

THE UNIVERSITY OF CHICAGO

ESSAYS ON LEARNING THROUGH PRACTICE

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE IRVING B. HARRIS
GRADUATE SCHOOL OF PUBLIC POLICY STUDIES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY

JUNCHEN FENG

CHICAGO, ILLINOIS

MARCH 2017

ProQuest Number: 10255230

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10255230

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Contents

List of Figures	vi
List of Tables	vii
Abstract	ix
Introduction	1
0.1 The Instructor’s Problem and the Dual Role of Practice	2
0.2 Mastery, Learning, and Practice Efficacy	3
0.3 The Bayesian Knowledge Tracing Model	5
0.4 Learner Engagement	7
0.5 Overview of the Thesis	9
1 Learning Through Practice	12
1.1 Practice Efficacy and Learning Process	14
1.1.1 The General Definition of Efficacy	15
1.1.2 The Working Definition of Efficacy	16
1.1.3 The Learning Process	18
1.1.4 The Learning Process with Learner Engagement	20
1.2 The Statistical Inference Model	20
1.2.1 An Overview of Hidden Markov Model	21
1.2.2 Hidden Markov Model As A Learning Model	22

1.2.3	The Bayesian Knowledge Tracing Model	23
1.2.4	Learner Heterogeneity	24
1.2.5	Learner Engagement	25
1.2.6	LTP as HMM	28
1.2.7	An Example of Input Data	28
1.3	Express Learning Theories with the LTP model	29
1.3.1	Zone of Proximal Development	30
1.3.2	Reinforcement Learning	31
1.4	Adaptive Practice Recommendation	31
2	Model Estimation	34
2.1	Prior	35
2.2	Full Conditional Likelihood of Augmented Data	36
2.3	State Augmentation	39
2.3.1	Latent Mastery Augmentation	39
2.3.2	Learner Type Augmentation	42
2.4	Parameter Update	42
2.4.1	Conjugate Posterior	43
2.4.2	Adaptive Rejection Sampling (ARS)	43
2.5	Label Switching	46
2.6	Posterior Inference	47
2.7	Simulation	48
2.7.1	Single Learner Type	48
2.7.2	Multiple Learner Types	50
3	Model Identification	53
3.1	The Literature Review	54
3.2	Sufficient Statistics of the LTP Model	56

3.3	Identification of the Learning Through Practice Model	57
3.3.1	The Practical Implication	58
3.4	Identification of the Bayesian Knowledge Tracing Model	60
3.4.1	Revisit Beck&Chang	63
4	Dynamic Selection Bias of Sample Attrition	67
4.1	Introduction	67
4.2	Sample Attrition and Dynamic Selection Bias	68
4.2.1	Dynamic Selection Bias	68
4.2.2	Dynamic Selection bias and Dependence Structure	70
4.2.3	Stop-by-Rule and Stop-by-Choice	73
4.2.4	Sign of Dynamic Selection Bias in Stop-by-Choice	74
4.3	Correct Dynamic Selection Bias	76
4.3.1	The LTP Model	76
4.3.2	Simulation	78
4.4	Case Study	81
4.4.1	The Learning Environment	81
4.4.2	Data Cleaning Procedure	83
4.4.3	Empirical Hazard Rate Curve	83
4.4.4	Estimated Pedagogical Efficacy	85
4.4.5	The Fallacy of “Learning Gain” from Learning Curve	87
5	Effort Choice and Efficacy Ranking	90
5.1	Introduction: Efficacy Vesus Effectiveness	90
5.2	Effort Choice in Randomized Control Trials	92
5.3	The LTP model	94
5.3.1	Key Assumptions	94
5.3.2	Consistent Estimation of Relative Efficacy	95

5.4	The Difference in Difference Estimator	96
5.4.1	Inferring Relative Efficacy from Relative Effectiveness	97
5.4.2	Efficacy Ranking by the DID Estimator	98
5.5	Simulation Study	102
5.6	Case Study	103
5.6.1	The Learning Environment	104
5.6.2	Raw data	105
5.6.3	The Design and the Implementation of The Experiment	106
5.6.4	Differential Effort Choice	107
5.6.5	The result	112
A	Notation	115
B	Simulation Parameters	117
B.1	MCMC Algorithm Convergence Demo: Single Type	117
B.2	MCMC Algorithm Convergence Demo: Multiple Types	118
B.3	The Dynamic Selection Bias Demo	119
B.4	The Effort-induced Measurement Error Demo	119
C	Answer Classification	120
	Bibliography	125

List of Figures

3.1	The Learning Curves (T=3)	63
3.2	The Sufficient Statistics (T=3)	64
4.1	Posterior Distribution of Pedagogical Efficacy:Simulation	80
4.2	The Interface of the Learning Game	81
4.3	Observed Learning Curves of Knowledge Points	82
4.4	Empirical Hazard Rates of Different Knowledge Points	84
4.5	Conditional Hazard Curve of the LTP model	85
4.6	Counterfactual Learning Curve:No Sample Attrition	87
4.7	Counterfactual Learning Curve: With Sample Attrition	88
5.1	Stages of a Level in the Learning Game	104
5.2	Descriptions of Experiment Problems	106
5.3	Boxplot of Time Spent by Answer Type	109
5.4	Distribution of Time Spent on Item without Effort by Answer Classification	109
5.5	Distribution of Time Spent on Item With Effort by Answer Classification . .	110

List of Tables

1.1	An Example of the LTP Input Data	29
2.1	Estimated Pedagogical Efficacy of the LTP Models	49
2.2	Estimated Parameter of the Mixture Model (Global)	50
2.3	Estimated Parameter of the Mixture Model (Higher Mode)	51
2.4	Estimated Parameter of the Mixture Model (Lower Mode)	52
3.1	BKT Parameters in Beck&Chang (2007)	63
3.2	Estimated Efficacy	65
3.3	Estimated Initial Mastery	65
3.4	Estimated Guess Rate	66
3.5	Estimated Slip Rate	66
4.1	Hazard Rates Conditional on Observed Response in Stop-by-Rule	78
4.2	Estimated Pedagogical Efficacy of Stop-by-Rule Sample Attrition	79
4.3	Conditional Hazard Rate of Stop-by-Choice Sample Attrition	79
4.4	Estimated Pedagogical Efficacy of Stop-by-Choice Sample Attrition	80
4.5	Estimated Pedagogical Efficacy of the Parametric Specification	86
4.6	Estimated Pedagogical Efficacy of the Non-parametric Specification	86
4.7	Percentage of Learning Gain Attributed to the Practice Efficacy	89
5.1	Simulation Result of the DID estimator	103

5.2	Simulation Result of the LTP estimator	103
5.3	Answer Composition - All Wrong	108
5.4	Effort Rate Choices - All	111
5.5	Effort Rate Choice - By Response in the Pre-test	112
5.6	Estimated Relative Effectiveness and Relative Efficacy in Aggregate Score . .	113
5.7	Estimated Relative Effectiveness and Relative Efficacy by Knowledge Components	114
C.1	Answer Composition (Percentage) - All Wrong	121
C.2	Answer Composition (Percentage) - Partial Correct	121
C.3	Answer Breakdown: Nonblank Answer	122
C.4	Answer Breakdown: Wrong Shape	122
C.5	Answer Breakdown: Wrong Circumference	123
C.6	Answer Breakdown: Wrong Area	123
C.7	Answer Breakdown: Slip	124

Abstract

The future of education is human expertise and artificial intelligence working in conjunction, a revolution that will change the education as we know it. The Intelligent Tutoring System is a key component of this future. A quantitative measurement of efficacies of practice to heterogeneous learners is the cornerstone of building an effective intelligent tutoring system that is able to generate practice recommendations adaptive to individual learner's progress. This thesis proposes a framework for defining and estimating the practice efficacy, which can be applied to a wide variety of learning processes.

When the mastery is assumed to be an ordinal variable, learning is defined as a probabilistic transition from lower level to a higher level. Practice items differ in the probabilities. The practice efficacy is the magnitude of such probability and thus a measure of the instructional value of the practice items. Had the mastery been directly observed, practice efficacy can be directly estimated from data. Unfortunately, the mastery is not observed and practice efficacy needs to be revealed by a statistical inference model based on observed data. The Bayesian Knowledge Tracing model (BKT), a special instance of the Hidden Markov model, formalizes the “practice makes perfect” learning process in a homogeneous learner population based only on the observed response. The Learning Through Practice (LTP) model is an extension of the BKT model by the introduction of learner heterogeneity and the inclusion of learner engagement. The LTP model can be used to describe a variety of learning processes, such as reinforcement learning and zone of proximal development.

Learner heterogeneity in practice efficacy is a major contribution to the literature of dynamic learning process and the foundation of an adaptive practice recommendation system. This thesis introduces two types of learner heterogeneity: the state heterogeneity and the type heterogeneity. The state heterogeneity means learners from different starting mastery level respond to a practice item differently. The type heterogeneity means learners from the same starting level can respond to a practice item differently.

Another important improvement over the classical BKT model is the inclusion of learner engagement. As digital education flourishes, learner behavior data are now widely available. The LTP model capitalizes on these new data to account for learner engagement in the learning process. The two features analyzed in this thesis are stopp decision and effort decision. The former refers to the learner’s decision to exit the practice sequence before it ends; the latter refers to the learner’s decision to choose a level of effort in solving the problem. The stop decision address the potential bias of differential sample attrition. The effort decision captures the key intuition that “no pain no gain”.

