Bloom's model of student and teacher factors and the extent to which they contribute to learning outcomes is demonstrated in Figure 1.



*Figure 1.* Estimated effect of cognitive entry behaviors, affective entry characteristics, and quality of instruction on variation in academic achievement. Numbers indicate the percent of variance accounted for by the selected variables in academic achievement. Adapted from *Human Characteristics and School Learning* (p. 169), by B. S. Bloom, 1978, New York, NY: McGraw-Hill.

## Analysis

Bloom's specific predictions were supported by early empirical research conducted in part by his graduate students<sup>7</sup> and described in his book on *Human Characteristics and School Learning* (1976). The bold predictions that Bloom asserts about his learning for mastery model captured the interest of researchers and educators and prompted numerous experiments designed to either replicate his success or falsify his theory. I will now discuss the early research on Bloom's model through the examination of a comprehensive meta-analysis<sup>8</sup> reviewing mastery learning in the first 20 years after its inception.

**Meta-Analysis.** Guskey and Pigott (1988) conducted a review of 46 studies on mastery learning programs, critically examining cognitive learning outcomes, student affect, and teacher variables (such as the teacher's attitude towards the mastery learning process). Strict selection criteria required that studies involved teacher-paced group instruction, that mastery learning treatment classes be compared to conventionally taught control classes (or that there be a time-series design), and that experiments be free from serious methodological flaws<sup>9</sup>. The analysis examined many variables, but for the sake of brevity, I have

---

<sup>7</sup> Bloom validates his model through the findings of studies conducted by his doctoral students, but much of the evidence that supports his model is from unpublished dissertations and is therefore not readily available for further scrutiny.

<sup>8</sup> The meta-analysis described is by Guskey and Pigott (1989) which examines studies that implement Bloom's mastery model. There was another meta-analysis on mastery learning that was published in 1990 by Kulik, Kulik, and Banger-Drowns, however, it confounded Bloom's group-based mastery model with Keller's personalized system of instruction that uses mastery criterion and did not report on results of the two different models separately. All of the studies testing Bloom's model in this review overlapped with the review from two years prior – every study on Bloom's model in this meta-analysis were already included in the meta-analysis by Guskey and Gates (1988). I, therefore, excluded this review from my analysis; however a summary of this analysis can be found in the appendix.

<sup>9</sup> Guskey and Pigott specify that effect sizes were calculated in order to quantify outcomes of the studies – when means and standard deviations were reported for both treatment and control groups, the difference between the means of treatment and control groups was divided by the standard deviation of the control group as suggested by Glass (1976). For time-series

limited my summary to cognitive variables (measured by student achievement and re-analyzed by subject area, grade level, and duration of study), affective variables looking at student affect and engagement, and instructor variables.

Achievement outcomes were most commonly measured by test scores on teacher-prepared examinations or by letter grade distributions. Retention of learned content was measured by re-testing students a few weeks or months after instruction was completed. Time-related variables measured the amount of time spent involved in instruction, amount of time spent on-task, or student's class attendance and attrition rates. Affect toward school was measured by examining student self-reported affect toward a subject, affect toward school, academic self-concept, grade expectations, and attribution assignments. Teacher variables were measured by examining teacher's expectations for student learning and teacher attitude towards the mastery learning process.

**Cognitive variables.** *Student Achievement:* Seventy-eight effect sizes were produced from the forty-three articles that measured student achievement because many articles contained studies of multiple grade levels or subject areas. Nearly all of the 78 effect sizes reported indicated that the application of group-based mastery learning strategies has positive effects on student achievement; however, the size of the effect varied considerably by study. A test of homogeneity done on the entire collection of studies produced a homogeneity statistic of 759.50 ( $df =$

---

designs or cases where results were not fully reported, effect sizes were calculated from  $t$  or  $F$ -statistics (per Glass, McGaw, & Smith, 1981; Hedges, & Olkin, 1985), and, to prevent possible overestimation of the true population effect size, values were corrected for bias per the procedure described by Hedges (1981) before conducting further analyses.

77,  $p < .001$ ), indicating that variation across the studies is larger than expected of studies that shared an underlying effect size: the studies have fundamentally different effect sizes. To determine what factors produced similar effect sizes, Guskey and Pigott examined the data by subject area, grade level, and length of the study.

*Subject of instruction:* Subject areas were broken into five categories (Psychology, Science, Social Studies, Language Arts, and Mathematics), and homogeneity statistics ( $H$ ) were calculated for each category to determine whether studies in that category shared an underlying effect size. Psychology, which included general and educational psychology classes, produced a weighted effect size of .41 ( $H = 39.18$ ,  $df = 5$ ,  $p < .001$ ). Science, which included biology, chemistry, and general science, produced a weighted effect size of .50 ( $H = 13.28$ ,  $df = 9$ ,  $p = .150$ ). Mathematics, consisting of algebra, cartography, fractions, geometry, graphs, probability, and general math, yielded a weighted effect size of .70 ( $H = 230.98$ ,  $df = 35$ ,  $p < .001$ ). Language Arts, comprised of English, grammar, reading, vocabulary, and foreign language, produced a weighted effect size of .60 ( $H = 326.02$ ,  $df = 13$ ,  $p < .001$ ). Social Science, which included economics, education, government, history, humanities, and general social studies, produced a weighted effect size of .53 ( $H = 18.19$ ,  $df = 8$ ,  $p = .033$ ). Although there was a great deal of variance among studies within these categories, a test of the overall differences among subject areas indicated that the effect size differs by subject ( $H_b = 127.73$ ,  $df = 4$ ,  $p < .001$ ).

*Grade level:* The grade level of the studies were broken into three groups: *elementary* included grades 1 through 8, *high school* included grades 9 through 12, and postsecondary studies were classified as *college*. While mastery learning had positive effects on all levels of education, the effects were larger for younger students. The weighted average effect size of the twenty-five effect sizes calculated in the studies of the elementary students was .94; of the sixteen effect sizes calculated in the studies involving high school students was .48, and of the thirty-seven effect sizes calculated in the studies involving college-level students was .41. While there was large variation in effect-size magnitude among studies within each grouping ( $Hw = 631.77$ ,  $df = 75$ ,  $p < .001$ ), a test of the differences between groups was statistically significant ( $Hb = 127.73$ ;  $df = 2$ ,  $p < .001$ ).

*Duration:* Studies were also grouped by the duration of the mastery program interventions. Six effect sizes were calculated from studies that implemented mastery learning for one week, twenty-nine effect sizes were calculated from studies that lasted 2 to 12 weeks, and forty-three effect sizes were calculated from studies that lasted longer than 12 weeks (usually a semester in length, none of which were longitudinal studies). The weighted average effect sizes for each of the groups were .78, .73, and .50, respectively; the differences between the groups were not statistically significant ( $Hb = 2.09$ ,  $df = 2$ ,  $p = .35$ ) and within each grouping the effect sizes were heterogeneous, concluding that duration of a study did not influence the size of the effect (a finding that ran

counter to the predictions of mastery theorists, who expected longer exposure to mastery practices to have greater effects).

**Affective variables.** *Student affect:* Thirteen studies measured student affect using a range of indexes to assess student feelings about the subject, including their attitudes about the importance of the subject, school in general, academic self-concept, and grade expectations. The varying affective indexes made calculating average effect sizes inappropriate – some measured student’s affect toward school, while others measured student’s grade expectations or feelings about the importance of the subject matter – however the results from each of the studies indicate an overall positive effect on affective outcomes for students in the mastery learning conditions, with weighted effect sizes ranging from .10 to 1.33. Students in the mastery groups were more confident in their abilities, enjoyed the subject area and felt it was more important, and accepted greater personal responsibility for their learning compared to those in the non-mastery groups.

*Engagement:* Eight studies used classroom observations to gather data on time spent engaged academically and time spent on-task. The weighted average effect size was .76 in favor of the mastery classes. There were also two studies that used student attendance and course attrition rates of college students to measure the time spent engaged academically. Classes taught using mastery had higher levels of attendance (a statistically significant difference when compared to the non-mastery courses), with an effect size of .38, and in seven of the eight



classes taught using mastery learning strategies, attrition rates were lower (average effect size of .85).

**Instructor variables.** *Teacher affect toward mastery:* Studies measuring teacher reactions found that teachers felt more positive about the philosophy and practices of mastery learning; the effect size for this attitude change of 1.67 (Okey, 1997). Teachers who implemented mastery learning programs were no longer able to predict which students would do well and which would struggle with material (Guskey 1982) – they alter explanations for what makes them effective teachers, devaluing personality factors ( $E = -.38$ ) and increasingly valuing teaching practices ( $E = 1.13$ ). These teachers feel better about their roles as teachers ( $E = .61$ ), accept greater responsibility for their students' learning ( $E = 1.25$ ), but also express less confidence in their teaching abilities ( $E = -.59$ ).

*Exam type:* Most studies analyzed in this meta-analysis used teacher-prepared criterion-referenced unit examinations rather than standardized norm-referenced examinations to test objectives. Part of the mastery protocol is to make the learning objectives clear, so it is expected that mastery students would excel in criterion-referenced exams that assess the specific objectives taught; but because control classes are not narrowly focused on the defined objectives of instruction, and standardized norm-referenced tests have more general, broad-ranging objectives, control classes are expected to have learned a broader range of learning goals and therefore excel on those. Three studies reported results on both criterion-referenced tests as well as norm-referenced tests; two of them found no

statistically significant differences in the magnitude of the effects of mastery programs, whereas one study (Anderson, Scott, & Hutlock, 1976) found that mastery classes performed better on criterion-referenced measures than control classes ( $E = .58$ ) and the same as control classes on norm-referenced measures ( $E = .04$ ).

Overall, the conclusion that Guskey and Pigott reached was that effect sizes varied considerably across studies. Studies on achievement found pooled effect sizes ranging from 0.02 to 2.96, with most frequently reported effect size ranging 0.4 to 0.5 (found in 10 studies) followed by 1.0 to 1.1 (found in 6 studies). The analysis found that although variation in the effect sizes across studies were large and did not share a common underlying effect size, results of Bloom's group-based mastery learning model consistently yielded positive effects on cognitive and affective outcomes for students as well as on many teacher variables.

### **Shift in Priorities**

Given the intense focus on Bloom's model in the initial years after it was introduced, and its generally positive results on both teacher attitudes and student affect and achievement, it seemed that Learning for Mastery would permeate classrooms across the nation, but surprisingly enough, interest in the model began to decline as the standards movement gained traction. The *Nation at Risk* publication in 1983 asserted the failures of the country's educational system and raised urgent concerns about our rank against other industrialized nations. This incited a wave of efforts, which called for the implementation of academic,

content, and performance standards, from the 1989 National Education Summit that called for development of standards for student performance, the 1994 Educate America Act to create content and performance standards and state assessments, and culminating in the No Child Left Behind Act of 2000 tying performance to federal funding (Overview of the Standards Movement, 2015; Nation at Risk, 1983). The waning interest in Bloom's model coincided with the shift in school priorities, narrowing classroom attention toward efficiently preparing students to pass state proficiency exams and standardized tests and away from classroom activities geared at exploration, problem solving, and exposure to a wider curriculum (Stecher, 2010). With criticisms of standards-centric approaches to education accumulating (Relic, 2000), it is natural to want to look for a new approach to instruction; however, in the search for effective practices for teaching and learning, it is also worthwhile to reconsider existing approaches that showed promise. After all, abandoning a model of instruction because of a shift in funding priorities is not the same as abandoning it because it has been proven ineffective.

The early literature around Bloom's model showed promise in improving learning outcomes, and although research examining Learning for Mastery has subsided since its last thorough meta-analysis in 1988, there is still a line of active research examining its potential. I will now provide an analysis of the more recent literature on Bloom's Learning for Mastery model that has been conducted.

### **Subsequent Research (post-1988)**

There was much research examining mastery learning when it was first introduced, but with the changing landscape of educational policy and the emergence of instructional technologies, interest in examining Bloom's theory has dwindled. In an effort to find more current studies that examine Bloom's mastery model, I searched ERIC, Web of Science, the databases and e-resources available through Harvard's LibX library, and Google Scholar for all articles related to mastery learning. Over 150 papers were identified; to find more recent literature that was not already discussed in the meta-analyses, I filtered the search to articles published after 1990; approximately 60 papers remained. I then scanned the abstract of each document and separated empirical research from expository papers, leaving approximately 34 studies. All papers were read to ascertain that Bloom's Learning for Mastery model was applied on group instruction; 15 studies met this criterion. One of the studies was of simulation-based mastery learning practices for medical training, and was excluded due to its lack of experimental design, leaving 14 studies in my analysis<sup>10</sup>.

All 14 studies will be reviewed, and the various ways in which Bloom's mastery model has been interpreted and implemented will be assessed to determine whether effects of the model vary based on study characteristics. I begin by providing descriptive characteristics of the included studies, followed by a description of the general outcomes of the studies. Finally, I examine studies by

---

<sup>10</sup> The original intent was to narrow the selection of studies to randomized studies conducted in classroom settings that tested Bloom's model, however, only six studies described random assignment, five of which were conducted in classroom settings, and only three of which described mastery criterion as designated by Bloom (to 80% or more), therefore instead of limiting this analysis to three studies, I broadened my criteria to include all studies that described using Bloom's LFM model.

four salient factors – mastery level, correctives<sup>11</sup>, content area, and duration of the study – to determine whether these are aspects of Bloom’s mastery learning model that hold particular importance.