Built on the previous learning theory, this thesis develops outlines a Monte Carlo Markov Chain(MCMC) algorithm to estimate parameters of the LTP model. The MCMC algorithm first augments the observed data with latent mastery by Forward Recursion Backward Sampling algorithm given the parameters, then uses Gibbs sampler to update the parameters given the augmented data. In the second step, if the conditional likelihood does not have a conjugate prior distribution, Gibbs sampler draws new parameters with Adaptive Rejection Sampling algorithm. Because a point estimation instead of the posterior distribution of parameters is used in practice, which is only valid if the model is identified from the frequentist view, this thesis also provides the necessary identification condition of the LTP model and the sufficient condition of the BKT model. The latter corrects a long-standing mistake in the literature to base parameter inference on the observed learning curve.

This thesis then provides two applications of the LTP model featuring different aspects of

learner engagement. In the first application, the LTP model is applied to a quiz dataset of two-digit multiplication and long division to correct for the dynamic selection bias. The average learning gain from repeated practice decreases by at least 50% for the two-digit multiplication and at least 75% for the long division. In the second application, the LTP model is applied to account for effort choice in a randomized control trial to rank efficacies of a practice question with or without a video instruction set. While the two question forms cannot be ranked by the Difference in Difference regression, the practice question with the video instruction is estimated to have a stronger efficacy by the LTP model after controlling for the differential effort choices.

Introduction

The future of education is human expertise and artificial intelligence working in conjunction, a revolution that will change the education as we know it. Learning through practice, or learning by doing, is among many areas where such collaboration is possible.

Practice, as defined by the Merriam-Webster dictionary, means “to do something again and again in order to become better at it”. Although repetition is a necessary condition for improvement, it is not sufficient. One must understand how practice, specifically problem solving as practice, as a learning process that improves mastery, the ability to solve similar problems. Practice efficacy measures how likely the mastery is improved by the practice item. Therefore, practice efficacy is the key to further our understanding of what to practice and how to practice so that a learner can get better at what she is practicing.

One key insight of this thesis is that practice efficacy is heterogeneous among learners, otherwise, the optimal system design would be one-size-fits-all. Learners differ in their starting mastery level and in their speed of learning. The definition, and the estimation, of practice efficacy, need to reflect these two characteristics. The Other key insight of this thesis is that practice efficacy coupled with learner engagement produces improvement of mastery because “no pain no gain”. A theory of learning must take both practice efficacy and learner engagement into consideration.

The major challenge to this endeavor is that mastery is not directly observed. The practice efficacy cannot be directly estimated from the observed data but has to be revealed by a

statistical model of the dynamic latent process based on the observed data. This thesis proposes such a statistical model to define and estimate the practice efficacy along with the learner engagement, called the Learning Through Practice model. Such model can be applied to describe a wide variety of learning processes.

0.1 The Instructor's Problem and the Dual Role of Practice

The necessity of studying practice efficacy stems from the nature of tutoring. If an algorithm recommends a sequence of practice questions to a learner, it should mimic what a good human instructor does when he interacts with a learner. The instructor assesses what the learner knows and what she does not know, then devises and executes a teaching plan. In the next interaction, the instructor uses the assessment again to gather feedback on the previous teaching plan, so that he can improve the current teaching plan. The two steps constitute an instruction loop that is repeated until the learner masters the material. If the instruction is limited to problem-solving as practice, it should also perform the two tasks. As an assessment tool, the result of the practice anchors inference of the learner's mastery. As an instruction tool, the process of practice elevates the learner's mastery. Computerized Adaptive Testing(CAT) offers a well-developed theory for assessment. CAT aims to measure the testee's ability with a pre-specified precision by using as few questions as possible. Although the underlying mathematical mechanisms differ, CAT favors items that sharply differentiates whether the testee's ability is above or below a certain threshold. If the testee's ability is above the threshold, she will solve the problem; otherwise, she will fail. Although it is a sound principle for assessment, it is not a sound strategy for instruction. Good practice enables a learner to solve problems that were previously out of her reach with the help of instructions or hints. A good practice question can be a bad exam question and vice versa.

In short, CAT is not sufficient to build a good intelligent tutoring system because it is not the optimal solution for creating practice as instruction. An understanding of how good a practice item in elevating learner’s mastery, or practice efficacy, is fundamentally important to design the instruction step of the tutoring system.

It is important to notice that the tailored instruction based on assessment is only necessary when learners are heterogeneous. Otherwise, there exists a one-size-fits-all lesson plan that an instructor can follow almost blindly. If efficacies of a practice problem to different types of learners are known, as more responses are observed, the system can better infer the learner’s information and choose the next practice item most effective to her type and mastery level. Therefore, a quantitative measurement of the practice efficacies to heterogeneous learners is key.

0.2 Mastery, Learning, and Practice Efficacy

The previous section highlights the importance of developing a measure of “practice efficacy”. This section conceptually defines efficacy in relation to a learning process. The learning process is defined in relation to the concept of mastery.

Mastery is defined as the unobservable (latent) capability to solve a problem or perform a task in a particular domain. Although the definition does not exclude transferable skill or creativity, this thesis focuses on the muscle memory aspect of the mastery: the ability to solve similar problems. For example, the mastery of solving two-variable equations requires the learner to solve any set of linear equations with two unknowns. However, it does not require the learner to recognize that such technique can solve for the price-quantity equilibrium of a one-good linear supply and demand market system. It should be emphasized that mastery is a conceptual construction that is not directly observable. The observed response to a problem is only a noisy measure of the true mastery. A learner without mastery can solve the problem

out of sheer good luck, while a learner with mastery may fail it due to bad luck. Although the mastery is unobserved, it can be inferred from the observed solution to the problem.

Learning is defined as a process in which a learner becomes capable of solving a problem that she is unable to previously. There are many modes of learning other than problem-solving, e.g. self-reflection on previous mistakes. This thesis only focuses on learning in terms of problem-solving. Practice is defined as solving a sequence of problems. The notion of “sequence” emphasizes the repetitive nature of practice, which is echoed in the Merriam-Webster’s definition. The old idiom “practice makes perfect” still rings true. Repetitive practice is the cornerstone of expertise in many professions (Ericsson et al., 1993), be it in sports, arts or science. The bad reputation of “teaching to the test” is a result of a misguided construction of mastery. If the test authentically represents the problems a learner needs to solve in order to be successful, teaching to the test is a good instructional strategy. This thesis will not venture to discuss how mastery should be defined, but rather to study the impact of the practice on learning given a pre-defined mastery and a set of problems.

To characterize learning, it is necessary not only to characterize the current level of the mastery but also the change of mastery. The characterization would have been an easy task had the mastery been directly observable or the observed response been a perfect signal of mastery. Consider a simple case that a learner solves two problems, failing the first time then succeeding the second time. If the outcome accurately reflects the state of latent mastery, one can infer the learner has no mastery while attempting the first time but has since gained mastery. Unfortunately, the observed response is only a noisy measure. Given the same sequence of observations as above, one can argue the learner has no mastery and she only got lucky the second time, or that the learner has always had mastery and only got sloppy the first time, or that the learner gains mastery by learning from her first failure. The fog of inference requires the analysis to take a stand on whether the level of mastery can change during the practice sequence. It is a subtle but critical difference. The analysis can either

view the practice sequence as a static description of the current mastery or view it as a dynamic trajectory of the historical mastery levels. The static view reduces the complexity of inference at the cost of eliminating the possibility of learning through practice.

The prevailing psychometrics literature takes the static view. The family of models derived from Item Response Theory (IRT) (Rasch, 1960; Carlson, 2013) requires local independence as a fundamental identification assumption (Lord, 1980): Conditioned on the latent ability, the response to each item is independent of one another. Although the local independence assumption does not explicitly state that the analysis is conditioned on a constant latent ability, IRT models commonly assume that it is the case. Because the static view excludes the possibility of learning, the IRT models are unsuitable for studying a learning process.

When taking the dynamic view of mastery, the pedagogical efficacy of the practice item quantifies its capability to change the latent mastery. The operational definition of efficacy depends on the operational definition of mastery and change. For example, if mastery is assumed to be a binary variable and change is defined as the transition from the non-mastery (0) to mastery (1), the efficacy can be defined as the probability of such change. If mastery is assumed to be a continuous variable and change is defined as the increment of the mastery score, the efficacy can be defined as the magnitude of such change.

When the mastery is defined as a discrete ordinal variable and learning is defined as a transition between levels of mastery, Hidden Markov Model (HMM) is a classical framework to model such dynamic latent process. The next section describes the Bayesian Knowledge Tracing model family, an HMM that describes the simplest learning process.

0.3 The Bayesian Knowledge Tracing Model

In learning analytics literature, the Bayesian Knowledge Tracing (BKT) model is a classic representation of the dynamic learning process (Corbett and Anderson, 1994). The BKT

model formalizes the intuition that “practice makes perfect”: A learner achieves mastery by repeated exercise in a probabilistic learning process. The BKT model family is the main user modeling engine in many intelligent tutoring systems (ITS), most notably the Cognitive Tutor (Aleven and Koedinger, 2002; Koedinger and Corbett, 2006; Ritter et al., 2007) by Carnegie Learning LLC and the ASSISTment (Pardos and Heffernan, 2010b) by Worcester Polytechnic Institute.

The Bayesian Knowledge Tracing model assumes that the learner’s mastery level has only two states: non-mastery and mastery. A learner’s initial state is characterized by the probability of having mastery. If the learner has no mastery, at each practice opportunity, she has a probability to transit from non-mastery to mastery. The transition probability is called learning rate in the literature and pedagogical efficacy in this thesis. If the learner attains mastery, she never regresses upon further practice. Because the observed response is a noisy measure of the latent mastery, the BKT model uses a guess rate to measure the probability that the learner guesses correctly when she has no mastery and a slip rate to measure the probability that the learner makes an accidental error when she has mastery.