**Descriptive characteristics.** The selection of fourteen studies on Bloom’s learning for mastery model included students ranging from 3<sup>rd</sup> graders in elementary school to college students in graduate level courses. Five studies were conducted with students in university, eight with students in secondary school (defined as grades 7 to 12), and one study was done in primary school with students from grades 3 to 6. The topics of instruction spanned across social studies, English, math (including developmental math, matrices, intermediate algebra, and general mathematics), biology, chemistry, physics, economics, computer technologies, and psychology. The duration of the studies ranged from a two-day period in which the students completed two units of instruction, to a four-semester period in which previously gathered student data was analyzed post hoc<sup>12</sup>. Three of the studies reported the implementation duration was over a semester or school term; there was ambiguity in four of the papers as to the duration of the studies, but the description of the process insinuated a semester-long duration. Ten of the studies were conducted internationally in Turkey,

---

<sup>11</sup> The meta analysis Guskey and Piggott did not examine the ways in which correctives were implemented because the studies they examined often failed to provide sufficient detail about the characteristics of feedback and corrective activities offered, however, they emphasized its importance when evaluating implementation of Bloom’s model. I make a point to evaluate corrective procedures because the studies I analyze here are further removed from the initial attention that surrounded the model, making them potentially more susceptible to variations and interpretations that can develop over time, so it worthwhile to consider whether effectiveness of the model differs by corrective practices.

<sup>12</sup> The post-hoc studies analyzed data that was retained from classes that had previously implemented Bloom’s learning for mastery model; the course and data collection practices were not part of the study design.

Kenya, India, Iran, Malaysia, or Hong Kong, and the remaining ones were conducted in the United States. One study was performed in a lab that replicated a classroom environment whereas the rest of the studies stated or suggested the experiments were conducted in an existing classroom setting. A brief overview of each of the studies can be found in the appendix.

**General findings.** The studies included in this analysis contained findings on model effects in three areas: cognitive outcomes (defined by student achievement on exams), student affect (defined by tests measuring motivation or attitudes), and time to mastery (defined by how long it takes a student to reach the mastery level defined by the study). Affective measures were determined by the researcher-prepared examinations in the studies by Mehar & Rana (2012) and Sankhian & Gahlawat (2014); the Learning Process Questionnaire (LPQ) was used by Lai & Biggs (1994) and Sadeghi & Sadeghi (2012), the Student Motivation Questionnaire (SMQ) was used by Changeiywu, Wambugu, & Wachanga (2010), Keter, Barchock, & Ng'eno (2014), and the Mathematics Attitude Inventory was used by Guzver & Emin (2004). All studies that measured achievement used tests developed by the researchers or teachers involved in the study.

The general finding among the studies in this analysis was that the learning for mastery approach produces better cognitive and affective outcomes than conventional group instruction, but there was considerable variation in size of the effects. On cognitive measures, some studies reported effect size ranging from 0.56 to 1.12, while other studies found no significant effect. The studies that

examined attitudes and motivation supported the notion that Bloom's learning for mastery approach produces better affective outcomes than conventional instruction, but studies that reported effect sizes varied from .31 to .76. Similar to the findings from the meta-analysis by Guskey and Pigott, there did not appear to be a consistency of effect sizes among the studies. The studies that examined the decreasing variability hypothesis, which claims there will be a convergence in the amount of time that it takes for students to learn, found no evidence in support of Bloom's theory that the students who started off slower catch up in speed.

The considerable variation in findings prompted a closer investigation of what factors within Bloom's model account for these differences. To determine whether the effectiveness of the model hinges on particular features – such as level of mastery specified, whether its effectiveness is dictated by the subject matter, the types of correctives employed, or the duration of implementation – these four factors were more closely examined within the studies to ascertain their explanatory power.

**Mastery level.** According to Bloom's model, units of instruction are intended to be arranged hierarchically such that latter units build on earlier units in a chronological manner, so as to equip students with the prerequisite knowledge necessary to understand the units as they progress through difficulty. While Bloom acknowledges that the level of mastery should be set depending on the nature of the material of instruction, performance of 80% on exams is often used to constitute mastery. Given the progressive nature of many subjects, in which a

latter unit often builds on the understanding of an earlier one, it is worthwhile to examine whether the higher levels of mastery lead to greater effects. By setting a higher threshold for mastery, it can be assumed that students must reach a more comprehensive understanding of the material that serves as the prerequisite for later units, thus resulting in better subsequent learning. With this in mind, I examined the studies to find out whether studies with higher levels of mastery produce better achievement or affective outcomes. To do this, I will group studies by mastery level, describe the overall findings, and then look at the different groups by student cognitive and affective outcomes.

Studies varied in the level of criterion for mastery – ranging from 60% to 100% – but the outcomes of those studies did not clearly cluster by mastery levels. Studies that required higher levels of mastery more frequently reported positive effects of mastery learning, but positive effects were not exclusive to studies with high mastery criterion. Six studies specified mastery criterion of 80% or higher. Five of these studies (Damavandi & Kashani, 2012; Guzver & Emin, 2004; Keter, Barchock, & Ng'en, 2014; Anderson, Barrett, & Butson, 1992; Hoon, Chong, & Ngah, 2010) found positive effects of mastery learning on measures of achievement and affect that were examined, while one study reported positive effects for only some students – namely “surface learners” but not “deep learners” (Lai & Biggs, 1994). Five studies set mastery criterion at below 80%. Two of these studies reported positive outcomes for mastery students (Changeiywu, Wambugu, & Wachanga, 2010; Miranda, 2014), one reported that mastery



benefitted “surface” but not “deep” learners<sup>13</sup> (Sadeghi & Sadeghi, 2012), and two reported no effects of Bloom’s mastery learning model (Livingston & Gentile, 1996; Martinz & Martinez, 1999). Two studies failed to mention the specific level of mastery criterion was used for the investigation, merely specifying that the mastery classes were taught according to Bloom’s mastery model; these studies both found that the mastery learning groups performed better on academic and affective measures (Sankian & Gahlawat, 2014; Mehar & Rana, 2012).

Student achievement, defined by grades or exam scores, was examined in eleven studies. Six of the studies used mastery criterion of 80% or greater. When compared against conventional instruction, students in mastery classrooms made significant gains in achievement in five studies that used 80% or greater mastery criterion (Damavandi & Kashani, 2010; Guzver & Emin, 2004; Anderson, Bennet, & Hutson, 1992; Hoon, Chong, & Ngah, 2010; Kazu, Kazu & Odzdemir, 2005); the remaining study indicated a significant gain in mean scores of “surface” learners but non-significant gains for “deep” learners (Lai & Biggs, 1994). Three studies that reported on student achievement used mastery criterion of below 80%; two of them found no significant difference on final exam scores (Martinez & Martinez, 1999; Miranda, 2014) – however the study by Miranda found significantly higher passing grades in the course for mastery students – and one reported that the mastery approach had a significant main effect on summative test

---

<sup>13</sup> This study categorized learners as either “surface” or “deep” learners based on their performance on a Learning Process Questionnaire that was given to all students. Raw scores were coded into deciles scale scores to classify the students in one of the two categories.

scores but that students with different approaches to learning (“deep” vs. “surface” learning) had differing effects (Sadeghi & Sadeghi, 2012). The two studies that did not specify a level of mastery criterion found significant gains in academic performance (Mehtar & Rana, 2012; Sankhian & Gahlawat, 2014).

Student affect (measured by surveys of motivation or attitude towards learning) was examined by eight studies. Five of the studies used mastery criterion of 80% or greater, two of the studies did not specify mastery criterion, and one study used below 80% criterion for mastery. Three of the five studies using 80% or higher mastery criterion (Anderson, Barrett, and Hutson, 1992; Guzver & Emin, 2004; Damavandi, & Kashani, 2010), as well as the single study that set mastery criterion below 80% (Changeiywu, Wambugu, & Wachanga, 2010), reported significantly higher affective scores for students in mastery conditions. One study using 80% or higher mastery criterion reported higher motivation of students in the mastery conditions but did not indicate anywhere in the study whether the difference was significant<sup>14</sup> (Keter, Barchock, & Ng’eno, 2014), and the other study using 80% or higher mastery criterion found through qualitative interviews of student affect that “surface” learners liked the mastery approach whereas “deep” learners did not (Lai & Biggs, 1994). Of the two studies that did not specify level of mastery, one reported significant differences in

---

<sup>14</sup> The study by Keter, Barchock, & Ng’eno (2014) was investigating differences in student motivation by gender. They found increased student motivation for those who received the mastery condition, but because their focus was on gender differences, they did not indicate whether the increase in motivation for the mastery students was statistically significant. While student motivation went up for both boys and girls, they reported no significant difference between the genders in their motivation to learn chemistry.

student attitude (based on researcher-developed attitude measures) between the mastery and control groups in favor of mastery (Mehtar & Rana, 2012), and the other noted that differing attitude levels of students did not correspond to differences in achievement (Sankian & Gahlawat, 2014).

In short, when examined by mastery criterion, studies with mastery levels of 80% and above reported gains in achievement scores more often than studies with mastery levels set at below 80%. On affective measures, most studies reported positive effects of the mastery model without trends by mastery level.

**Subject of instruction.** In the studies selected, the subject of instruction included: five studies of mathematics, one of physics, two of chemistry, one of biology, one of basic information technologies, one of English, one on economics, and one on social studies. To make comparisons consistent with the earlier meta-analysis, groupings were assigned to resemble Guskey and Pigott's distinctions of psychology, mathematics, language arts, and social sciences as closely as possible. The grouping of mathematics includes developmental mathematics, algebra, matrices, and physics. Science includes biology and chemistry. Language arts includes one study of English, psychology uses one study on the psychology of learning and instruction, and social sciences includes economics, social studies, and basic computer technologies.

Mastery learning was reported to have a positive effect on social science (Kazu, Kazu & Ozdemir, 2005; Mehtar & Rana, 2012; Sankian & Gahlawat, 2014), chemistry courses of science (Keter, Barchock, & Ng'eno, 2014;

Damavandi & Kashani, 2010), and five of the six mathematics studies (Hoon, Chong, Ngah, 2010; Anderson, Barrett, & Hutson, 1992; Guzver & Emin, 2004; Changeiwu, Wambugu, & Wachanga, 2010; Miranda, 2014<sup>15</sup>). The study with biology reported diverging results for “surface” and “deep” learners such that mastery learning was beneficial for only a subset of students (Lai & Biggs, 1994), a result that was also found in the study of English (Sadeghi & Sadeghi, 2012). The study teaching psychology (Livingston & Gentile, 1996), and one study of mathematics (Martinez & Martinez, 1999) did not find support of Bloom’s theory.

Bloom’s theory is said to attain optimal results for subjects in which units can be hierarchically ordered to ensure preceding material is mastered before subsequent material is taught. Mathematics and science generally lend themselves to being sequentially organized, and are also the subject areas in which mastery has been reported to be valuable, however studies in the social sciences, which often contain topics that do not require strict ordering of units that build upon one another, reported positive effects of mastery as well.

As part of the implementation of his model, Bloom calls for the course content in mastery classes to be divided into smaller, digestible units that are sequentially ordered. The effects of mastery do not appear to be limited to subjects that inherently require sequential ordering of content in which latter units

---

<sup>15</sup> The study of math by Miranda (2014) found that the effects of mastery with computer-assisted instruction were not significantly better than computer-assisted instruction without mastery on final exam scores; however, the students in the mastery with computer-assisted instruction had significantly higher passing rates in the course than the computer-assisted condition without mastery. This finding suggests that the mastery model alone (without the use of the computer assisted instruction) did not have significant effects on the final exam but did have significant effects on academic performance in the class overall.

depend on the understanding of earlier ones; it can also be effective for courses that build-in a sequential arrangement of instructional units.

**Duration.** Bloom posits that student affect in educational contexts plays an important role in learning. He theorized that implementation of his Learning for Mastery model would lead to improvements in student affect and therefore enhance learning outcomes. According to Bloom, providing quality of instruction that has proper feedback and correctives allows the students to have experiences of successful learning, which in turn lead the student to be more interested in embarking on another experience of learning. These experiences of success serve to not only motivate students to progress through the units of content but also provide the cognitive foundation and encouragement necessary for the student's academic outcomes to increase, leading to a reduction in achievement differences over time. The amount of time required for these improvements to be seen was not specified, but it is worthwhile to examine whether results differ by study duration.

Similar to the categorizations used by Guskey and Piggott (1988), short-duration studies that lasted one week or less were grouped together, medium-duration studies lasting two to ten weeks were grouped together, and long-duration studies that examined one semester to two years were grouped together.