The Bayesian Knowledge Tracing model breeds a family of learning models that employ the same structural representation of the learning process. The main extension of the BKT model is to relax the strong assumption of learner homogeneity by allowing learners to be different in their initial mastery probability (Pardos and Heffernan, 2010b), the slip and guess rate (Baker et al., 2008) or the learning rate (Lee and Brunskill, 2012; Yudelson et al., 2013). However, none of the authors question the assumption of two-state latent mastery. Learner heterogeneity can be modeled much more parsimoniously when this assumption is relaxed.

The model identification is another major research interest in the BKT literature. Beck and Chang (2007) first argued that the BKT model is not uniquely identified. They claimed that different parameter sets with very different interpretations of the learning process result in the same learning curve, therefore the model is not identified. Their work motivated some later

development of the BKT model to use individualized parameters to address the identification issue (Baker et al., 2008; Rai et al., 2009; Pardos and Heffernan, 2010b). Unfortunately, Beck and Chang are wrong about the BKT model being unidentified, because different parameter sets produce the same learning curve but different likelihoods. In short, the identification of the BKT model is not properly analyzed in the literature.

This thesis intends to address both problems in the current research: the restrictive latent model structure and the lack of formal analysis on the model identification.

0.4 Learner Engagement

Instead of defining all these concepts and constructing an abstract model, an alternative simple approach is to measure the instructional efficacy as the difference in the probabilities of correctly answering a question before and after the pedagogical intervention. One such example is the “Learning Gain” metric used by the Khan Academy, a popular K-12 online learning platform, to evaluate the instructional quality of its videos (Faus, 2015). If a learner is engaged in learning for as long and as focused as the pedagogical method requires, such measure is qualitatively the same as the pedagogical efficacy defined by a structural model in a large sample. However, the learner engagement is not perfect in low-stakes learning environments, such as school lectures or self-learning with a computer tutor.

There is an emerging literature on the intensity and the duration of learner engagement. As for the intensity of learner engagement, Ryan Becker (2004b; 2004a) first introduced the concept of “gaming the system” to describe the learner’s off-task behavior when using an intelligent tutoring system. Later research shows that such off-task behavior is not only prevalent during the instruction (Baker et al., 2010) or using the hints (Koedinger and Aleven, 2007; Wixon et al., 2012) of the intelligent tutoring system but also in classroom learning (Pardos et al., 2013). As for the duration of learner engagement in the context of practice,

Murray et al (2013) show that the learner’s practice sequence length varies substantially in the Cognitive Tutor and such differential attrition biases the inference of learning progress if only the aggregate data are used.

Imperfect learner engagement implies that a practice problem may improve learner’s mastery under the ideal condition, but would fail to do so in reality. The distinction between efficacy, the impact of such method under ideal conditions, and effectiveness, the impact of such method under realistic conditions, is well understood by the medical literature (Flay, 1986; Glasgow et al., 2003; Flay et al., 2005). Although effectiveness is the ultimate goal, the research may better understand the nature of the ineffectiveness by separating inefficacious trials and efficacious but ineffective trials. The development of an intelligent practice system can also benefit from the same insight. If the practice has no efficacy, the material should be eliminated from the instructional content pool. If the practice is efficacious but the learner is not engaged, the instructor (or the learning product manager) can provide more incentives and improve the interface to better engage the learner.

However, in contrast to the well-controlled lab environment in medical research, learning data are usually collected in an imperfect learning environment where learners do not always exert their best effort. The challenge is to separate the efficacy from the effectiveness in such dataset. There have been attempts to account for engagement intensity (Feng et al., 2009) and duration (Pelánek et al., 2016) within the Bayesian Knowledge Tracing framework. However, these adjustments are ad hoc and do not have a good theoretical foundation. This thesis incorporates the learner’s engagement as a part of the Learning Through Practice model. By doing so, it provides an analysis of the bias as well as the method to correct it.

0.5 Overview of the Thesis

The first chapter lays out the learning through practices (LTP) model. This chapter first develops a theory of learning and a definition of practice efficacy that highlights learner heterogeneity. Because mastery is latent, it then proposes the Hidden Markov Model (HMM) to be the statistical inference framework that reveals the practice efficacy from the observed data. It proceeds to introduce the Bayesian Knowledge Tracing model (BKT) as the simplest case of HMM. Based on the BKT model, this chapter extends the inference to incorporate both learner heterogeneity and learner engagement. Learner engagement includes two aspects: the stop decision describes the engagement length and the effort decision describes the engagement intensity. It closes with a short discussion how LTP model can be used to express various learning theories and provide guidance to adaptive learning recommendation.

The second chapter describes the Monte Carlo Markov Chain (MCMC) algorithm that estimates the parameters of the LTP model. The MCMC algorithm first augments the latent states with Forward Recursion Backward Sampling algorithm given the parameters, then uses the Gibbs sampler to update the parameters given the augmented data. In the second step, if the conditional likelihood has a conjugate prior distribution, the Gibbs sampler draws from the conjugate posterior distribution. If the conditional likelihood does not have a conjugate prior distribution, the Gibbs sampler draws new parameters by the Adaptive Rejection Sampling algorithm. The MCMC also requires rank order conditions to prevent label switching. This chapter uses simulation data to demonstrate that the MCMC algorithm can estimate the parameters of a learning model that have three states of latent mastery, two learner types, the stop decision, and the engagement decision with the reasonable precision.

The third chapter discusses the identification assumptions of the LTP model. By reparametrizing the LTP model as a multinomial distribution, the sample frequencies of the joint distribution of item assignments, observed responses, stop decisions and effort decisions are sufficient statistics of the system and moment conditions for identification. It can be proved that a local

optimum parameter set exists only if the number of the parameters are smaller than or equal to the number of moment conditions, and that Jacobian matrix of the moment conditions at the optimum solution has full column rank. This necessary condition for identification puts an upper limit on the number of model parameters. In the special case of the Bayesian Knowledge Tracing model, this chapter provides the sufficient identification conditions by solving the moment conditions explicitly.

The fourth chapter applies the LTP model to analyze the dynamic selection bias of estimated efficacy due to selective sample attribution. Different lengths of practice sequence are a common feature in a learning dataset. This chapter shows that selective sample attrition is not a sufficient condition to bias the estimation of practice efficacy of the BKT model. If learners exit the practice sequence based on an exogenous rule (stop-by-rule), even though it may be a selective attrition, both the LTP model and the LTP model consistently estimate the pedagogical efficacy. If learners exit the practice sequence based on their own choice (stop by choice), only the LTP model consistently estimates the pedagogical efficacy. The chapter applies the LTP model to a quiz dataset on two digit multiplication and long division. In that dataset, the majority of the observed change of the success rate may be attributed to selective sample attrition rather than practice efficacy. The selective attrition accounts for at least 50% of the observed success rate increment in the two digit multiplication quiz and 75% of that in the long division quiz.

The fifth chapter applies the LTP model to evaluate efficacy in low stake learning environment where measure errors (i.e. frivolous wrong response) due to a lack of effort abound. Randomized Control Trial is an important method to evaluate relative the pedagogical efficacy of different practice materials. Usually, the data are analyzed with a Difference in Difference (DID) regression which measures the relative effectiveness of the practice materials. This chapter shows that the relative effectiveness may not have the same ranking order of relative efficacy when the effort-induced measurement error is present. It further argues that the LTP model

that accounts for mastery-dependent effort decision correctly recovers the relative efficacy of the experimental data. The chapter applies the LTP model to an experiment that compares the efficacies of practice questions with or without a video instruction set. The case study details the classification of effort-induced error and the evidence of differential effort rates between two groups. Whereas the DID estimator shows no significant difference between two questions, the LTP model strongly suggests that the question with video instruction has a superior efficacy when controlled for effort input, which is consistent with the pedagogical expert's prior expectation.

Chapter 1

Learning Through Practice

The ultimate goal of the Learning through Practice (LTP) model is the selection of a sequence of practice problems for the purpose of enhancing a learner's mastery. To achieve this goal, this chapter provides a learning model that defines practice efficacy and a statistical model to draw inference about it from observed data.

The focal parameter of the chapter is practice efficacy (ℓ). Practice efficacy is defined as the probability that a learner reaches a higher level of mastery after attempting a practice item, or a practice problem, once. ℓ ranges between 0 and 1. When $\ell = 1$, the practice efficacy is perfect, a learner achieves higher mastery for sure after one attempt. When $\ell = 0$, the practice efficacy is null, a learner never achieves higher mastery no matter how much attempts she makes. Practice items differ in their efficacy, which is the rationale for engineering item assignments (A) of a practice sequence.

Designed for this purpose, the Bayesian Knowledge Tracing model (BKT) model assumes the simplest structure for practice efficacy. It postulates that the mastery only has two states and learners are homogenous in their response to items. The BKT model masks key learner heterogeneities that are important to the understanding of the learning process and the engineering of an adaptive learning strategy. In contrast to BKT, this chapter assumes

that there are more than two levels of mastery and learning rates differ depending on the mastery level, which is called state heterogeneity. This chapter also assumes that learners differ in their learning rate conditional on the same mastery level, which is called type heterogeneity. Formally, the LTP model postulates a learner's mastery level is an ordinal variable X , learner's type is a nominal variable Z , occasion within the practice sequence is a positive integer t and the item id is a positive integer j . Practice efficacy expressed the probability of a learner of type z ascends to mastery level n from the base level m after exposing to item j as $\ell_j^{z;m,n} = P(X_t = n | X_{t-1} = m, Z = z, A_t = j)$.