Two of the three short-duration studies found positive results on achievement; the two subjects were science (chemistry) and mathematics (Hoon, Chong, & Ngah, 2010; Damavandi & Kashani, 2010). Four of the five medium-

duration studies found positive results of Bloom's model, these studies were in the subjects of math (Changeiywo, Wambugu & Wachanga, 2010; Guzver & Emin, 2004; Anderson, Barrett & Hutson, 1992) and social studies (Sankhian & Gahlawat, 2014). In the short- and medium-duration studies, which were in English and science (biology) respectively, the remaining studies found that mastery was beneficial for only a subset of students (Sadeghi & Sadeghi, 2012; Lai & Biggs, 1994). Six studies comprised the long-duration category, four of which found positive effects of mastery; the subjects of these studies were math (Miranda, 2014), science (chemistry) (Keter, Barchock, & Ng'eno, 2014), and social science (Mehar & Rana, 2012; Kazu, Kazu & Ozdemir, 2005). Two long-duration studies – one in math, the other in psychology – did not find support for Bloom's model, however, it should be noted that the studies were examining Bloom's diminishing variability hypothesis<sup>16</sup>, not affective or achievement variables (Martinez & Martinez, 1999; Livingston & Gentile, 1996).

When evaluating by the duration of implementation, the studies used in this analysis indicated that the mastery model was effective most often in medium-duration studies, which range from 2 to 10 weeks; however, of studies that examined achievement or affective variables, the long-term studies were most effective as they all reported gains for mastery students.

---

<sup>16</sup> Bloom hypothesized that over time, exposure to mastery learning procedures would reduce variability in achievement and time such that all students are performing at high levels and that time needed for corrective activities is minimal.

**Correctives.** The feedback and correctives that students are afforded are an important part of the quality of instruction they receive. Bloom's model calls for students to receive feedback about their misconceptions and be given various forms of alternate explanations, activities, and examples that differ from, enhance, and clarify the original instruction to help students correct their mistakes. Very few studies, however, report on the types of feedback and correctives that students receive, and in some cases, do not clarify whether students repeat corrective measures until they master the content, or whether there are limited opportunities to reach mastery before moving on to subsequent material. The assorted interpretations what constitutes quality of instruction prevented us from making defined boundaries for grouping studies; features of the studies will instead be described, and corresponding results conveyed.

Twelve studies mentioned how feedback and correctives of Bloom's model were implemented, however not all of them specified how feedback and corrective were executed in the study. Seven studies specified some aspect of the feedback and corrective process implemented in the study (Livingston & Gentile, 1996; Martinez & Martinez, 1999; Changgeiywu, Wambugu, & Wachanga, 2010; Guzver & Emin, 2004; Lai & Biggs, 1994; Hoon, Chong, & Ngah, 2010; Sadeghi & Sadeghi, 2012), five studies described the feedback and corrective process in the description of Bloom's model and described the treatment as the implementation of his model (Kazu, Kazu & Ozdemir, 2005; Damavandi, & Kashani, 2010; Keter, Barchock, & Ng'eno, 2014; Anderson, Barrett, & Hutson,

1992; Miranda, 2014), and two studies did not mention of feedback and correctives in either the description of Bloom's model or the description of the study executed (Sankhian & Gahlawat, 2014; Mehar & Rana, 2012).

It was explicitly stated or suggested<sup>17</sup> in seven of the studies that feedback and correctives were continued until the students reached mastery, and in all of those studies positive results of the mastery learning model were reported (Kazu, Kazu & Ozdemir, 2005; Changgeiywu, Wambugu, & Wachanga, 2010; Damavandi, & Kashani, 2010; Keter, Barchock, & Ng'eno, 2014; Anderson, Barrett, & Hutson, 1992; Miranda, 2014; Hoon, Chang, & Ngah, 2010). Four studies examined affective characteristics. Two of those studies found that students in mastery classes had a greater motivation toward learning the subject matter (Changeiywu, Wambugu, & Wachanga, 2010; and Keter, Barchock, & Ng'endo, 2014); one study reported significant changes on some dimensions of affect (Damavandi & Kashani, 2010), and the other study did not specify whether the increase in affect of mastery students was significant, noting only that there was no significant difference between boys and girls in their attitude toward the course (Keter, Barchock, & Ng'eno, 2014). Four studies measured achievement via exam scores or passing grades (Kazu, Kazu & Ozdemir, 2005; Damavandi, & Kashani, 2010; Miranda, 2014; and Hoon, Chong, & Ngah, 2010). All four of the studies reported significantly better achievement results of the mastery classes,

---

<sup>17</sup> Two of the five studies explicitly stated that students repeated the corrective activities until they reached mastery; the other three studies did not explicitly state how many attempts students were able to engage in corrective activities, but it was assumed to be unlimited because in all three cases they explained that Bloom's model requires feedback and correctives be repeated until the student reaches mastery and described the study as following Bloom's model.



with the mean of mastery groups performing 0.56 to 0.87 standard deviations better than the mean of control groups (Hoon, Chong, & Ngah, 2010; Kazu, Kazu, & Ozdemir, 2005).

Studies that limited the number of times students would be allowed to engage in corrective activities before progressing to the next unit reported mixed results. Effect sizes of 0.96, and 1.35 were reported on achievement measures in studies that did not mention any aspect of the feedback and correctives implemented (Sankian, & Gahlawat, 2014; Mehar & Rana, 2012). In studies that specified students received two rounds of feedback and corrective activities before moving on, results varied: one study found no significant difference between mastery and control groups on achievement and did not examine affective measures (Martinez & Martinez 1999), and the other, which examined both cognitive and affective measures, reported 1 standard deviation increase on summative exam scores for the mastery students, and 0.76 standard deviations higher on attitudes measures of mastery students over the control classes taught by conventional instruction (Guzver & Emin, 2004).

Overall, studies that reported the largest improvements allowed students as many iterations of feedback and correctives as necessary to master material before advancing to the next unit. The type of feedback or variety of corrective activities were not specified, but the studies that required students to reach mastery before progressing further found positive results from the use of mastery learning procedures. The studies that did not allow sufficient opportunity for students to

master the unit content before moving on, giving only one or two rounds of feedback and correctives before advancing them to the next unit, occasionally found mixed results about the effectiveness of the mastery model.

**Summary of subsequent research.** In the selection of studies analyzed, Bloom's mastery model was most effective in conditions where it was implemented for longer than two weeks, where mastery criterion was set at 80% or higher, and in which correctives that were repeatedly offered until the student reached mastery. No consistent effect was observed across studies; however this could be a result of the varying degrees to which aspect of the model were implemented. A table of the studies, participants, and findings, is presented in Table 1 in the Appendix.

### **Criticisms of Bloom's model**

**Pace and rate of learning.** Critics of mastery learning primarily point to constraints on class time and teacher effort as barriers to implementation. Bloom's group-based LFM mastery model of instruction requires the teacher to set the pace of class instruction, preventing the class from advancing to the next unit until mastery of the previous unit of instruction has been achieved by nearly all students. The group-paced stipulation can result in unnecessarily stunting the progress of the faster learners in the class; this delay is claimed to be harmful for the learning of the faster students because it moves the class at a slower pace through less content and replaces advancement with redundant busy work (Slavin, 1987). A study done by Johnson and Henning in 1979 examining content covered

in elementary schools noted that fifth grade students in the group-based mastery program were exposed to less than half of the objectives assigned to the grade level (Arlin, 1984). Critics argue that stunting the progression of the entire class in order for a few struggling students to catch up is not an effective use of school time, and requiring teachers to work with these students after class hours is an unrealistic expectation of teachers.

Slavin (1987) points out that studies in which LFM approaches were effectively tested to produce the kinds of results Bloom predicted required a great deal of extra time by the instructor for individual tutoring. Students who did not attain 80% were tutored after class until they met the mastery criteria, giving them more instructional time than students who did not require tutoring. The Learning for Mastery model assumes certain protocols, for example that content be broken up into smaller units of instruction, which can result in longer the initial classroom instruction time. In studies by Arlin & Webster (1983), the mastery students received about twice as much instruction time as non-mastery students, and in other studies the initial instructional time was about 20% to 33% greater on average (or one class session per week) for mastery students than non-mastery groups (Anania, 1981). Requiring the teacher to alter instructional practices and provide additional individual tutoring outside of class de-incentivizes the adoption of the master model, particularly if this time burden is expected of them consistently throughout the semester.

Mastery theory posits that the rate of learning is an alterable characteristic that is a result of instructional conditions; the amount of instructional time for the students who do corrective activities is expected to diminish as they master the material, which serves as prerequisites for later content (Block 1972; Guskey & Gates, 1986). Bloom acknowledges that some students may require additional instructional attention early on to ensure that they are equipped with the cognitive entry characteristics to succeed, but explains that as they progress through the units, the amount of time on correctives will diminish. To test this hypothesis, Arlin (1984) conducted a study to examine the amount of remedial time needed for students in mastery classes and concluded that differences in time for each unit did not diminish, and that additional time needed for the slower students to attain mastery remained stable over the course of the 10 instructional units. Although this study was often referenced as evidence of the ineffectiveness of Bloom's model, an examination of Arlin's data and results suggest a different conclusion. The data provided in Arlin's study actually showed a statistically significant decrease in the amount of remedial time students spent on correctives over the ten instructional units, such that time needed in unit 9 was one-fourth of what was necessary for unit 2 (Guskey, & Pigott, 1988).

Bloom's hypothesis was supported by several studies (Block, 1970; Merrill, Barton, and Wood, 1970; Anderson, 1976) who all found diminishing time differences as students progressed through units of study using the mastery model. Aviles (2001) notes that there appears to be a link between the classroom structure

of mastery elements and the amount of instructor time spent. An effective course structure, in which the strengths of the learning environment can be leveraged to facilitate the mastery learning process, can reduce the amount of time the instructor needs to spend with students individually. The high amounts of tutoring time with individual students while others independently work on corrective or enrichment activities, could be reduced to create a more effective mastery learning environment that does not produce lengthy delays between units of instruction.

**Use of instructional time.** In addition to the potential need for greater instructional time, a challenge to implementing Bloom's mastery model is the time that instructors must invest to prepare learning units, correctives, and quizzes that establish the instructional environment for mastery learning. According to Bloom, mastery procedures performed in appropriate instructional conditions can raise the performance of 90% of the class by two standard deviations when taught in individual-student tutoring conditions, and one standard deviation when taught in mastery group-instruction. Many group-based mastery studies, however, have fallen short of these expectations, calling into question whether the student gains, particularly on norm-referenced tests ( $E = .04$ ), are worth the teacher effort (Guskey & Pigott, 1988). The same pool of studies, however, showed that on criterion-referenced measures the mastery groups performed much better than control classes ( $E = .58$ ).

The discrepancy between criterion-referenced measurements (which compare performance to a standard rather than the performance of other students)

and the standardized norm-referenced measurements (the normal curve) can be attributed to the difference in the objectives of the test – standardized tests being known to cover a broad range of knowledge whereas the criterion-referenced tests are designed to cover the specific knowledge covered in the units of class (Aviles, 2001). Criterion-referenced tests may produce performance distributions that deviate from the normal curve because mastery learning theory posits that all students are capable of high levels of achievement under appropriate learning conditions, which does not necessarily lend itself to a normal curve (Stufflebeam, Madaus, & Kellaghan, 2000).

If the act of specifying the learning objectives in mastery classes limited students learning, then standardized test scores, which evaluate broad ranging objectives, would favor control groups; however this is not the case. All studies that evaluated performance of students on both standardized test and locally-created tests found either no statistically significant difference in the magnitudes of the effects, or differences that favored the mastery classes (Guskey & Pigott, 1988; Kulik et al, 1990). The preparatory effort required of teachers can be seen as an investment that can pay off over the years, as the course is repeatedly taught. When materials are modularized into smaller units and correctives are created to re-explain the content in a variety of ways, they can be refined and reused in later iterations of the course without many additional time demands, while presumably still providing the gains to student learning year after year.

**Equity of educational opportunity.** A philosophical concern that often arises among critics of mastery learning is around educational equity of resources and traditional grading: A learning environment in which some students utilize more time and supplemental corrective instruction than others does not deliver equal resources to all students. Also, if students are provided instructional environments that allow everyone to reach levels of mastery, then traditional letter grades no longer carry the same meaning; alternate ways of differentiating students in their rank against their peers would need to be documented (Muller, 1975; Cox & Dunn, 1979; Arlin 1984). Bloom argues that rate of learning is an alterable characteristic, and access to a quality instructional environment that affords students the time and resources necessary to reach high levels of learning is the ultimate objective of education, and the means by which to provide equal educational opportunity (Brandt, 1976). His focus was not on how to rank and order students, but rather, on how to support all students as they work toward mastery.

### **Synthesis of Part I**

Bloom did not want to be too prescriptive in his model so that it can be applied to any course of instruction; however, the lack of specificity about certain facets allows for interpretations that lead to ineffective, or at minimum sub-optimal, implementation. It seems clear that implementing Bloom's model can lead to significant cognitive and affective improvements for students in group instruction, but it does not seem that attaining 2-sigma gains is doable with fidelity

to his model as it exists. Bloom posed the 2-sigma challenge as a call to educators and researchers to find approaches for to group instruction that personalize the learning experience in ways that they benefit all students in the classroom. His group-based mastery model has provided a promising starting point that gets us part of the way there, but it needs more clarity around key leverage points that can lead to more systematic implementation in order to yield more consistent results.