The learning process can be defined in relation to the pedagogical efficacy and item assignments. In this thesis, it is characterized by the evolution of density of mastery levels over time, rather than the evolution of mastery levels per se. As long as practice items differ in their efficacies, a learning process cannot be defined without the knowledge of item assignment, because different orderings of items lead to different orderings of transition probabilities across values of X , and thus different densities, even given the same starting density. Formally, let the probability of a learner with type z has mastery level k at the first practice occasion be $\pi_1^{z;k} = P(X_1 = k | Z = z)$. This is the prior belief of the learner mastery profile. Given the item assignments ($\mathbf{A} = \{A_1, \dots, A_T\}$), $\pi_{t,\mathbf{A}}^{z;n}$ be the density of a type z learner with respect to the mastery level n at occasion t is:

$$\pi_{t,\mathbf{A}}^{z;n} = \sum_{m=0}^n \pi_{t,\mathbf{A}-1}^{z;m} \ell_{A_t=j}^{z;m,n} \quad \forall t > 1$$

Had the mastery (X) and the learner type (Z) been observed, both initial density ($\pi_1^{z;k}$) and the practice efficacy ($\ell_j^{z;m,n}$) could be estimated from data. Unfortunately, neither of them is observed. The A statistical model is therefore needed to reveal it. The Hidden Markov Model (HMM) is a classical framework to describe such a dynamic latent variable model. The HMM model has two components: the hidden layer and the observed layer. The hidden layer describes the evolution of latent states in the form of a Markov Chain. The observed

layer describes the generation of observed data based on latent states.

The hidden layer consists of two elements, the initial density of latent states and the state transition matrix. In the context of the learning process, the initial density of the hidden layer is the prior belief of the density of a learner’s initial mastery. The lower right diagonal of the state transition matrix is 0 because of the no-regression assumption. The upper right diagonal of the state transition matrix is filled with corresponding practice efficacy.

The observed layer only involves the observed response (Y) in the classical BKT model. This chapter expands the observed data to include learner engagement, which includes the stop decision (H) and the effort decision (E). The generation of the observed data ($O = \{Y, E, H\}$), which are discrete variables, is specified as a multinomial distribution whose probability mass function depends on the latent mastery level $P(O|X = k)$. Importantly, when effort choice is involved, this thesis makes a critical assumption that “no pain no gain” which argues that a learner’s mastery cannot be elevated unless she tries. Therefore, it also affects the evolution of the latent state. Let E_t denote a learner makes a valid effort at occasion t , given the item assignments (A), the learning process is hence

$$\pi_{t,\mathbf{A}}^{z;n} = \pi_{t-1,\mathbf{A}}^{z;n} + E_t \sum_{m=0}^{n-1} \ell_j^{z;m,k} \pi_{t-1,\mathbf{A}}^{z;m} \quad \forall t > 1$$

1.1 Practice Efficacy and Learning Process

The core of the Learning Through Practice model (LTP) is practice efficacy. It characterizes the how practice items boost a learner’s mastery. This section starts with a general definition built on an ordinal representation of mastery, then outlines the assumptions leading to a simplified working definition used in this thesis.

1.1.1 The General Definition of Efficacy

Let the mastery (X_t) at practice occasion t be represented as a unidimensional ordered discrete variable with M_x number of states. M_x is a positive integer. In this thesis, it can take value either 2 or 3. $t \in \{1, 2, \dots, T\}$, where T is the max length of the practice sequence. The learning is not timed by clock time, but by the number of practice problems done, or the practice occasion. In the following analysis, “time” and “practice occasion” are used interchangeably.

The unidimensionality specification avoids the complexity of representing the response as a function of multiple inputs. If the latent mastery is multi-dimensional, the likelihood function of the observed response depends on the question being a single solution with sequential reasoning, multiple solutions with single step reasoning, or multiple steps with sequential reasoning. In addition, each reasoning can house one or more components of the mastery. The unidimensionality assumption also avoids the explosion of pedagogical efficacy parameters. If the latent mastery is multi-dimensional, the state transition matrix of one dimension is unlikely to be independent of all other dimensions. The number of the parameters of the transition matrix explodes exponentially as the dimension of the latent mastery grows.

In an ordinal mastery representation, learning can be defined as ascending from a lower level to a higher level. Practice efficacy is thus defined as the probability of such ascension. In addition to the learner type and mastery levels, a general definition of efficacy requires encoding the learning context, which includes item assignments, responses, and practice occasion. Let j practice item id, $j \in \{1, 2, \dots, J\}$, where J is the total number of items in the question bank. Let Y_t denote the observed response at practice occasion t , which can take value 0 or 1. Let A_t denote the item assignment at practice occasion t . $A_t = j$ means that a learner encounters item j at practice occasion t . Formally, the general efficacy can be written as

$$\ell_{A_1, \dots, A_t; Y_1, \dots, Y_t; t}^{z; m, n} \equiv P(X_t = n | X_{t-1} = m; Z = z; A_1, \dots, A_t; Y_1, \dots, Y_t; t)$$

An important assumption of the learning process is that a learner never regresses on mastery, or she never forgets what she has learned.

Assumption 1:

$$\ell_{A_1, \dots, A_t; Y_1, \dots, Y_t; t}^{z; m, n} = 0 \quad \forall m > n$$

1.1.2 The Working Definition of Efficacy

The highly contextualized general definition of practice efficacy results in too many parameters to be estimated. For the sake of feasible inference, a working definition of practice efficacy needs to be decontextualised. In this thesis, the working definition is

$$\ell_j^{z; m, n} \equiv P(X_t = n | X_{t-1} = m; Z = z; A_t = j)$$

Compared to the general definition, the working definition of the practice efficacy removes the dependencies on observed response (Y_1, \dots, Y_t) , previous item assignments (A_1, \dots, A_{t-1}) , and practice occasion (t) . This subsection discusses the practical implications of these simplifications.

For a particular practice sequence with length T , denote the joint responses as $\mathbf{Y}_{1,T} = (Y_1, \dots, Y_T)$. Similarly, $\mathbf{X}_{1,T}$ is the joint latent masteries, and $\mathbf{A}_{1,T}$ the joint item compositions. Let $\mathbf{y}_{1,T} = (y_1, \dots, y_T)$ be the realized response sequence and P denotes the probability mass function where $P(\mathbf{Y}_{1,T} = \mathbf{y}_{1,T}) = P(Y_1 = y_1, \dots, Y_T = y_T)$. Similarly, $\mathbf{x}_{1,T}$ is the realized joint latent masteries, and $\mathbf{a}_{1,T}$ the realized joint item compositions. For simplicity, when referring to the whole practice sequence, the underscript $1, T$ is dropped throughout the thesis.

Assumption 2: Pedagogical efficacy does not depend on responses conditional on the previous latent mastery.

$$\begin{aligned}
& P(X_t = n | X_{t-1} = m; Z = z; \mathbf{A}_{1,t-1}, A_t = j; \mathbf{Y}_{1,t-1}, Y_t; t) \\
& = P(X_t = n | X_{t-1} = m; Z = z; \mathbf{A}_{1,t-1}, A_t = j; t) \quad \forall \mathbf{Y}_{1,t-1}, Y_t
\end{aligned}$$

It may strike some readers as odd to assume learner learns at the same rate whether or not she solves the problem or not. This is the exact critique of the performance factor analysis model (PFA) (Pavlik Jr et al., 2009). However, what stylized fact does the Assumption 2 fails to account for? The proponents of the PFA may argue that Assumption 2 does not generate a positive correlation between successes. This critique is not entirely correct because the successes are positive correlated unconditional on the latent mastery. More previous success implies higher mastery and consequently higher success rate in the future practice. Assumption 2 claims independence only after conditioning on the latent mastery. The proponents of the PFA may be right to argue that Assumption 2 does not generate high enough positive correlation with an LTP model of binary latent mastery. However, a larger magnitude of the positive correlation may be achieved by allowing the latent mastery to have more states and a positive correlation between efficacy and the state of latent mastery. In short, the Assumption 2 greatly reduces the complexity of parameter learning without significantly impairs the model’s explanatory power on a learning dataset.

Assumption 3: There is no complementarity or substitution effect in the item composition.

$$\begin{aligned}
& P(X_t = n | X_{t-1} = m; Z = z; \mathbf{A}_{1,t-1}, A_t = j; t) \\
& = P(X_t = n | X_{t-1} = m; Z = z; A_t = j; t) \quad \forall \mathbf{A}_{1,t-1}
\end{aligned}$$

Assumption 3 rules out scaffolding by preparing a learner with a string of easy problems to solve the final difficult problem. It also rules out a decreasing pedagogical efficacy in the case of naive repetition. If items in the sequence are nearly identical, the learner learns from

subsequent practice opportunities with equal probability even if she does not learn from the first attempt.

Assumption 4: The pedagogical efficacy is independent of the sequence position.

$$\begin{aligned} P(X_t = n | X_{t-1} = m; Z = z; A_t = j; t) \\ = P(X_t = n | X_{t-1} = m; Z = z; A_t = j) \quad \forall t \end{aligned}$$

Assumption 4 claims that a learner has the same probability of learning whether the item is the first one in the sequence or the last one. It essentially assumes that the learner is a paragon of grit and positive psychology. She is never frustrated by failures, never bored by repetition, and can focus as long as it takes.