The developments of modern technologies and advances in research over the past two decades offer a wealth of possibilities that were not available at the time that Bloom developed his model, and with the shift of schools toward standardization, very little attention has been directed to improving or re-examining the model in light of these new tools and insights. To evaluate whether a model of group instruction can provide an individualized experience for all students within the structure of the existing school system, we must acknowledge the aspects of Bloom's model that are effective, and give proper consideration to the advancements in technology and research that can augment and amplify the model to get us there.

## **Part II: Looking Forward**

There is a depth of knowledge outside of Bloom's learning for mastery model that is vital to its effectiveness. In this section I identify four areas that have mature lines of research, are compatible with Bloom's model, and that deserve empirical attention if we intend to find methods of group instruction that are as effective as individual tutoring. I identify areas of ambiguity in Bloom's



model and describe how the insights from more recent literature and available learning technologies can enhance and amplify its effects. I discuss existing frameworks that can be incorporated into Bloom's model as it currently exists, to provide a more systematic approach to addressing four key leverage points – design, entry characteristics, feedback, and cooperative learning – that could potentially produce the 2-sigma gains within the context of group instruction.

### **Design of the learning environment**

Bloom developed his model with consideration to pacing differences, acknowledging that student speed does not reflect ability. He included redundancies within his framework to accommodate differences in learning speed – students who require more time with the material beyond initial group instruction are provided corrective instruction, which offers additional exposure to the material. He did not consider leveraging the design of instructional environments to broaden the reach of the initial group-based teacher-paced instruction. Any model that intends to produce an equitable learning experience must account for differences among individual learners. Teachers often bear the burden of individualizing instruction for students in an effort to personalize the learning experience, but this is an unrealistic expectation to place on teachers within the existing school structure.

A modern, well-regarded framework that has recently been adopted in classrooms as a way of providing group instruction that benefit all students is Universal Design for Learning (ULD). UDL is an instructional framework that

distinguishes three principles for curricular design: providing multiple means of representation for a student to become acquainted with the content, multiple means of engagement that provide the appropriate amount of challenge in tapping into the student's interests, and multiple means of expression for students to demonstrate what they have learned. Unlike the model of differentiated instruction, in which the teacher is responsible for adapting the instructional materials, goals, and activities for the specific needs of each student in the classroom, UDL places emphasis on the initial design of instructional materials to ensure that it accommodates as broad a range of students as possible. The aim of UDL is to ensure that the curriculum is designed to include multiple means of representation from the onset so that the teacher are not left shouldering the burden of adapting materials to match the needs of individual students.

According to Bloom's model, the teacher delivers initial group instruction, and offers feedback and correctives that provide alternate explanations in ways that differ from the original. His model allows for, but does not require, the initial instruction to incorporate multiple representations. Since students differ in the ways they approach new information – some find it easier to grasp observable phenomena, others find verbal explanations, or introspection, or some combination of presentations, easier to grasp when encountering new information – therefore limiting the presentation of initial content to one form of delivery preferences students who are inclined toward that approach. Incorporating a variety of representations in the initial instruction creates a more equitable learning

environment in which students are afforded greater access to the material regardless of the approach that is most favorable to their learning (Felder & Brent, 2005).

Initial instruction is not the only place in which various representations would be valuable. In Bloom's model, students who have not mastered the material provided in the initial instruction engage in corrective activities. The correctives are not merely a redirection to the explanation provided earlier, but rather, incorporate any number of activities that reestablish and clarify the initial instruction in various different ways. Multiple representations are beneficial for learning when they are used for distinct purposes: to provide information that is complimentary to the original instruction, when they constrain possible misinterpretations, and when they facilitate a deeper understanding of the material (Ainsworth, 1999). Delivering multiple forms of representation in the initial instruction and offering correctives in which the various representations provide targeted support to clarify the learner's misconceptions ensures the course content is delivered in an unbiased fashion for the students in the class, allowing appropriate fit between each student's most effective approach to acquiring new information and the classroom instruction and corrective activities.

Bloom claims that quality of instruction explains up to 25% of differences in achievement. Flexible design of the learning environment enhances the quality of instruction built into the curriculum. I hypothesize that the inclusion of flexible design in the learning environment, which provides multiple representations of the

course content both in the initial instruction and corrective activities, would increase the effect above and beyond that which is currently found in the mastery literature on group instruction.

### **Entry Characteristics**

The degree to which a student is sufficiently equipped with the background knowledge to embark on a new learning task plays an important part in whether the new material is learned. Bloom has shown that up to 50% of variation in school achievement can be accounted for by the prerequisite background knowledge pertaining to the learning task, and while he emphasizes the importance of addressing it, he does not provide a systematic approach to dealing with this critical influence on outcomes. Often times, particularly in high school and beyond, an entire course is declared as a prerequisite, leaving ambiguity in what facets of the course are actually critical to the upcoming learning task, and exposing students who may have a rudimentary understanding of the material, perhaps only sufficient to attain a barely passing grade in the prerequisite course, to learning tasks that build on concepts they have only somewhat acquired. Given the important role of cognitive entry characteristics in the mastery model, a systematic approach to ensuring students are sufficiently equipped with the prerequisite knowledge required to succeed in the learning tasks ahead is required.

Technological advancements in the last decade have enabled circumstances that Bloom could not have envisioned for his model. The adoption of flipped classrooms – a model of instruction that requires students watch pre-recorded

lectures at home and reserves school time for in-class exercises – has resulted in a number of online repositories of video lectures ranging in length and topic. The availability of video lectures allows instructors to identify and assign relevant lectures that contain expected background knowledge required of students. However, stand-alone lectures are not sufficient in themselves; students' understanding of the instruction must be assessed, and feedback and corrective activities made available, to ensure that the students have sufficiently acquired the cognitive characteristics they are expected to possess at the start the course. Rather than necessitating an entire course be a prerequisite, which often contains additional content that is irrelevant to the upcoming learning tasks, existing online resources that are modularized by concepts and short units of instruction, can be selectively assigned to ensure that students have command of the essential background knowledge that is truly required. The prevalence of online resources and the ease with which they can be created and widely accessed in the digital space, allows the instructor to compile a variety of explanations and activities that support the acquisition of cognitive entry characteristics. *Khan Academy*, a dashboard of instructional video repositories and practice exercises for students to progress through at their own pace, and *Assistments*, an online tutoring platform with built-in flexible assessments and that provides immediate feedback to students while also apprising the instructor of student progress, utilized together as the flipped classroom approach for prerequisites, allow instructors to systematically ensure that students enter the course with the knowledge base

required to succeed. Although younger students may not be developmentally prepared to utilize the technology in this manner, older students, generally beyond elementary school age, are able to take advantage of this approach.

While Bloom emphasizes the importance of ensuring students possess the background knowledge required for the learning at hand, his model does not provide guidelines for how these cognitive characteristics are to be acquired, leaving open to interpretation the ways in which such an important facet of the model should be accounted for. I hypothesize that, particularly for older students, when specific background knowledge is expected, utilizing a flipped classroom approach, in which students are supplied with modularized units of prerequisite material prior to the start of the course, and utilizing flexible, instructor-guided, online assessments to ensure students are appropriately equipped with the knowledge to succeed in the learning tasks ahead, would increase the effect size above and beyond that which is currently found in the group instruction literature testing Bloom's model.

### **Actionable feedback**

Accurately assessing what a student knows is essential for recognizing whether mastery has been attained, and providing appropriate feedback in response to the assessment is critical to furthering a student's understanding. When left unspecified, feedback means different things to different people, requiring by definition only that some reaction to an activity be provided. Bloom describes the need for cues, feedback, and correctives to be incorporated into the

learning experience so that students can reach mastery, but while he does specify that they should be calibrated to the needs of the student, he does not describe specifics with regard to the qualities that the feedback should possess in order for it to be effective and powerful. Any model in which feedback is required should outline a systematic approach for providing it to ensure that it is effective.

In a review of over 500 meta-analyses that included 180,000 studies, John Hattie (1999) found that of the more than 100 factors that influenced achievement, feedback in the classroom (for example, instructor-provided cues or reinforcements) was among the top five influences, producing an average effect size of 0.79. The analysis revealed that the most effective forms of feedback delivered cues and reinforcements in the form of video-, audio-, or computer-assisted instruction that are directly related to the goal of the task. Effective feedback ensures that it is targeted, appropriate to the level of the student, and actionable. As such, the student must be able to ascertain the goal, the extent of progress made toward the goal, and what must be done in order to make better progress through the feedback provided (Hattie, & Timperley, 2007). Based on these insights, Hattie developed a framework for feedback to enhance learning that outlines the important characteristics for appropriate execution. His model asserts the levels of task performance, process of understanding, regulatory processes, as well as the individual, must be carefully considered because they matter for the kinds of affects that result from the feedback.

Bloom describes feedback as a component of quality of instruction, distinguishing it as an important piece that involves delivering appropriate cues, encouragement, and reinforcement specific to the needs of the learner. Although his model implies that the instructor is to calibrate the verbal and non-verbal cues to match the needs of each student in the class, he does not make explicit what features the feedback must possess to be meaningful to the student. The wrong kind of feedback, for example, emphasis on knowledge the student has not yet acquired, or careless presentation of the feedback, can lead to frustration and cause the student to give up rather than persist with the learning task (Howie, Sy, Ford, & Vaicente, 2000; Deci, Koestner, & Ryan, 1999). Adopting a framework for feedback, that lays out how to gauge the task complexity, use feedback to promote active information processing, and ensure it is not threatening to the student at the self level, is important to ensuring the feedback that is provided is effective in guiding the student toward learning. I hypothesize that by systematically employing Hattie's feedback framework within the existing context of Bloom's LMF model will give effect sizes above and beyond mastery alone.

### **Cooperative Learning**

So far I have discussed how to address facets of Bloom's model that often pose the challenges for group instruction. However, there are attributes of group environments that lend themselves to enriched learning experiences; this fourth piece leverages the group to facilitate peer social interactions around learning. Situations in which students work together in small groups are known to enhance



the learning experience and improve student confidence (Springer, Stanne, & Donovan, 1999). Collaborative learning develops interpersonal skills and encourages greater social and intellectual involvement of the students through the discussion process. Exposing students to different interpretations, explanations, or answers forces them to re-evaluate their own viewpoints and consider other strategies for learning (Dooly, 2008). Competition and cooperation both lend themselves to collaborative group environments, and though they have differing effects on student attitudes, the research shows that both cooperative and competitive conditions can produce better learning outcomes than isolated learning experiences without peer interaction (Guzver and Emin, 2005).

Rather than finding ways to offset group instruction, here I am considering a case in which the group environment may be the key to better results. Bloom designed his model around group instruction, allowing cooperative learning to be integrated into its existing structure; but it requires a framework for collaboration that enables its systematic implementation. One conceptual framework on which cooperative lessons can be built is Learning Together. Developed by Roger and David Johnson (1994), this framework identifies the conditions under which cooperative efforts are expected to produce better results than individual conditions: perceived positive interdependence, face-to-face interaction, individual accountability towards goals, use of interpersonal skills, evaluation of group-functioning in service of future group effectiveness. This general template of criteria can be integrated into the class structure to fit its specific circumstances,

making it an ideal fit candidate for a framework that can work within Bloom's existing model.

Group work, when structured correctly, can have both affective as well as cognitive benefits for students, and although Bloom acknowledges the importance of the student's affective experiences with learning, he only mentions the option of incorporating peer-tutoring in small group environments as part of repertoire of corrective activities instructors can choose to offer students. Bloom purports that affective characteristics explain 25% of the variance in achievement outcomes, and collectively with cognitive characteristics, up to 65%. Given the importance of student affect and how that influences later learning episodes, and in light of the advantages of cooperative learning conditions, incorporating an existing template for group work would be beneficial. I hypothesize that the adoption of a cooperative group work structure, such as the Learning Together framework (Johnson & Johnson, 1994), into Bloom's mastery model will produce greater effect sizes than have been found without group work.

## **Synthesis of Part II**

Bloom's 2-sigma challenge sought ways to replicate the effectiveness of individual tutoring under conditions of group instruction. His Learning for Mastery model was one attempt at doing so, and having shown to make important gains in achievement over conventional instruction practices, it has proven to be a promising starting point, but the many interpretations of its implementation have produced varying effect sizes, emphasizing the need for more structure in key

points of leverage. I highlighted four areas within Bloom's model that, if refined to incorporate the advancements in research and technologies of today, could create the conditions of group instruction that personalize the learning experience in a way that benefits all of the individuals in the class. I hypothesize that by incorporating facets of existing frameworks to address issues of design, entry characteristics, feedback, and cooperative learning, to create more systematic, flexible, and effective learning environments, Bloom's model could produce the 2-sigma gains predicted.

### **Conclusion**

The architecture of the current public educational system was based on a group instruction approach, and achievement differences have been a consistent problem since its early inception. While there have been disagreements as to whether the differences in performance are attributable to innate capacities, in which better-performing students were simply smarter, or to environmental factors, in which better-performing students were from more affluent living circumstances, it has been generally accepted that the schools are not able to influence the underlying cause. The education system was never intended to nurture the minds of the students and develop their potential (Peterson, 2010); rather, it was designed to sort students by their comparative performance and rank them on the degree to which they acquire information provided.