1.1.3 The Learning Process

For any particular learner, the learning process is characterized by the evolution of mastery levels, $\{X_1, X_2, \dots, X_T\}$. When two consecutive levels differ ($X_t \neq X_{t+1}$), learning happens. The thesis does not take this approach to characterize the learning process. For one thing, learning is a probabilistic event, therefore any particular realization of learning (or lack thereof) is a noisy signal of practice efficacy. For another thing, an analyst may be interested in mastery at the population levels for each type of learners. Instead, this thesis characterizes the learning process as a change of density of mastery levels over time. Let the probability of the “typical” learner of type z has mastery level k at practice occasion t be $\pi_t^{z,k} = P(X_t = k | Z = z)$. Let the density of mastery for learner type z at occasion t be noted as $\mathbf{\Pi}_t^z = \{\pi_t^{z,0}, \dots, \pi_t^{z,M_X-1}\}$ where $\sum_{k=0}^{M_X-1} \pi_t^{z,k} = 1$ and $0 \leq \pi_t^{z,k} \leq 1$. The generic characterization of the learning process is $\mathbf{\Pi}_1^z, \dots, \mathbf{\Pi}_T^z$.

The generic characterization of the learning process masks the important role of item assignments. If practice efficacy is known, the expected learning process is deterministic

in the sense that it can be expressed explicitly by the prior belief of the density of the initial mastery (Π_1^z), practice efficacies ($\ell_j^{z;m,n}$) and item assignments ($\mathbf{A} = \{A_1, \dots, A_t\}$). In another word, given practice efficacies, it may be possible to formulate an item assignment \mathbf{A}^* that is optimal to a particular learner's estimated mastery profile $\widetilde{\Pi}_1^z$, which is the goal of the thesis.

Consider the learning process of a learner encountering item j at practice occasion t . If her starting level of mastery is known for sure, say m , her new expected mastery profile can be expressed as

$$\begin{aligned}\Pi_{t,j}^z &= P(X_t | X_{t-1} = n, A_t = j, Z = z) \\ &= \begin{cases} 0 & \text{if } X_t < m \\ \ell_j^{z;m,X_t} & \text{if } X_t \geq m \end{cases}\end{aligned}$$

Instead, her exact mastery of level is unknown, but the density of her mastery is known. The probability that she reaches a higher mastery, say n , can be calculated by computing the joint likelihood of mastery at two states then integrating out the previous mastery. The tricky part is that the starting density depends on the previous item assignments A_1, \dots, A_{t-1} , except for the first practice occasion where the prior belief is employed. To save the notation from an abuse of symbols, this thesis simply notes the item assignment scheme as \mathbf{A} . The new density is therefore

$$\begin{aligned}\pi_{1,\mathbf{A}}^{z,n} &= \pi_1^{z,n} \\ \pi_{t,\mathbf{A}}^{z,n} &= \sum_{X_{t-1}} P(X_t = n, X_{t-1} | \mathbf{A}) \quad (\forall t > 1) \\ &= \sum_{X_{t-1}} P(X_t = n | X_{t-1}, A_t = j) P(X_{t-1} | \mathbf{A}) \\ &= \sum_{k=0}^n \ell_j^{z;m,k} \pi_{t-1,\mathbf{A}}^{z,k} + \sum_{k=n+1}^{M_x-1} 0 \pi_{t-1,\mathbf{A}}^{z,k} \quad (1.1)\end{aligned}$$

The last term ($\sum_{k=n+1}^{M_x-1} 0\pi_{t-1,\mathbf{A}}^{z,k}$) of equation (1.1) is implied by the no forgetting assumption.

1.1.4 The Learning Process with Learner Engagement

One important aspect of the learner engagement is the effort decision (E_t), which is a binary variable. Motivated by the intuition of “no pain no gain”, this thesis makes a strong that learning does not happen unless a learner actually tries

Assumption 5:

$$\pi_{t,\mathbf{A}}^{z,n} = \pi_{t-1,\mathbf{A}}^{z,n} \quad \text{if } E_t = 0$$

With the addition of the effort decision, the full learning process can be described as

$$\pi_{t,\mathbf{A}}^{z,n} = \begin{cases} \pi_1^{z,n} & \text{if } t = 1 \\ \pi_{t-1,\mathbf{A}}^{z,n} + E_t \sum_{m=0}^{n-1} \ell_j^{z;m,k} \pi_{t-1,\mathbf{A}}^{z;m} & \text{if } t > 1 \end{cases}$$

1.2 The Statistical Inference Model

Because the mastery of a learner is not directly observed but has to be inferred from observed data, this section develops a statistical model to make the inference. Hidden Markov model (HMM) is a classical framework that describes a dynamic latent process. The HMM has a hidden layer and an observed layer. The hidden Layer describes the evolution of the latent state while the observed layer describes the generation of observed data. The Bayesian Knowledge Tracing model is a special case of the HMM model. This thesis extends the BKT model in two ways: the introduction of learner heterogeneity and the inclusion of learner engagement. The extended model is named as Learning Through Practice model (LTP) to differentiate it from the BKT model.

1.2.1 An Overview of Hidden Markov Model

Let the latent state be S , the observed data be O . Let the parameters of the hidden layer be Θ_S , the parameters of the observed layer be Θ_O . Let all the parameters be $\Theta = \{\Theta_S, \Theta_O\}$. The key challenge for HMM is to infer $P(\Theta_S, \Theta_O|O)$ when S is not observed. The key insight HMM is that S can first be augmented to produce full likelihood then be integrated out to get the posterior parameter distribution.

Let R_Θ denote the parameter space of Θ , $F(\Theta)$ denote the prior distribution of parameters. The HMM argues that posterior distribution of parameter $P(\Theta|O)$ can be obtained by

$$P(\Theta|O) = \frac{P(O, \Theta)}{\int_{R_\Theta} P(O, \Theta) dF(\Theta)}$$
$$P(O, \Theta) = \sum_S P(O, S, \Theta)$$

Whereas the latent variable is sampled from the posterior distribution

$$P(S|O, \Theta) = \frac{P(O, S, \Theta)}{P(O, \Theta)}$$

Obviously, the augmented likelihood ($P(O, S, \Theta)$) plays a key role in the HMM scheme. It is the combined result of the hidden layer and the observed layer. The hidden layer is characterized by the joint likelihood $P(S, \Theta_S)$. It is usually factored as

$$P(S, \Theta_S) = P(S|\Theta_S)P(\Theta_S)$$

where $P(S|\Theta_S)$ is informed by the theory about the hidden states while $P(\Theta_S)$ is the prior. The conditional likelihood $P(O|S, \Theta_O)$ characterizes the generation of the observed layer, where the dependence of S is made explicit. The complete-data likelihood is thus

$$P(O, S, \Theta) = P(O|S, \Theta_O)P(S|\Theta_S)P(\Theta)$$

where S is regarded as the mastery and learner type.

Here is a brief summary of the inference routine, which is described in much greater detail in Chapter Two:

- (1) From prior parameter distribution sample $\hat{\Theta}$
- (2) Initialize the latent state \hat{S} by $P(S|\Theta_S)$
- (3) Sample new parameters $\hat{\Theta}'$ by $P(\Theta|O)$
- (4) Sample new latent states \hat{S}' by $P(S|O, \Theta)$
- (5) Repeat step (3)-(4) until converges

1.2.2 Hidden Markov Model As A Learning Model

When applying the Hidden Markov model framework to the learning process, the latent states are the learner type and the learner mastery, while responses and other observed behavior are the observed data. The hidden layer describes the learning process or the evolution of latent mastery. The observed layer laid out the data generating process driven by the latent mastery. In the following analysis, the specific structure of S, O, Θ_S, Θ_O will be laid out explicitly.

It should be noted that although the mastery is an ordinal variable, the statistical model for the hidden layer is a latent class model, rather than an ordered multinomial logit model. The ordered logit uses the ranking of the cutoff points of the continuous latent variable to impose the order while here the no-regression assumption (Assumption 1) imposes the order. Readers familiar with the literature of Item Response Theory (IRT) may also wonder why the latent mastery is not continuous. The different choices of operational definition are reflected in the shape of item characteristic curve (ICC). The ICC of the LTP model is a step function,

whereas the ICC of the IRT model is a smooth sigmoid function. A priori, it is difficult to assert which ICC describes the true data generating process better. There are two scenarios in which the step function fits better. In the first scenario, the true mastery is a discrete variable and the true ICC is a step function. In the second scenario, the true mastery is a continuous variable but the true ICC is a mixture of sigmoid functions. In this case, a single sigmoid function may not fit as well as a flexible step function. As the number of discrete states increases, the step ICC will eventually outperform the single sigmoid ICC. Therefore, assuming the latent mastery as a discrete variable is not very restrictive as long as the number of states is allowed to vary.

1.2.3 The Bayesian Knowledge Tracing Model

The BKT model is a special instance of the Hidden Markov model. To describe the latent state (mastery), the initial state density and the state evolution (efficacy) needs to be specified. To describe the observed data (response), the observation matrix needs to be specified.

The BKT model adopts the simplest definition of practice efficacy possible. It only admits one learner type and two mastery states. Consequently, there exists one practice efficacy for each item $\ell_j^{1;0,1} = P(X_t = 1 | X_{t-1} = 0, Z = 1, A_t = j)$. For the rest of the thesis, denote the BKT efficacy as ℓ , according to the convention.

It is necessary to have a prior belief of the probability that a learner has mastery at the first practice occasion. Formally, such prior belief can be written as $\pi_1 = P(X_1 = 1)$. π_1 is called prior knowledge in the learning analytics literature, but initial density in this thesis to be consistent with the Markov Chain Monte Carlo algorithm introduced in Chapter Three. Once π and ℓ are known, the unconditional state density at each practice occasion t can be recursively defined as

$$\pi_t = (1 - \pi_{t-1})\ell + \pi_{t-1} \quad \forall t > 1$$

The inference of latent mastery is based on observed responses, Y_1, \dots, Y_T . The response of each practice occasion is a noisy measure of the concurrent mastery. The measure is noisy because a learner without mastery can answer correctly due to luck while a learner with mastery can answer incorrectly due to carelessness. The probability of a blind-luck correct is called guess rate, $g = P(Y_t = 1|X_t = 0)$. The probability of a careless incorrect is called slip rate $s = P(Y_t = 0|X_t = 1)$.

In short,

$$\begin{aligned} S &= \{X\} \\ O &= \{Y\} \\ \Theta_S &= \{\pi, \ell\} \\ \Theta_O &= \{s, g\} \end{aligned}$$

1.2.4 Learner Heterogeneity

The efficacy of the BKT model is homogeneous for all learners given the same item. The LTP model allows an item to have different effects on different learners. The learner heterogeneity has already been built into the working definition of practice efficacy. One type of heterogeneity is state heterogeneity: Given the same type (z) and same mastery goal (n), learners from different starting mastery state (k, m) progress at different speeds. $\ell_j^{z,k,n} \neq \ell_j^{z,m,n}$. The other type of heterogeneity is type heterogeneity: Given the same origin (m) and same mastery goal (n), learners of different types (z, z') progress at different speed. $\ell_j^{z,m,n} \neq \ell_j^{z',m,n}$.

The density of learner type is noted as $\alpha_z = P(Z = z)$, which sums up to 1. The generation of the observed response only depends on a learner's mastery level, not on her type. Intuitively,

it means that learners at the same level produce similar responses, no matter how fast they get to that level. The observed response only takes value 0/1 in this thesis. The conditional probability of the observed response is

$$c_j^k \equiv P(Y_t = 1 | X_t = k, A_t = j) \quad (1.1)$$

For example, in the BKT model, $g_j = c_j^0$ and $s_j = 1 - c_j^1$.

1.2.5 Learner Engagement

So far, this chapter has implicitly assumed that the learner can engage in learning for as long as the instructor wishes and as focused as the learning task requires. Both the duration and the intensity of learner engagement are imperfect in a low stake learning environment, which is typical of most applications of the intelligent tutoring system. The learner engagement is not only another set of observed data that can be used to infer latent mastery, but engagement also influences the learning process directly. Therefore, it is important to include learner engagement as a component of the LTP model. This subsection includes two aspects of learner engagement: The stop decision describes the duration of learner engagement and the effort decision describes the intensity of learner engagement.

1.2.5.1 Event Sequence

The addition of learner engagement makes the data generating process more complicated than the vanilla BKT process where only latent mastery and observed response are involved. The following event sequence clarifies the new learning process.

1. A learner is presented with a practice question.
2. The learner exerts a level of effort based on her state of latent mastery

-
3. The learner produces a response based on the effort level and her state of latent mastery.
 4. The learner receives feedback on the observed response.
 5. If the learner has exerted effort, she learns probabilistically. Otherwise, she does not learn.
 6. The learner can choose or be forced to exit. If the learner continues, repeat from (1); else data collection stops.

1.2.5.2 The Effort Decision

The effort (E_t) is assumed a binary choice with value 0 for not exerting effort and 1 for exerting effort. Conditional on the type of the learner, the state of latent mastery and the item id, the probability of exerting effort is constant.

$$\gamma_j^{z;k} \equiv P(E_t = 1 | Z = z, X_t = k, A_t = j) \quad (1.2)$$

Furthermore, assume that if the learner does not exert effort, the response is always completely incorrect.

Assumption 6: $P(Y_t = 0 | E_t = 0, A_t = j) = 1 \quad \forall j, t$

1.2.5.3 The Stop Decision

The stop decision (H_t) denote whether a learner stop practice sequence at occasion t . $H_t = 0$ means a learner continues to practice whiel $H_t = 1$ means a learner is no longer practicing. The hazard rate of occasion t (η_t) is the probability that a learner stops at occasion t conditional on she continues at $t - 1$.

$$\eta_t = P(H_t = 1 | H_{t-1} = 0)$$

Whether hazard rates depend on the response or the latent mastery merits a brief discussion. The issue is examined in details in Chapter Five. There are two types of stop decisions: stop-by-rule and stop-by-choice. An example of stop-by-rule is the “X-Strike” rule: A learner always keeps practicing unless she is forced to stop after accumulating X successes/failures. An example of stop-by-choice is the differential impact of boredom and frustration (Baker et al., 2010): A learner without mastery is frustrated by further practice while a learner with mastery is bored by it. In the case of stop-by-rule, the stop decision can only depend on the responses because they are what the system observes. In the case of stop-by-choice, the stop decision can reasonably depend on either the response or the latent mastery, but more likely the latter. Therefore, this thesis equates the stop-by-rule decision with a response dependent hazard model and the stop-by-choice decision with a mastery dependent hazard model. In the following analysis, Let the dependence states be S . $S = \{Z, X\}$ or $S = \{Y\}$.

This thesis assumes that the conditional hazard rate is independent of item characteristics. This is largely a by-product of the simple stop-by-rule accounting method: Most stop-by-rule decision weighs each correct or incorrect answer equally regardless of its item. A more nuanced item-dependent hazard function requires the knowledge of a specific learning product. For the sake of generalization, such feature is excluded from the LTP model

Assumption 7: $P(H_t = 1 | H_{t-1} = 0, S, A_t = j) = P(H_t = 1 | H_{t-1} = 0, S, A_t = l) \quad \forall j, l$

Furthermore, the functional form of hazard rate curve can be specified as parametric or nonparametric. A nonparametric hazard function is analogous to that of the observed response or effort choice. A different parameter is assigned to the hazard rate of each state at each practice occasion. Formally, the nonparametric model postulates

$$\eta_t^s \equiv P(H_t = 1 | H_{t-1} = 0, S_t = s) \quad (1.3)$$

In contrast, a parametric hazard function has one set of parameters for each state. The popular choice is the proportional hazard model, which allows for different base rates (λ_s) and different duration dependence (β_s) among states.

$$\eta_t^s \equiv \lambda_s e^{\beta_s t}$$

If the data generating process of the stop decision is parametric, both specifications are consistent but the parametric specification is efficient. Otherwise, the non-parametric specification is consistent while the parametric specification is inconsistent.

1.2.6 LTP as HMM

The previous two subsections expand the BKT model to cover both learner heterogeneity and learner engagement. This subsection summarizes the LTP model in terms of the ingredients of the HMM model.

$$S = \{X, Z\}$$

$$O = \{Y, E, H, A\}$$

$$\Theta_S = \{\pi, \ell, \alpha\}$$

$$\Theta_O = \{c, \gamma, \eta\}$$

1.2.7 An Example of Input Data

Table 1.1 shows an example of the input data, which is used in Chapter Six. The first column is user id. The second column is item id. The third column is the practice occasion, which

ranks the items into a sequence. The fourth column to the sixth column is observed data, responses, effort decisions and stop decisions respectively. The learner 30002 made a valid effort in the first item ($E_1 = 1$) but got it wrong ($Y_1 = 0$). She gave up on the second item ($Y_2 = E_2 = 0$). However, she made an attempt for the third item and got it right ($E_3 = Y_3 = 1$). The practice sequence is only three items long so the first two stop decision was no 0 ($H_1 = H_2 = 0$) while the third stop decision was yes ($H_3 = 1$).

Table 1.1: An Example of the LTP Input Data

User ID(i)	Item ID(j)	Occasion(t)	Response(Y)	Effort(E)	Stop(H)
30002	Q_10201056649366	1	0	1	0
30002	Q_10201058056988	2	0	0	0
30002	Q_10201056666357	3	1	1	1

1.3 Express Learning Theories with the LTP model

Because the LTP model relaxes the constraint of the restrictive structural representation of the Bayesian Knowledge Tracing model, it can have an arbitrary representation of the learning process, to the extent that the data allows for unique identification. To avoid abusing the LTP model as a pure data mining exercise, the representation of the learning process should be guided by learning theories. By the same token, because the LTP model is capable of expressing the testable implications of learning theories, the researcher can decide if the learning data fit her choice of the learning theory. Although this thesis does not carry out the empirical test, it is meaningful to point out the possibility of doing such test.

1.3.1 Zone of Proximal Development

The zone of proximal development (Vygotsky, 1978) postulates three types of relationship between the mental development of a learner and the mental requirement of a learning task. If the learner's mental development lags the mental requirement of the learning task, the learning is slow, the engagement is low, and the performance is poor. If the learner's mental development falls shorts of the requirement of the learning task if when she works independently but within her reach under the guidance or in collaboration, the learning speeds up, the engagement is high, and the performance improves. This is the zone of proximal development. If the learner's mental development exceeds the requirement of the task, the performance is high albeit there is no learning.

This theory naturally calls for a three-state latent mastery ($M_x = 3$). It has empirically testable implications on the observed response and the effort rate. If the learner is not ready to learn ($X = 0$), she will fail the problem with high probability, exert little effort, and has no chance to suddenly master the skill.

$$P(Y_t = 0|X_t = 0) \approx 1$$

$$P(E_t = 0|X_t = 0) \approx 1$$

$$P(X_t = 2|X_{t-1} = 0) \approx 0$$

If the learner has achieved mastery ($X = 2$), she will not fail the problem completely.

$$P(Y_t = 0|X_t = 2) \approx 0$$

If the learner is in the zone of the proximal development ($X = 1$), the learning process has no obvious constraints implied by Vygotsky's theory.