Benjamin Bloom, however, felt differently about the role and responsibility of schools in nurturing students. He believed that all students could reach high

levels of achievement, and that it was the schools' responsibility to create the learning conditions to facilitate this. After demonstrating that individual tutoring can produce two-sigma gains in achievement, which proved that top performance is possible for all students when given the right learning conditions, he developed a Learning for Mastery model that leverages effective aspects of individual tutoring and applies them to group instruction situations in an effort to replicate its two-sigma gains.

Various studies examined Bloom's model, particularly in the years following its inception, finding it very helpful in improving achievement scores and raising student affect, and yielding impressive effects on learning, but it fell short of the two-sigma gains that Bloom claimed possible. Rather than refining the model and continuing to improve on it over the years, however, interest in Bloom's mastery model dwindled as the focus of educators and policy makers shifted towards standards-based education. Various efforts have been made to raise student achievement, with educational leaders acknowledging the need for personalization of the educational experience, but the nebulous definitions of what that entails has led to alternative approaches that don't fit within the traditional public school structure that currently exists. Advocates of school reform argue the need for transformational change, claiming the current structure of schools is fundamentally flawed, and calling for the design of an entirely new educational system. Undergoing this kind of ambitious overhaul, however, takes time and exposes students to preventable risks and inadvertent consequences. Before

abandoning the current established structure, it is worthwhile, and in fact necessary, to carefully consider whether it is possible to provide a quality, equitable, individualized learning experience for all students within the constraints of the existing educational system while preserving its core structure of group instruction.

In this paper I have examined whether there are methods of group instruction that can be implemented within the constraints of the existing educational system to yield high levels of achievement for all students within a classroom. To this end, I analyzed Bloom's Learning for Mastery model as a starting point, identified points of ambiguity that could explain its shortcomings, and proposed four areas – design, entry characteristics, feedback, and cooperative learning – with frameworks that are supported in the literature and implementable by advancements in technologies. I hypothesized that incorporating the suggested frameworks within Bloom's Learning for Mastery model could produce the 2-sigma gains in student achievement that he envisioned.

There is no doubt that the centuries-old system of education currently in place brings with it major challenges, such as deeply rooted legacy policies and administrative obstacles that must also be addressed, but overcoming such political barriers does not necessitate a fundamentally different structure of education. At its core, the human and social elements of group-based teacher-paced instruction have a great deal of value that, if designed and executed well, can produce an equitable and quality learning environment. The challenge lies not

in the constraints that the structure of the current system imposes, but rather, the thinking of the individuals that operate within the system. Entrenched beliefs about student capacity provide convenient explanations for why achievement differences exist. A change of mindset alone is not sufficient for closing the achievement gap, a model of instruction that nurtures the development of all of the students is critical. While peripheral aspects of the educational system may require adjustment, its core structure can be preserved and leveraged to deliver a personalized educational experience that provides all students with learning opportunities that support their potential, nurture their growth, and develop their skills.

## References

- Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education, 33*, 131-152.
- Anania, J. (1981). "The Effects of Quality of Instruction on the Cognitive and Affective Learning of Students" Unpublished dissertation, University of Chicago.
- Anderson, L W. (1976). An empirical investigation of individual differences in time to learn. *Journal of Educational Psychology, 68*, 226-233.
- Anderson, S. A., Barrett, C., & Huston, M. (1992). A mastery learning experiment. (Tech. Rep.). Yale, MI: Yale Public Schools.
- Annenberg Institute for School Reform at Brown University (15 August 2015) retrieved from <http://annenberginstitute.org/>
- Anonymous. Discuss the rise and fall of the use of Keller Plan (PSI) instruction in sciences in the United States. Retrieved on 10 June 2015 from (<http://webs.wichita.edu/depttools/depttoolsmemberfiles/belder/Question%20%20-%20Keller%20Plan.htm>)
- Arlin, M. (1984). Time, equality, and mastery learning. *Review of Educational Research, 54(1)*, 65-86.
- Arlin, M. & Webster, J. (1983). Time cost of mastery. *Journal of Educational Psychology, 75(2)*, 187-195.
- Aviles, C. B. (2001). *Creatively adapting mastery learning and outcome-based education to the social work classroom*. Paper presented at the Joint Conference of the International Federation of Social Workers and International Association of Schools of Social Work, Montreal, Canada.
- Block, J. H. (1972). Student learning and the setting of mastery performance standards. *Educational Horizons, 50(4)*, 183-191.
- Block, J. H. (1970). The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished dissertation, University of Chicago. Retrieved from <http://search.proquest.com.ezpprod1.hul.harvard.edu/docview/302573418?accountid=11311>

- Bloom, B. (1968). Learning for Mastery. *Regional Education Laboratory for the Carolinas and Virginia; 1(2)*, 1-12.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Bloom, B. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership*, 4-17.
- Bonk, C. J., Lee, M. M., Reeves, T. C., Reynolds, T. H. [eds.] (2015). *MOOCs and Open Education Around the World*. London: Routledge.
- Brame, C., (2013). Flipping the classroom. Retrieved on 15 August 2015 from <http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/>.
- Brandt, R. (1976). On mastery learning: An interview with James H. Block. *Educational Leadership*, 584-589.
- Bravmann, S. (2004). Two, four, six, eight, let's all differentiate: Differential education: yesterday, today, and tomorrow. *Johns Hopkins School of Education*:  
<http://education.jhu.edu/PD/newhorizons/strategies/topics/Differentiated%20Instruction/differentiate/>
- Burke, J. A. (1983). "Students' Potential for Learning Contrasted Under Tutorial and Group Approaches to Instruction" Unpublished Dissertation, University of Chicago.
- Buskist, W., Cush, D., and DeGrandpre, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education*, 1, 215-234.
- <http://www.campbell.edu/content/662/overviewpaper.html>  
accessed on 01 August 2015.
- Carroll, J. B., & Spearritt, D. (1967). *A study of a "Model for School Learning."* Center for Research and Development on Educational Differences, Harvard University.
- Cavanagh, S. (2014). What is 'Personalized Learning'? Educators Seek Clarity. Education Week. Retrieved on 15 August 2015 from <http://www.edweek.org/ew/articles/2014/10/22/09pl-overview.h34.html>



- Changeiywo, J. M., Wambugu, P. W., & Wachanga, S. W. (2010). Investigations of students' motivation towards learning secondary school physics. *International Journal of Science and Mathematics Education, 9*, 1333-1350.
- Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the early childhood longitudinal study (ECLS) Kindergarten to first grade sample. *Journal of Educational Psychology, 98*(3), 489-507.
- Dooly, M. (2008). Constructing Knowledge Together. Extract from *Telecollaborative Language Learning. A guidebook to moderating intercultural collaboration online*. Bern: Peter Lang. (21-45)
- Cook, D. A. (1990). PSI: Two unorthodox studies. *The Behavior Analyst, 13*(1), 67-11.
- Coleman, M. R. (2001). Curriculum differentiation: Sophistication. *Gifted Child Today, 24*(2), 24-26.
- Coleman, M. R., Hughes, C. E. (2009). Meeting the needs of gifted students within an RtI framework. *Gifted Child Today, 32*(3), 14-17.
- Cox, W. F., & Dunn, T. G. (1979). Mastery learning: A psychological trap? *Educational Psychologist, (4)*1, 24-29.
- Damavandi, M. E., & Kashani, Z. S. (2010). Effect of mastery learning method on performance and attitude of the weak students in chemistry. *Procedia Social and Behavioral Sciences, 5*, 1574-1579.
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, (125)*6, 627-668.
- Dunne, J. D. (2002). Behavior analysis: No defense requires. *Electronic Journal for Inclusive Education, 1*(6), 1-13.
- Eccles, J. Wigfield, A., Harold, R. D., Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary schools. *Child Development, 64*(3), 830-847.

- Education News (2013). American Public Education: An Origin Story. Accessed on 25 June from <http://www.educationnews.org/education-policy-and-politics/american-public-education-an-origin-story/>
- <https://www2.ed.gov/pubs/NatAtRisk/risk.html> accessed on 01 August 2015
- Ehri, L. C., Rosenthal, J. (2010). Spelling of words: A neglected facilitator of vocabulary learning. *Literacy Development and Enhancement Across Orthographies and Cultures* (pp. 137-152). Springer US.
- Ehri, L., and Rosenthal, J. (2007). Viewpoints: Spellings of words: A neglected facilitator of vocabulary learning. *Journal of Literacy Research*, 39(4), 389–409.
- Ellis, E., Gable, R. A., Gregg, M., & Rock, M. L. (2008). REACH: A framework for differentiating classroom instruction. *Preventing School Failure*, 52(2), 31-47. [http://libres.uncg.edu/ir/uncg/f/M\\_Rock\\_REACH\\_2008.pdf](http://libres.uncg.edu/ir/uncg/f/M_Rock_REACH_2008.pdf)
- Eisner, E. W. (2000). Benjamin Bloom. *Prospects: The Quarterly Review of Comparative Education*, XXX(3), 1-7. Retrieved on 15 June 2015 from <http://www.ibe.unesco.org/publications/ThinkersPdf/bloome.pdf>.
- Federation for Community Schools. Community School Transformation: A guide for schools, districts, parents, and community members. Retrieved on 15 August 2015 from of [http://www.communityschools.org/assets/1/AssetManager/WkshpI\\_GregHall\\_GuideCS\\_Transformation.pdf](http://www.communityschools.org/assets/1/AssetManager/WkshpI_GregHall_GuideCS_Transformation.pdf)
- Felder, R. M., & Brent, R. (2005). Understanding student differences. *Journal of Engineering Education*, 94(1), 57-72.
- Friso-van den Bos, I., Kroesbergen, E. H., Van Luit, J. E. H., Xenidou-Dervou, I., Jonkman, L. M., Van der Schoot, M. & Van Lieshout, E. C. D. M. (2015). Longitudinal development of number line estimation in mathematics performance in primary school children. *Journal of Experimental Child Psychology*, 134, 12-29.
- Gardner, W. (15 September 2010). Unintended consequences of school choice. Retrieved on 22 January 2016 from [http://blogs.edweek.org/edweek/walt\\_gardners\\_reality\\_check/2010/09/unintended\\_consequences\\_of\\_school\\_choice.html](http://blogs.edweek.org/edweek/walt_gardners_reality_check/2010/09/unintended_consequences_of_school_choice.html)

- Gallup, H. F. & Allan, R. W. (2003). Concerns with some recent criticism of the Personalized System of Instruction (PSI).
- Guskey, T. R. (2001). *Benjamin S. Bloom's contributions to curriculum, instruction, and school learning*. Paper presented at the annual meeting of the American Educational Research Association Seattle, WA.
- Guskey, T. R. & Gates, S. L. (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. *Educational Leadership*, 75-80.
- Guskey, T. R. & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. *The Journal of Educational Research*, 81(4), 197-216.
- Goldberg, W. A., Prause, J., Lucas-Thompson, R., Himsel, A. (2008). Maternal employment and children's achievement in context: A meta-analysis of four decades of research. *Psychological Bulletin*, 134(1); 77-108.
- Grant, L., & Spencer, R. (2003). The Personalized System of Instruction: Review and Applications to Distance Education. *The International Review Of Research In Open And Distributed Learning*, 4(2). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/view/152/233>
- Hattie, J. (1999) *Influences on student learning*. Inagural Lecture: Professor of Education, University of Auckland. Retrieved on 01 August 2015 from <https://cdn.auckland.ac.nz/assets/education/hattie/docs/influences-on-student-learning.pdf>
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Heward, W. L., & Dunne, J. D. (1993). A teleconference with professor Fred S. Keller. *The Behavior Analyst*, 16, 341-345.
- Hoon, T. S., Chong, T. S., Ngah, N. A. B. (2010). Effect of an interactive courseware in the learning of matrices. *Educational Technology & Society*, 13(1), 121-132.
- Howie, E., Sy, S. Ford, L. & Vicente, K. J. (2000). Human-computer interface design can reduce misperceptions of feedback. *System Dynamics Review*, 16(3), 151-171.