1.3.2 Reinforcement Learning

The observation that people repeat actions that reward them with pleasure and avoid actions that punish them with pain is well-established in behavioral science. Reinforcement learning is not only a stylized fact of how we learn (Anderson, 2000) but may also be the biological foundation of how we learn (Holroyd and Coles, 2002). In the context of learning through practice, reinforcement learning means that the more successes a learner get, the more engaged she is; the more failures a learner gets, the less engaged she is.

For the practice duration, if the stop decision depends on the latent mastery and the hazard rate is negatively correlated with the latent mastery, the LTP model generates the following pattern: The more successes the learner has in the preceding sequence, the more likely she is going to continue practice; the more failures the learner has in the preceding sequence, the less likely she is going to continue practice. This pattern arises because a higher rate of success implies a higher level of mastery, and consequently a lower probability to stop.

The same is true for the practice intensity. If the effort rate is positively correlated with the latent mastery, unconditional on the latent mastery, the more success the learner has enjoyed, the more effort she is likely to put into practice; the more failures the learner has suffered, the less effort she is likely to exert. This pattern arises because a higher rate of success implies a higher level of mastery and consequently a higher probability to exert effort.

1.4 Adaptive Practice Recommendation

The classical Bayesian Knowledge Tracing (BKT) model cannot generate adaptive practice recommendations within a knowledge/skill domain. This somewhat surprising result is a consequence of the homogeneous learning assumption imposed by the BKT model: If the learner has no mastery, she learns with a constant learning rate. If the learner has mastery, she does not learn. To highlight the problematic implication of this assumption, consider

an extreme case that a college freshman and a first grader learn first order differentiation. According to the assumption of the BKT model, if both learners have not learned the skill, they learn at the same rate for any given material.

This defect does not harm the Intelligent Tutoring Systems (ITSs) developed in the United States too much because major US ITSs focus on between domain objective individualization rather than within domain practice individualization. Use the math learning as an example. The ITS tries to sequence learning objectives that the learner needs to master in an optimal way. For instances, Khan Academy builds a knowledge tree and encourages learners to pass the test on parent nodes first before proceeding to their children nodes. The Cognitive Tutor follows a similar strategy to break a large learning goal into smaller skill-building blocks. Assessments and Learning in Knowledge Spaces (ALEKS) system adaptively changes the next learning objective given the learner's accomplished objectives. For objective individualization, the key question is what pre-requisite objectives the learner should achieve to ensure that the current objective is attainable (Doignon and Falmagne, 2012) and if such process does describe how people learn to acquire certain skills (Anderson, 2013). Within the learning objective, the practice question is usually generated by an algorithm rather than selected from a pre-existing content library. These computer-generated practice items are highly substitutable, thus the question is whether the practice engine is effective rather than which question the engine generated is most effective (Ritter et al., 2007). Therefore, practice individualization within a knowledge domain is not the major concern for the learning analytics community.

However, for some education service providers, the main problem is not objective individualization but practice individualization. Consider the problem of helping a teacher compile a homework from an existing content pool. The learning objective sequencing is set by the teacher (or the curriculum), while the item sequencing is left to the algorithm. The BKT model family is not very useful for solving the item sequencing problem, so some service providers, such as Knewton, resort back to the literature of computerized adaptive testing,

which may not be a good framework for instructional recommendations because it assumes no learning possible in the first place.

The LTP model addresses the problem of practice individualization by introducing learner's heterogeneous gain from practice. The construction of practice efficacy ($\ell_j^{z;m,n}$) defines two types of learner heterogeneity. The state heterogeneity refers to differential learning gains for learners with different levels of mastery conditional on the same learner type. The type heterogeneity refers to differential learning gains for learners with different types conditional on the same mastery level. Reconsider the previous “learning first order differentiation” example. With three levels of mastery inspired by Vygotsky's zone of proximal development, the first grader is defined as level-1 mastery who is not ready to learn, the college freshman is defined as level-2 mastery who is ready to learn but not proficient in the skill yet. The instruction has little efficacy on the level 1 learners but some efficacy on the level-2 learners. This is called state heterogeneity. For the college freshmen, some may be more versed in the mathematical reasoning than others, therefore the instruction may have different efficacies on learners with different susceptibilities. This is called type heterogeneity.

The learner engagement component of the LTP model introduces another dimension to practice individualization. Because item differs in its effort appeal condition on the type and the mastery state of the learner, there exists state heterogeneity and type heterogeneity in the effort appeal of the item as well. In general, the optimal practice item is both high in practice efficacy and strong in effort appeal. Heterogeneities on both aspects can give rise to a rich set of recommendation strategies.

Chapter 2

Model Estimation

The previous chapter outlines the Learning Through Practice model as a Hidden Markov Model (HMM). This chapter details the Markov Chain Monte Carlo (MCMC) algorithm that estimates the HMM parameters. The general strategy of estimating a hidden Markov model (HMM) with MCMC is to first augment the observed data with latent states given parameters then update parameters with Gibbs sampler given the augmented data.

To review, let S denote the latent states, O the observed data, Θ_S the parameters of the hidden layer, Θ_O the parameters of the observed layer, and Ψ the prior distribution of parameters. The goal is to estimate $P(\Theta_S, \Theta_O | O)$. The challenge is that latent states are not observed in the dataset. The solution that HMM proposes is to augment the latent states first to calculate the full likelihood then integrate them out later to get the posterior parameter distribution.

If S is augmented, the posterior parameter distribution can be calculated as

$$P(O, S, \Theta) = P(O|S, \Theta_O)P(S|\Theta_S)\Psi(\Theta)$$

$$P(O, \Theta) = \sum_S P(O, S, \Theta)$$

$$P(\Theta|O) = \frac{P(O, \Theta)}{\int_{R_\Theta} P(O, \Theta) dF(\Theta)}$$

The latent states S are either sampled from the prior distribution at the inception of the MCMC algorithm or sampled from the posterior distribution of states by

$$P(S|O, \Theta) = \frac{P(O, S, \Theta)}{P(O, \Theta)}$$

The latent states can be sampled by two algorithms: the Brute Force algorithm (BF) or the Forward Recursion Backward Sampling algorithm (FRBS). The BF algorithm is intuitive and flexible while the FRBS algorithm is computational economic. The parameters are drawn according to Gibbs sampler. Gibbs sampler iteratively draws parameters from their full conditional likelihood. When the posterior distribution is conjugate to a prior distribution that is easy to sample from, the Gibbs sampler is very efficient. When the posterior distribution has no conjugate prior, the parameters are drawn by Adaptive Rejection Sampling algorithm (ARS). The ARS algorithm avoids the costly evaluation of the full conditional likelihood while still guarantees to sample from the correct posterior distribution.

2.1 Prior

Most parameters of the LTP model follow either multinomial distribution or Bernoulli distribution. Since Bernoulli distribution is a special case of multinomial distribution, without loss of generality, this section discusses only the prior distribution for multinomially distributed parameters. They include:

-
- (1) The mixture density of learner type, $(\alpha_1, \dots, \alpha_{M_Z})$. $0 < \alpha < 1$
 - (2) The initial state density of learner type z , $(\pi_0^z, \dots, \pi_{M_X-1}^z)$, $0 < \pi < 1$
 - (3) The practice efficacy of item j conditional on previous latent mastery state k of learner type z , $(\ell_j^{z;k,k+1}, \dots, \ell_j^{z;k,M_X-1})$. It should be noted that the length of the density vector depends on k .
 - (4) The conditional response density of item j for mastery level k (see eq 1.1), $(1 - c_j^k, c_j^k)$.
 - (5) The conditional density of item j for mastery level k of learner type z (see eq 1.2), $(1 - \gamma_j^{z;k}, \gamma_j^{z;k})$.
 - (6) The conditional hazard rate of item j at practice occasion t conditional on the dependence state s (see eq 1.3), $(1 - \eta_t^s, \eta_t^s)$. It should be noted that this prior is only used when the hazard function is assumed to nonparametric.

The prior distribution of the parameters listed above is Dirichlet distribution $Dir(0.5, \dots, 0.5)$. The dimension of the prior distributions match that of the parameters. This prior is a special of the Jeffery prior that is non-informative and invariant to scaling of the parameters.

The prior distribution of the proportional hazard model parameters, i.e. the baseline hazard rate and the duration dependence, is the uniform distribution. The range of distribution cannot be set *a priori* because the hazard rate at the end of the sequence must be no larger than 1. This chapter revisits this issue in details when describing the Adaptive Rejection Sampling (ARS) algorithm.

2.2 Full Conditional Likelihood of Augmented Data

For HMM model, the two key posterior distributions that support the iteration of sampling is

$$P(\Theta|O) = \frac{P(O, \Theta)}{\int_{R_\Theta} P(O, \Theta) dF(\Theta)} \propto P(O, \Theta)$$

$$P(S|O, \Theta) = \frac{P(O, S, \Theta)}{P(O, \Theta)}$$

Therefore, both posterior distributions depend on the calculation of $P(O, S, \Theta)$. Given parameters, the key ingredient is the conditional likelihood of augmented data ($P(O, S|\Theta)$).