- Hymel, G. M., & Dyck, W. E (1993). *The internationalization of Bloom's learning for mastery: A 25-year retrospective-prospective view*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.
- Johnson, R. T., & Johnson, D. W. (1994). Overview of cooperative learning. Retrieved on 01 August 2015 from [http://clearspecs.com/joomla15/downloads/ClearSpecs69V01\\_Overview%20of%20Cooperative%20Learning.pdf](http://clearspecs.com/joomla15/downloads/ClearSpecs69V01_Overview%20of%20Cooperative%20Learning.pdf)
- Johnson, K. R., & Ruskin, R. S. (1977). *Behavioral Instruction: An evaluative review*. Washington, DC.: American Psychological Association.
- Kazu, I. Y., Kazu, H., & Ozdemir, O. (2005). The effects of mastery learning model on the success of the students who attended “usage of basic information technologies” course. *Educational Technology & Society*, 8(4), 233-243.
- Keter, K. J., Barchock, H. K., & Ng'eno J. K. (2014). Effects of cooperative mastery learning approach on students' motivation to learn chemistry by gender. *Journal of Education and Practice*, 5(8), 91-97.
- Kulik, Kulik, & Bangert-Drowns (1990). Effectiveness of mastery learning programs. *Review of Educational Research*, 60(2), 265-299.
- Kulik, J. A., Kulik, C., & Carmichael, K. (1974). The Keller plan in science teaching. *Science, New Series*, 183(4123), 379-383.
- Kulik, J., Kulik, C. L. C., and Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's Personalized System of Instruction. *American Psychologist*, 34, 307-318.
- Kulik, J. A., & Kulik, C. L. C. (1989). Meta-analysis in education. *International Journal of Educational Research*, 13(3), 221-340.
- Lai, P., & Biggs, J. (1994). Who benefits from mastery learning? *Contemporary Educational Psychology*, 19, 13-23.
- Lawrence-Brown, D (2004). "Differentiated Instruction: Inclusive Strategies For Standards- Based Learning That Benefit The Whole Class". *American Secondary Education*, 32(3): 34-62.

- Levy, H. M. (2008). Meeting the needs of all students through differentiated instruction: Helping every child reach and exceed standards. *The Clearing House*, 81(4), 161-164
- Livingston, J. A., Gentile, J. R. (1996). Mastery learning and the decreasing variability hypothesis. *The Journal of Educational Research*, 90(2), 67-74.
- Martinez, J. G. R. & Martinez, N. C. (1999). Teacher effectiveness and learning for mastery. *The Journal of Education Research*, 92(5), 279-285.
- Mehar, R. & Rana, A. (2012). Effectiveness of Bloom's mastery learning model on achievement in economics with respect to attitude towards economics. *Journal of All India Association for Educational Research*, 24(1), 1-13.
- Merrill, M. D., Barton, K., & Wood, L. E. (1970). Specific review in learning a hierarchical imaginary science. *Journal of Educational Psychology*, 61(2), 102-109.
- Merritt, E. (1953). Adapting instruction to individuals. *Educational Leadership*, 11(3), 187-188. <http://www.ascd.org/publications/educational-leadership/dec53/vol11/num03/toc.aspx>
- Metropolitan Center for Urban Education, NYU Steinhardt School of Culture, Education, and Human Development. (2008). *Culturally Responsive Differentiated Instructional Strategies*. New York University, NY.
- Miranda, J. (2014). The efficacy of an interactive computer system for teaching developmental mathematics to college students. *FIU Electronic Thesis and Dissertations. Paper 1148*.
- Mission: School Transformation retrieved on 15 August 2015 from <http://www.transformtexas.org/>
- Mueller, D. J. (1975). Mastery learning: Partly boon, partly boondoggle. *Teacher Education*, 3(11), 13-15.
- National Alliance for Public Charter Schools. (2009). A growing movement: America's largest charter school communities. Retrieved on 15 August 2015 from [http://www.publiccharters.org/wp-content/uploads/2014/12/2014\\_Enrollment\\_Share\\_FINAL.pdf](http://www.publiccharters.org/wp-content/uploads/2014/12/2014_Enrollment_Share_FINAL.pdf)
- NYU Steinhardt (2008). Culturally responsive differentiated instructional strategies. Retrieved on 01 August 2015 from

<http://steinhardt.nyu.edu/scmsAdmin/uploads/005/120/Culturally%20Responsive%20Differentiated%20Instruction.pdf>

[http://www.parentsassociation.com/education/standards\\_movement.html](http://www.parentsassociation.com/education/standards_movement.html) accessed on 01 August 2015.

Only a teacher: Schoolhouse pioneers. (n. d.). In *PBS*. Retrieved on 04 March 2016 from <http://www.pbs.org/onlyateacher/horace.html>

Peterson, P. E. (2010). *Saving Schools: From Horace Mann to Virtual Learning*. Cambridge: Belknap Press of Harvard University Press. pp 21-36.

Plomin, R., McClearn, G. E., Smith, D. L. , Vignetti, S., Chorney, M. J., Chorney, K., Venditti, C. P., Kasarda, S., Thompson, L. A., Detterman, D. K., Daniels, J., Owen, M., McGuffin, P. (1994). DNA markers associated with high versus low IQ – The IQ quantitative trait loci (QTL) project. *Behavior Genetics*, 24(2), 107-118. DOI: 10.1007/BF01067815

Relic, P. D. (2000). The trouble with the standards movement. *Parents Association*. Accessed on 01 August 2015 from [http://www.parentsassociation.com/education/standards\\_movement.html](http://www.parentsassociation.com/education/standards_movement.html)

Reville, P. (2015). The next steps in education reform. *The Boston Globe, Opinion*. Accessed on 20 August 2015 from <https://www.bostonglobe.com/opinion/2015/01/20/the-next-steps-education-reform/jYR7gFXaKtjKo6dU6gsiXL/story.html>

Rutledge, L. (2003). Differentiated instruction. Retrieved on 10 June 2015 from <https://wvde.state.wv.us/institutional/PD/Differentiated%20Instruction.pdf>

Sadeghi, A., & Sadeghi, A. (2012) Relevance of mastery learning in teaching of English. *Creative Education*, 3(1), 41-44.

Sankian, A., & Gahlawat, I. (2014). Effect of Bloom's mastery learning model and attitude towards social studies with respect to achievement. *Intellectual Quest*, 1, 52-65.

Sherman, J. G. (1977). Individualizing instruction is not enough. *Educational Technology*, 17, 56-60.

Silberman, R. (1978). The Keller Plan: A personal view. *Journal of Chemical Education*, 55(2), 97-98.

- Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, 57(2), 175-213.
- <http://www.sonoma.edu/users/p/phelan/423/standards.html>, accessed on 01 August 2015.
- Springer, L. Stanne, M. E., and Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21-51.
- Stecher, B. (2010). Performance assessment in an era of standards-based educational accountability. Stanford Center for Opportunity Policy in Education.
- Strong, R. W., Silver, H. F., and Perini, M. J. (2001). Making students as important as standards. *Educational Leadership*, 59(3), 56-61.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). Evaluation models: Viewpoints on Educational and Human Services Evaluation (2<sup>nd</sup> ed.). MA: Kluwer Academic Publishers. (Page 121)
- Subban, P. (2006). Differentiated instruction: A research basis. *International Education Journal*, 7(7), 935-947.
- The School Reform Initiative (accessed 15 August 2015 retrieved from <http://www.schoolreforminitiative.org/>)
- Tomlinson, C. A. (2000). Differentiation of instruction in the elementary grades. *ERIC Clearinghouse on Elementary and Early Childhood Education Champaign IL*. ED443572.
- Tomlinson, C. A. (2004). Sharing responsibility for differentiating instruction. *Roper Review*, 26(4), 188-200.
- Tomlinson, C. A. (2003). Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching. Alexandria, VA: Association for Supervision and Curriculum Development.
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, 79, 192-227.

- Trogdon, E. W. (1980). An exercise in mastery learning. *The Phi Delta Kappan*, 61(6), 389-391.
- Virgil S. Ward (1986). Theory in the practice of differential education for the gifted, *Roeper Review*, 8:4, 263-271, DOI: 10.1080/02783198609552986
- Wagoner, J. L.; Haarlow, W. N. (2002). "Common School Movement." *Encyclopedia of Education*. Retrieved 04 March 2016 from Encyclopedia.com:<http://www.encyclopedia.com/doc/1G2-3403200146.html>
- Willis, S., & Mann, L. (Winter, 2000). Differentiating instruction: Finding manageable needs to meet individual needs. *Curriculum Update: Association for the Supervision and Curriculum Development*. Retrieved on June 10, 2015 from <http://www.ascd.org/publications/curriculum-update/winter2000/Differentiating-Instruction.aspx>
- Wormeli, R. (2003). Differentiating Instruction: A Modified Concerto in Four Movements Retrieved June 10, 2015 from <http://www.ldonline.org/article/5679/>
- Yildiran, G. & Aydin, E. (2004). The effects of mastery learning and cooperative, competitive and individualistic learning environment organizations on achievement and attitudes in mathematics. *Journal of the Korea Society of Mathematical Education Series D: Research in Mathamatical Education*, 9(1), 55-72.
- Zucker, A. A. (2009). Transforming schools with technology. *National Association of Independent Schools: Independent Schools Magazine*. Retrieved on 15 August 2015 from <http://www.nais.org/Magazines-Newsletters/ISMagazine/Pages/Transforming-Schools-with-Technology.aspx>



## Appendix

Table 1. Overview of the studies included in the analysis.

Authors (year)	Subjects	Content area	Mastery level	Study duration	Results on achievement and affective measures	Correctives
Livingston, & Gentile (1996)	(Graduate school) 376 total for study 1. 85 from study 1 for study 2.	Psychology of learning and instruction	78%	Long: four semesters' worth of data analyzed	This study found no support for Bloom's decreasing variability hypothesis	Instructor provided corrective feedback twice: 1) to the group after exam, 2) Students who did not attain mastery attended remedial session with review of content and Q&A time. If not mastered after test2, made appointment for individualized assistance and task was changed (given take-home paper assignment)
Martinez & Martinez (1999)	(College) 80 total from four sections	Math (Algebra)	70%	Long: one semester	No significant difference on achievement measures. This study found no support for Bloom's decreasing variability hypothesis.	Treatment had 3 total chances to master test. Corrective feedback for tests B and C were given in teacher's office (no mention of corrective feedback after test version A)
Kazu, Kazu, & Ozdemir (2005)	(University) 217 total	Information technology	100%	Long: one semester	Achievement test results: Mean score of mastery group was 0.87 SD better than the average control	Study describes assessing can vs cant do computer tasks, so feedback/correctives would be provided until students reach 100%
Changeiywo, J. M., Wambugu, P.W., & Wachanga S.W (2010)	(Secondary school) 161 total from four schools	Math (Physics)	60%	Medium: three weeks	On affective measures: Mastery students had higher motivation towards learning $F(3, 156)=34.46, p<.05$	Specified that the cycle of correctives and feedback is continued until mastery is met
Mehar, R., and Rana, A. (2012)	(9th grade) 100 total from two schools	Economics	x	Long: one semester	On achievement measures: the average of the mastery group was 0.96 SD higher than the average of the conventional group. Difference in mean gain scores of mastery group and conventional group was 17.169 ( $p < 0.01$ ). On attitude measures, difference in means of the mastery and conventional groups was 4.485 ( $p < 0.05$ ).	Did not specify, and did not describe feedback and correctives when discussing Bloom's model.
Damavandi, M. E. & Kashani, Z. S. (2010)	(11th grade) 40 total	Science (Chemistry)	80%	Short: 1 unit on atom structure	On achievement measures: mastery learning is more effective on chemistry performance than traditional instruction ( $p < 0.01$ ). On affective measures: mastery learning model has effect on change of attitudes in some dimension ( $p < 0.05$ ). No sig relationship between chem performance of control and treatment groups with low levels of learning.	Described that students that don't reach specified level are taught again and receive exam until all students pass, and some of them reach mastery. (it is somewhat confusing if 'pass' means mastery or simply not failing, the sentence is not clear)

Guzver & Emin (2004)	(Junior high) 158 total in six classes	Math	85%	Medium: three weeks	On summative achievement scores: difference between mastery and control groups was 1.116 SD favoring mastery. The difference between achievement levels of all mastery classes and the conventional classes is significant at the .01 level ( $q=8.475$ ). On attitudes towards math: .7577 SD higher ML. Students in ML conditions have significantly more positive attitudes towards math than students in conventional instruction ( $p<.01$ ) ( $q=4.918$ ). ML account for 15.36% of variance on mathematics attitudes.	If mastery not reached after first test, a parallel form of the test is given after corrective procedures which included worksheets.
Keter, Barchock, & Ng'eno (2014)	(Secondary school) 205 total	Science (Chemistry)	80%	Long: one semester	On affective measures: increase in motivation was not significantly different between boys and girls.	Describes Bloom's model of correctives and feedback until mastery is met, so it is assumed that the treatment adheres to this.
Lai & Biggs (1994)	(9th grade) 223 total from three classes	Science (Biology)	80%	Medium: nine weeks	On achievement measures: Mastery learning raised the mean score of "surface" learners by 18 points (which was significant) but only had marginal and non-significant gains of 6 points or less on "deep" or "no bias" learners. On affective measures: The "surface" learners liked mastery approach, but "deep" learners didn't.	Corrective exercises designed to help students identify their errors. Enrichment activities included tutoring classmates who needed corrective instruction. After test B, everyone moved to the next unit.
Anderson, Barrett, and Hutson (1992)	(Elementary school) 94 total across six classrooms	Math	80%	Medium: 2 units in one semester	On achievement measures: The means of pre-test to post-test for Unit1 were 75% to 85% (effect size .66), and for Unit2 were 75% to 83% (effect size .56) ( $p<0.0005$ ). On affective measures: mean of self-concept of ability increased from 3.5 to 3.8 (effect size .31)( $p<.01$ ).	Suggested in the study that feedback and correctives were provided until mastery was reached, but that's not made explicit
Miranda, J. (2014)	(College) 210 total (7 classes of 30 students each)	Math	70%	Long: one semester	No significant difference on the adjusted means of final exam scores of students of ML+MML classrooms vs only MML classrooms, but there was a significant difference between the proportion of students in ML+MML who passed the class compared to MML	Instructor monitors student's MyMathLab use and assigned specific homework exercises for extra reinforcement until student masters that topic
Hoon, Chong, Ngah (2010)	(Secondary school) 262 total from four schools	Math (matrices)	80%	Short: two days	Effect size of CML towards CCL was .5604, indicating that an individual learner in CML had a .5604 SD increase.	Student who failed to meet the required performance level received supplementary instruction and corrective activities immediately after each question until the requirement was met
Sadeghi & Sadeghi (2012)	(college) 240 total	English	70%	short: five days	Only reported means, did not report SD. Results of summative tests: both approaches (surface and deep) and treatment had significant main effects on the scores ( $F=3.33$ , $p < 0.05$ , $F=5.06$ , $p < 0.05$ )	specified 1: The test was given approximately once per cycle for the purpose of feedback typically taking about 15 minutes to complete and was marked by the subject teacher concerned and returned to students in the next class session.

Sankhian & Gahlawat (2014)	(9th grade) 100 total from 2 schools	Social studies	x	medium: 5 lessons	On achievement test: the average performance of mastery group on the post-test was 1.35 SD higher than average performance of students in conventional instruction. On affective measures: Differing attitude levels toward social studies did not correspond with differences in achievement scores $F=2.086$ , $p > 0.05$	did not specify, and did not describe feedback and correctives when discussing Bloom's model.
----------------------------	--	-------------------	---	----------------------	---	---

### Meta-Analysis by Kulik, Kulik, & Banger-Drowns

Kulik, Kulik, and Banger-Drowns (1990) conducted a meta-analysis of the effectiveness of mastery learning programs in which they evaluated 108 studies that used either Bloom's learning for mastery program, or Keller's personalized system of instruction program. Criteria for inclusion required that studies be field evaluations comparing an experimental group using mastery procedures and a control group, that studies report quantitative results from which effect sizes could be calculated, and that studies be free from serious methodological flaws. Outcome variables measured were student achievement (indicated by end-of-instruction examinations), retention (indicated by a follow-up exam some time after the instruction is completed), student attitudes towards instruction and subject area, course completion (measured by attrition), and amount of time needed for instruction. Similar to Guskey & Piggott (1998), effect sizes were used to express outcomes on a common scale per formulas provided by Glass et al. (1981).

Kulik and colleagues reported on 108 mastery studies, which included 72 using Keller's PSI and 36 using Bloom's LFM approaches, affirming that the difference in effects of the two methods was non-significant,  $t(102) = 1.50$ ,  $p > .10$ . In an effort to focus on Bloom's model, data on the LFM results, we extracted data from the tables and

charts provided, and LFM-specific effect sizes were calculated and conveyed; however, it was not possible to calculate the statistical significance of the effects based on the data available.

*Student achievement:* Of the 103 studies that measured student examination scores, 67 studies reported statistically significant differences, all of which were in favor of the mastery groups. The average effect size of all 103 studies was .52, and the standard error of the mean was 0.033 ( $t(102) = 15.78, p < .001$ ), indicating that the average student in a control class performed at the 50<sup>th</sup> percentile whereas in a mastery program the average student performed at the 70<sup>th</sup> percentile. With specific attention to the 36 studies using Bloom's model, 34 produced results that favored the mastery treatment, 24 (or 71%) of which were statistically significant, and produced an average effect size of 0.59.

LFM students performed similar to conventionally instructed students on standardized tests, and better than conventionally taught students on locally developed tests.

*Affective Variables:* Student attitudes toward the subject were examined by fourteen studies across both mastery processes, all but two of which reported more positive attitudes in mastery classes, producing an average effect size of 0.40,  $t(13) = 3.08, p < .01$ . Effect sizes of the three LFM studies ranged from -.09 to .83 (the average effect size of the three LFM studies was 0.38)

*Teacher variables:* The review reported on 18 studies that examined student attitudes and perception towards the method of instruction and found an average effect size of .63, reflecting more positive attitudes among students in the mastery classes about

the instruction they were provided ( $t(17) = 4.50, p < .001$ ). Three of these studies used the LFM model, and produced effect sizes of .27 in two instances and .24 in the other.

Overall, Kulik and colleagues concluded that the LFM approach has consistently stronger educational effects than other programs (see Kulik & Kulik, 1989 for meta-analysis of 40 areas of educational research). However, after careful scrutiny of the data in our attempt to isolate effects for LFM programs, we found errors in calculations and several inconsistencies between the descriptions of the data and that which was provided in the tables, therefore our attention to this analysis has been limited to relevant findings that did not contain questionable calculations.

### **Differentiated Instruction**

In contrast to the teacher-oriented approach, differentiated instruction assumes a student-oriented approach to classroom instruction. Differentiated instruction is a philosophy of teaching that intends to cater to the diverse needs of students in a classroom.

The idea of differentiating teacher instruction in order to meet the educational needs of the individual students in the classroom has been around for decades (Merritt, 1953; Ward, 1986; Bravmann, 2004), but it is Carol Ann Tomilson who has become the leading figure in the area of differentiated instruction through her many works articulating the aims and approaches of a classroom that abides by its philosophy. According to Tomilson (2000), differentiated instruction is a process by which the teacher can ensure that the content a student learns, the way in which it is taught, and the

manner in which learning is demonstrated is appropriately aligned to the student's readiness, interest, and preferred mode of learning (Tomilson, 2004, p188).

Under the tenants of differentiated instruction, all students learn the same principle ideas, but the teacher adjusts process, pace, and approaches to instruction to match the needs of the diverse learners in the classroom (Tomilson, 2000; Tomilson, 2003 via Subban, 2006). Accomplishing this requires teachers to know the readiness, interests, and backgrounds of each of their students in order to engage them in effective ways that create links between the content and the student's lives in an engaging manner (Coleman, 2001; Strong et al., 2001; Suban, 2006). The teacher is expected to plan the content and process of each lesson to support the learning styles and individual needs of the students in the classroom (Lawrence-Brown, 2004; Tomilson, 2001) which in turn creates a learning experience that is appropriately calibrated level of content, speed of learning, and activity of interest for that student (Subban, 2006).

Based on cognitive psychology research of student achievement (McTighe & Brown, 2005; Ellis, Gable, Greg, & Rock, 2008.), the differentiated instruction framework identifies four elements that can be adjusted once the teacher appropriately identifies the readiness, interest, and learning profile of students. These four elements are: 1) content, which entails what the student will learn and how that content will be accessed, 2) process, which entails the activities that the student engages with in order to understand the content, 3) products, which refers to the culminating work the student will use to demonstrate understanding of the content learned, and 4) learning environment, defined as the way the classroom works and feels for the student (Tomlinson, 2000).

Differentiated Instruction it is intended to be a flexible, dynamic process that gets adjusted based on the particular needs of the students in the classroom (Steinhardt, 2008; Rutledge, 2003). While there is not a specific recipe or instructional strategy for effectively implementing differentiated instruction, broad suggestions for application of the ideas have been recommended, such as clustering students into zones according to their learning styles (Bafile, 2004), continually monitoring the progress of each student in order to adjust instructional strategies and materials as necessary (Coleman & Hughes, 2009 via Whipple, 2012), or consulting a blueprint of suggested activities through continual iterations in the classroom (Ellis, Gable, Greg, & Rock, 2008).

In the practice of differentiated instruction, teachers must discern the readiness of students in their classrooms, assign tasks that match each student's learning style, and ensure activities are calibrated to the level of challenge appropriate for maintaining interest and engagement (Tomlinson, 2003; Subban, 2006). It is the teacher's responsibility to meet students where they are in their education and move them as far as possible on their journey (Levy, 2008 via Whipple 2012), even if that sets differing task-completion expectations of students within the same classroom (Lawrence-Brown, 2004).

While differentiated instruction is a flexible framework that allows teachers to creatively adjust elements of the learning environment to meet to the needs of the students in their classrooms, the tremendous demand on the teacher is also its primary shortcoming. In order for differentiated instruction to be effective, the teacher needs to know each student intimately enough to place them in appropriate learning groups (Bafile, 2004), assign them engaging and relevant materials (Subban, 2006), and guide them from where they are in their understanding to as far as they can get on their

educational path (Levi, 2008; Whipple, 2012). Developing a relationship with the student that is sufficient for acquiring that kind of relevant student information is a time-consuming process; and by the time the teacher has learned enough about a student to be effective in applying the suggestions of this framework the semester is nearly over.

Differentiated instruction is described as an effortful process that is a career-long pursuit. Designing learning environments and materials that support differentiated instruction is an energy-consuming endeavor, and becoming savvy with it requires three or more years (Metropolitan Center for Urban Education, 2008).

Many teachers interested in meeting the needs of their students are overwhelmed by the workload involved or feel the challenge is insurmountable without the sustained staff development and peer support necessary for effective implementation (Ellis, Gable, Greg, & Rock, 2008; Willis & Mann, 2000). Others may believe they are implementing effective approaches but in fact their strategies are not producing outcomes (Subban, 2006).

Differentiated instruction not only demands a lot from the teacher by way of preparation, effective implementation requires the support of district and school leadership as well as sustained programs for staff development (Metropolitan Center for Urban Education, 2008). Many educators have accepted the philosophy of differentiated instruction, but are unable to adequately execute it because of such factors beyond their control.

### **Keller's Personalized System of Instruction (PSI)**



A method of instruction that is less demanding on teachers is Keller's personalized system of instruction. The Personalized System of Instruction (PSI) approach proposes a method of teaching that does not require the instructor to deliver lectures to the class on the content of the course (Keller, 1968; Grant & Spencer, 2003; Eyre, 2007). The method was established in 1962 by Fred S. Keller as an application of the theories of operant conditioning applied to classroom environment to improve student learning outcomes (Eyre, 2007). Keller studied alongside B.F. Skinner in the 1920s and after obtaining his psychology degree from Harvard, he applied the principles of reinforcement learning to Signal Corps personnel in a military training center during World War II (Heward, & Dunne, 1993; Dunne, 2002). That experience shaped his thinking about instruction, assessment, and division of labor in the classroom, and led him to design the Personalized System of Instruction, which is claimed to produce better learning outcomes for students, while also providing a more enjoyable learning experience than traditional instruction (Buskist, Cush, and DeGrandpre, 1991; Johnson & Ruskin, 1977; Kulik, Kulik & Cohen, 1979; Kulik, Kulik, & Carmichael, 1974).

The Keller Personalized System of Instruction involves five features that are deemed essential for it to be effective for all students in the classroom. These five features are: 1) emphasis on mastery of units of content, 2) self-paced progression, 3) written materials of instruction, 4) use of peers and proctors for tutoring, and 5) offering teacher lectures as reward (Eyre, 2007; Kulik, Kulik, & Carmichael, 1974). The focus on written materials is intended to facilitate student self-pacing through the course content. The content is divided into topics, or units, that correspond to chapters in the course text, and are provided to students alongside stated objectives to be used as study guides for the

unit exams. Students are expected to read through the unit material on their own, either independently in the classroom (which is often managed like a study hall with a proctor available for clarification of passages during this time) or outside of class. When the student feels that the content has been learned, the unit exam is administered. In Keller's own application of the model he sent students to a separate room that had proctors and cubicles for test taking (Keller, 1968). Here the proctors probe student answers on correct or incorrect questions, provide feedback regarding the student's test performance, and offer individualized tutoring in areas of weakness (Eyre, 2007). The role of the proctors includes administration of exams, clarification of confusing passages in the text, and also advising students on which material needs most attention based on exam performance. This kind of personalized attention from the proctor serves as positive reinforcement for the student to continue working on the material until the unit test is mastered.

A unique feature of the PSI model is that there is a division of labor between the proctors, assistants, and the instructor in order to ensure students can get assistance when they need it without delay. The proctors are often peers who have already mastered the material and serve as tutors, the assistant is the intermediary between the proctors and the instructor for matters of tutoring, course material, and record-keeping, with the teacher serving a somewhat managerial role over the learning of the student (Gallup & Allan, 2003 via website). The instructor's role is to select and organize the course content into units of study and explanation for the students, construct the examinations, conduct final evaluation of student progress, and provide occasional lectures to the class. Lectures are not used for presenting new material or clarifying misconceptions, but rather, they are

intended to serve as rewards for students who have already demonstrated mastery of that content (Keller, 1968) so as to demonstrate usefulness or applicability of the material that was learned.

Soon after its inception, variants of the PSI model were implemented in various classes with favorable results, earning Keller awards for his contributions to the fields of Education and Psychology in the 1970s (see meta-analysis by Keller, Keller, & Cohen, 1979 for details). One of the challenges of Keller's model, however, is its implementability in general classrooms, particularly outside of higher education, where teachers don't have complete autonomy over most aspects of their course. Instructors in K-12 institutions rarely have the resources to distribute their teaching responsibilities among proctors and assistants; without this division of labor, the instructor must individually tutor each student in the classroom that requires an explanation (Trogon, 1980), at least at the start of the semester until students that have mastered early units can serve as proctors and tutors for their slower-performing peers (Sherman, 1977).

Flexibility of student self-pacing is one of the essential components of Keller's PSI model. Keller described to the students enrolled in his course that, "You will not be held back by other students or forced to go ahead until you are ready" (Keller, 1968, p. 80) and that students may take unit exams as many times as necessary without penalty to their grade. There is no limit to the number of times a student can take any given unit exam, the only stipulation is that a student cannot move forward until the unit is mastered (indicated by a passing grade on that unit exam). Acknowledging the time constraints of a semester-long course, he goes on to explain that the final examination held at the end of the term will cover all of the content provided to the students. Even in Keller's own

class, which faithfully executes his model, self-pacing is constrained by the duration of the semester; students who do not independently progress at a sufficient pace are penalized for their delay in the end. Often times, however, students who do not progress at a reasonable pace due to procrastination or repeated testing are inclined to drop out of the course, leading to higher attrition rates among PSI courses (Silberman, 1978, Cook, 1990; website). In an effort to implement Keller's Personalized System of Instruction within the limitations of the standard classroom, many teachers had to alter aspects of the model to account for the practical constraints of their teaching situations, but Keller himself noted that if teachers are unwilling to try the whole system exactly as it's prescribed (with adequate support and resources), then it shouldn't be tried at all (Howard & Dunne, 1993).

### **Brief Overview of Studies**

Changeiywo, Wambugu, & Wachanga (2010) required the cycle of correctives and feedback to be repeated until mastery of physics in their examination how mastery influenced the motivations of students to learn physics. Even though the mastery was set at 60%, after the three-week period of the experiment, in which the students covered three units of physics, the students in the mastery learning classrooms reported higher motivation towards learning  $F(3, 156)=34.46, p<.05$ .

Hoon, Chong, & Ngah (2010) taught matrices in a two-day study using computer-assisted technologies to facilitate cooperative learning, mastery learning, and cooperative+mastery learning conditions. In this study, the mastery condition allows

students to move at their own speed whereas the cooperative+mastery condition involves group-pacing in which students had to wait until all group members achieved 80% mastery before moving on, and those who had already mastered material were able to help those who had not. The cooperative+mastery condition is a better reflection Bloom's model than the mastery condition, in which the student works entirely independently in this study. Prior to commencing the study, researchers made sure that all students in all groups had necessary prerequisites to understand the math instruction on matrices, and those who did not score 80% on the prerequisite exam were provided interactive courseware to bring them up to speed. During the study, students in both of the mastery conditions received immediate supplementary instruction and corrective activities repeatedly until 80% mastery was met.

Results of this study reflect significant differences on gain scores between students in the three different conditions  $F = 20.155, p < 0.025$ ; the effect size of the mastery condition towards cooperative instruction was 0.5603. Cooperative+mastery learning, which better reflected Bloom's model, showed stronger effects; the effect size of cooperative+mastery learning towards cooperative instruction was 0.8778, indicating a student in the cooperative+mastery condition had a .88 standard deviation increase. When separated by abilities (determined by the nationally normed Penilaian Menengah Rendah (PMR) standardized examination of Malaysia) students deemed "low ability" did not achieve significantly higher gain scores in the cooperative+mastery condition whereas they did in the mastery condition,  $F = 10.093, p < 0.025$ , and students deemed of "high ability" showed no significant differences in gain scores in either of the conditions,  $F = 2.221, p > 0.025$ .

Anderson, Barrett, and Hutson (1992) taught two units of mathematics to 3<sup>rd</sup>, 5<sup>th</sup>, and 6<sup>th</sup> graders across 6 classrooms using multiple group pre- and post-test design with the sample acting as its own controls, implementing Bloom's model with 80% criterion levels for mastery. Particular focus was directed to teacher training to ensure proper implementation of the model of instruction and student achievement and affect (using measures of self-concept of ability) were assessed. With regard to achievement scores, the pre-test to post-test means for unit 1 went from 75% to 85% (effect size of 0.66), and for unit 2 went from 75% to 83%, yielding an effect size of 0.56 ( $p < 0.0005$ ). The mean of self-concept of ability increased from 3.5 to 3.8, corresponding to an effect size of 0.31 ( $p < 0.01$ ).

Guzver and Emin (2004) taught three units of mathematics over the course of three weeks, examining the effects of mastery learning in cooperative, competitive, and individualistic learning environments on achievement and attitudes in mathematics. Students who did not reach 85% criterion after the first exam received corrective procedures via worksheets that they had one hour to practice before taking the second version of the exam. After the second version of the exam, all students progressed to the next unit whether 85% mastery was reached or not. There was a significant difference between achievement of the mastery learning classes and the conventional learning classes at a 0.01 alpha level. On summative exam scores, the effect size difference between the mastery groups and the control groups was 1.12 standard deviations in favor of mastery learning. Cooperative learning conditions performed 0.41 standard deviations

better than competitive, and 0.42 standard deviations better than individualistic. Mastery learning under cooperative conditions performed 1.56 standard deviations above conventional class with individualistic learning conditions.

On scores of student attitudes towards mathematics (measured by MATE), students in the mastery learning groups have significantly more positive attitudes towards math than those under conventional instruction,  $p < 0.01$ . An effect size difference of 0.76 standard deviations was found between mastery conditions and conventional classes, in favor of mastery learning. Mastery conditions account for 15.36% of variance on mathematics attitudes. The difference between cooperative learning conditions and individualistic environments was 0.58 standard deviations, and between cooperative and competitive environments was 0.45 standard deviations, both cases in favor of cooperative conditions. The largest effect size (1.41 standard deviations) was between mastery class under cooperative conditions and the control class under competitive conditions.

Kazu, Kazu, and Ozdemir (2005) taught the usage of basic information technologies to college students over one school term. Experimental and control groups were selected using cluster analysis to maintain objectivity. In the experimental group, units of instruction were taught using Bloom's mastery model; each unit ended with a follow-up test and correction training was provided for students who did not meet mastery criterion. Corrective education mentioned in this study was performed outside of class. This study found that there were significant differences on final test scores ( $t = -4.31, p < 0.05$ ) and achievement points ( $t = -2.82, p < 0.05$ ) in favor of the mastery group.

The mean score of achievement test results for the experimental group was by 0.87 SD better than the average control.

Keter, Barchock, & Ng'eno (2014) taught chemistry in 52 secondary schools to examine whether mastery learning impacted student motivation, and whether effects were differential by gender. Treatment groups received Cooperative Mastery Learning Approach treatment (CMLA): units with predetermined objectives, students work through each unit in an organized fashion, and pass unit quiz with score of 80% or higher. Students who scored below mastery level receive tutoring, peer mentoring, small group discussions, or supplementary assignments, and additional time until they reach mastery. There were no gender differences in motivation at the start of the study ( $t(102) = 0.602$ ,  $p > 0.05$ ) or after mastery treatment  $t(100) = 0.310$ ,  $p > 0.05$ ; the study found that the motivation of both male and female students increased as a result of the treatment but there was no mention of whether this increase was significant.

Lai & Biggs (1994) taught Biology in a 9-week study that covered four units, where mastery was set at 80% criterion and corrective exercises were designed to help students identify their own errors and enrichment activities included tutoring classmates. Students who did not reach mastery were given a second version of the unit test, after which everyone progressed to the next unit regardless of mastery attainment. The Learning Process Questionnaire was used to determine whether students used surface, deep, or no bias learning approaches. Approaches to learning were found to be significant  $F(2, 154) = 3.26$  at the 0.05 alpha level. The treatment main effect was also



found to be significant  $F(1, 154) = 5.06$ , as was the interaction of approach by treatment,  $F(2, 154) = 3.12$ . Mastery classes outperformed control classes in summative tests, but students with bias towards surface learning did comparatively better than others. The study found an interaction between students learning approach by unit test, ( $F = 7.17, df = 7, p < 0.01$ ) indicating performance of surface learners progressively improved over four units whereas performance of deep learners progressively deteriorated. Treatment main effect favors mastery condition, but there was a difference in which kinds of learners benefited from the mastery approach.

Follow up interviews with 16 of the students in the study (8 surface and 8 deep learners) about the affective experience of mastery learning found that none of the surface learners disliked the mastery approach, and most of the deep learners disliked the mastery approach. This study suggests that mastery promotes surface learning, has no benefit in terms of improving cognitive skills or analytic power, and is only beneficial for average to low ability student students.

Livingstone and Gentile (1996) analyzed four semesters worth of existing data from a graduate course on the psychology of learning and instruction to evaluate the validity of Bloom's decreasing variability hypothesis, which suggests that using the mastery model will cause the variability in student performance to reduce over time such that nearly all students will perform at high levels. For this analysis, students who did not reach the 78% mastery criterion by the first test were considered slow. After the first exam, students who did not attain mastery attended remedial session in which the instructor provided corrective feedback to the group before administering a second

version of the exam, after which the entire group progressed to the next unit. Meanwhile, students who did not reach mastery after the second exam also received individualized assistance on the un-mastered content and were given take-home assignment, which served as the alternate exam (this individualized assistance and take-home assignment process continued until the student met the mastery criterion. The study concluded that the decreasing variance hypothesis was not supported. A second study identified a subset of the students from the first sample that enrolled in a follow-up course, and analyzed additional units that were learned under the mastery process; this second study corroborated the first in that there was no evidence for smaller variances as a result of successive mastery experiences.

Martinez & Martinez (1999) taught intermediate algebra to college students using 70% mastery criterion. Corrective instruction was provided in the teacher's office for students that did not mastery the second or third version of the exam, but there was no description for how feedback and correctives were administered for students who did not reach mastery on the first exam. This study did not find significant differences in learning outcomes between the mastery and control conditions ( $F(1, 62) = 0.22, p > 0.05$ ), but noted that mastery groups required twice as much of the teacher's time. The authors inspected the final exam scores of the control classes were compared against conventional class outside of the experiment and found statistically significant differences supporting the contention that the control condition was not representative of traditional classes, and that the teacher – rather than procedure – affected student performance.

Miranda (2014) taught developmental mathematics, a college remedial class, comparing mastery based computer assisted instruction, traditional computer assisted, and traditional instruction conditions over one semester. Mastery criterion was set at 70%; students who did not reach mastery were given feedback, correctives, and targeted homework exercises using a computer program (MyMathLab), which allows the teacher to monitor their progress toward mastery. The study found a significant difference in posttest scores among the three groups,  $F(2, 169) = 3.463, p = 0.034$ . There were no significant difference on the adjusted means of final exam scores of students in the mastery based computer assisted classrooms compared to the computer-assisted classrooms, but the proportion of students who received a passing grade was significantly higher in the mastery based computer assisted classes. The mastery learning computer assisted group had significantly better passing status over the control condition, Pearson  $\chi(2, N = 120) = 12.724, p = 0.002$ , Cramer's  $V = 0.326$ .

Sadeghi & Sadeghi (2012) taught English using 70% mastery criterion in an experiment that lasted 5 days. One test was given per unit, feedback was marked on the exam, and was returned to the student in the next class period. Students who mastered the unit instructed their classmates who had not yet obtained mastery; after the second version of the exam is administered, all students proceed to the next unit of instruction. Students were categorized into approaches, as "deep" or "surface" learners based on the Learning Process Questionnaire. This study found main effects of treatment on summative test scores ( $F = 3.33, p < 0.05$ ). Significant main effects for student learning approach or for the test occasion were not found, but there was an approach by test

interaction ( $F = 7.17, p < 0.01$ ), indicating that mastery benefits surface learners but not deep learners.

Damavadi & Kashani (2010) conducted a study with 40 students of age 16 to examine the effect of the learning for mastery approach on performance and attitude of students in chemistry. For performance, mastery learning was found to be more effective on performance of weak students in high levels of learning ( $F = 42.2, p < 0.01$ ), and on some attitude measures students in mastery conditions experienced an increase in positive attitudes toward chemistry learning ( $p < 0.05$ ).

Sankian and Gahrlawat (2014) conducted a study to test the effectiveness of Blooms LFM model on attitudes toward social studies achievement. 100 students in 9<sup>th</sup> grade social studies across two schools participated in the study, one school was assigned control (Government High School - GHS) and the other assigned mastery treatment (Government Model Secondary School- GMSS). Five unit lessons were developed according to Bloom's mastery model. Pre-test and post-test scores were used to ascertain performance gain, and attitudes test was created and administered to all students prior to the pre-test. The study found that the students in the mastery learning classes performed better than students taught using conventional instruction, performance does not vary by attitude levels (high attitude vs low attitude), and there is no interaction between performance through different approaches to instruction (mastery vs conventional) and attitude levels. Mean gain scores for mastery group is 25.72, and for conventional instruction is 18.48.

Mehar and Rana (2013) tested 100 students in 9<sup>th</sup> class of economics from two schools in Chandigarh using 2x2 factorial design; examined attitude levels and achievement levels (both attitude and achievement exams were developed exams by the investigator of the study). They found that on achievement measures, the average of the mastery group was .96 standard deviations higher than the average of the conventional group. Difference in mean gain scores of mastery group and conventional group was 17.169 ( $p < 0.01$ ). On attitude measures, difference in means of the mastery and conventional groups was 4.485 ( $p < 0.05$ ).