To review, the HMM structure of the LTP model is

$$S = \{\mathbf{X}, Z\}$$

$$O = \{\mathbf{Y}, \mathbf{H}, \mathbf{E}, \mathbf{A}\}$$

$$\Theta_S = \{\pi, \ell, \alpha\}$$

$$\Theta_O = \{c, \gamma, \eta\}$$

where $\mathbf{Y} = Y_1, \dots, Y_T$, $\mathbf{E} = E_1, \dots, E_T$, $\mathbf{H} = H_1, \dots, H_T$, and $\mathbf{X} = X_1, \dots, X_T$. The conditional likelihood of the augmented data is

$$\begin{aligned} P(O, S|\Theta) &= P(\mathbf{Y}, \mathbf{A}, \mathbf{E}, \mathbf{H}, \mathbf{X}, Z|\Theta) \\ &= P(\mathbf{H}|Z, \mathbf{X}, \mathbf{Y}, \Theta)P(\mathbf{Y}|Z, \mathbf{X}, \mathbf{E}, \mathbf{A}, \Theta)P(\mathbf{E}, \mathbf{X}, \mathbf{A}|Z, \Theta)P(Z|\Theta) \end{aligned}$$

This thesis assumes that item assignments are exogeneous to the analysis. To simplify the notation, omit \mathbf{A} for the rest of the chapter. Therefore, the conditional likelihood of the augmented data is

$$P(O, S|\Theta) = P(\mathbf{H}|Z, \mathbf{X}, \mathbf{Y}, \Theta)P(\mathbf{Y}|Z, \mathbf{X}, \mathbf{E}, \Theta)P(\mathbf{E}, \mathbf{X}|Z, \Theta)P(Z|\Theta) \quad (2.1)$$

In contrast, the conditional likelihood of the observed data marginalizes over the latent states from $P(O, S|\Theta)$.

$$\begin{aligned}
P(O|\Theta) &= \sum_Z \sum_{\mathbf{X}} P(O, S|\Theta) \\
&= \sum_{Z=1}^{M_Z} (P(Z) \sum_{X_1} \cdots \sum_{X_T} P(\mathbf{Y}, \mathbf{E}, \mathbf{H}, \mathbf{X}, Z|\Theta))
\end{aligned}$$

Therefore, the MCMC algorithm hinges on how to compute equation 2.1. The following analysis breaks it into separate parts.

The conditional hazard rate depends on the dependence structure and the functional form. If the stop decision depends on the latent mastery, the proportional hazard rate model is

$$P(\mathbf{H}|Z, \mathbf{X}, \Theta) = \prod_{t=1}^T \prod_{k=0}^{M_X-1} (\lambda_{z;k} e^{\beta_{z;k} t})^{I(H_t=1, X_t=k, Z=z)} (1 - \lambda_{z;k} e^{\beta_{z;k} t})^{I(H_t=0, X_t=k, Z=z)}$$

The nonparametric hazard rate model is

$$P(\mathbf{H}|Z, \mathbf{X}, \Theta) = \prod_{t=1}^T \prod_{k=0}^{M_X-1} (\eta_t^{z;k})^{I(H_t=1, X_t=k, Z=z)} (1 - \eta_t^{z;k})^{I(H_t=0, Z=z, X_t=k)}$$

Similarly, if the stop decision depends on the response, the nonparametric hazard rate are

$$P(\mathbf{H}|\mathbf{Y}, \Theta) = \prod_{t=1}^T \prod_{r=0}^1 (\eta_t^r)^{I(H_t=1, Y_t=r)} (1 - \eta_t^r)^{I(H_t=0, Y_t=r)}$$

The conditional likelihood of observed response

$$P(\mathbf{Y}|Z, \mathbf{X}, \mathbf{E}, \Theta) = \prod_{t=1}^T \prod_{k=0}^{M_X-1} \prod_{j=1}^J (1 - c_j^k)^{I(A_t=j, Y_t=0, X_t=k, E_t=1)} (c_j^k)^{I(A_t=j, Y_t=1, X_t=k, E_t=1)}$$

The joint conditional likelihood of effort and latent mastery is

$$\begin{aligned}
P(\mathbf{X}, \mathbf{E}|Z, \Theta) &= P(X_1|Z, \Theta) \prod_{t=1}^T P(E_t|X_t, Z, \Theta) \prod_{t=2}^T P(X_t|X_{t-1}, E_{t-1}, Z, \Theta) \\
P(X_1|Z, \Theta) &= \prod_{k=0}^{M_X-1} (\pi^{z;k})^{I(X_1=k, Z=z)} \\
P(E_t|X_t, Z, \Theta) &= \prod_{j=1}^J (\gamma_j^{z;k})^{I(A_t=j, Z=z, X_t=k, E_t=1)} (1 - \gamma_j^{z;k})^{I(A_t=j, Z=z, X_t=k, E_t=0)} \\
P(X_t|X_{t-1}, E_{t-1}, Z, \Theta) &= \prod_{j=1}^J \left\{ \left[\prod_{k=1}^{M_X-2} \left(1 - \sum_{n=k+1}^{M_X-1} \ell_j^{z;k,n} \right) 1^{I(A_{t-1}=j, X_{t-1}=k, X_t=n, Z=z, E_t=1)} \right] \right. \\
&\quad \left. \left[\prod_{m=1}^{M_X-2} \prod_{n=k+1}^{M_X-1} (\ell_j^{z;m,n})^{I(A_{t-1}=j, X_{t-1}=m, X_t=n, Z=z, E_t=1)} \right] \right\}
\end{aligned}$$

Last but not least, $P(Z|\Theta) = \alpha_z$, which is the state mixture density.

2.3 State Augmentation

2.3.1 Latent Mastery Augmentation

This subsection describes two ways of augmenting the state of latent mastery. The brute force algorithm computes the likelihood of the augmented data, marginalizes over the nuance states, imputes the conditional state transition probability, and draws states accordingly. In contrast, the forward recursion and backward sampling algorithm computes the local conditional state transition probability recursively and samples the state accordingly. Both methods sample the states backward for better state mixture (Scott, 2002).

2.3.1.1 The Brute Force Algorithm

Given parameters Θ , the joint likelihood $P(\mathbf{Y}, \mathbf{E}, \mathbf{H}, \mathbf{X}|Z, \Theta)$ can be calculated. Thus it is trivial to calculate the following quantities:

$$P(X_t = n, Z, O, \Theta) = \sum_{X_1} \cdots \sum_{X_{t-1}} \sum_{X_{t+1}} \cdots \sum_{X_T} P(\mathbf{X}, Z, O, \Theta)$$

$$P(X_{t-1} = m, X_t = n, Z, O, \Theta) = \sum_{X_1} \cdots \sum_{X_{t-2}} \sum_{X_{t+1}} \cdots \sum_{X_T} P(\mathbf{X}, Z, O, \Theta) \quad (2.2)$$

With the joint probability, calculate the conditional probability used in sampling

$$P(X_T = n|Z, O, \Theta) = \frac{P(X_T = n, Z, O, \Theta)}{\sum_{k=0}^{M_X-1} P(X_T = k, Z, O, \Theta)}$$

$$P(X_{t-1} = m|, X_t = n, Z, O, \Theta) = \frac{P(X_{t-1} = m, X_t = n, Z, O, \Theta)}{P(X_t = n, Z, O, \Theta)}$$

The state is sampled by the following steps:

1. Draw the state at the end of the sequence (T) from a multinomial distribution with probability mass function $P(X_T = k|Z, O, \Theta)$
2. Given the state drew at next sequence ($t+1$), draw the current state from a multinomial distribution with probability mass function $P(X_{t-1} = m|, X_t = n, Z, O, \Theta)$

2.3.1.2 The Forward Recursion and Backward Sampling Algorithm

In essence, the Forward Recursion Backward Sampling (FRBS) algorithm is identical to the brute force algorithm. However, instead of exhausting all the state combinations and marginalizing at each sequence, the FRBS algorithm uses a recursive formula and the conditional independence property of first-order Markov chain to reduce the computation complexity.

The key observations is that it is not necessary to condition on all observed variables in equation 2.2. Because of the first order markov independence,

$$X_t \perp\!\!\!\perp \mathbf{X}_{t+2,T} | X_{t+1}, Z$$

Therefore

$$X_t \perp\!\!\!\perp \mathbf{Y}_{t+2,T}, \mathbf{E}_{t+2,T}, \mathbf{H}_{t+2,T} | X_{t+1}, H_t, Z$$

Therefore

$$P(X_t | X_{t+1}, Z, O, \Theta) = P(X_t | X_{t+1}, Z, \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{H}_{1,t+1}, \Theta)$$

it is sufficient to sample backward according to

$$P(X_t | X_{t+1}, \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{H}_{1,t+1}, Z, \Theta) = \frac{P(X_t, X_{t+1} | \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{H}_{1,t+1}, Z, \Theta)}{\sum_{X_{t+1}} P(X_t, X_{t+1} | \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{H}_{1,t+1}, Z, \Theta)}$$

The key is to calculate the partial conditional joint state density:

$$P(X_t, X_{t+1} | \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{H}_{1,t+1}, Z, \Theta)$$

This quantity can be calculated by recursive method. Define the partial conditional marginal state density $\tilde{\pi}_t^{z;k}$ and the partial conditional joint state density $\tilde{p}_t^{z;m,n}$.

$$\tilde{\pi}_t^{z;k} = P(X_t = k | Z, \mathbf{Y}_{1,t}, \mathbf{E}_{1,t}, \mathbf{A}_{1,t}, \mathbf{H}_{1,t}, \Theta)$$

$$\tilde{p}_{t+1}^{z;m,n} = P(X_t = m, X_{t+1} = n | Z, \mathbf{Y}_{1,t+1}, \mathbf{E}_{1,t+1}, \mathbf{A}_{1,t+1}, \mathbf{H}_{1,t+1}, \Theta)$$

The recursive algorithm is:

1. Given $\tilde{\pi}_t^m$, calculate the partial conditional joint density

$$\tilde{p}_{t+1}^{z;m,n} = \tilde{\pi}_t^{z;m} P(X_{t+1} = n | X_t = m, E_{t-1}, Z)$$

$$P(Y_{t+1} | X_{t+1}, E_{t+1}) P(E_{t+1} | X_{t+1} = m, Z) P(H_{t+1} | X_{t+1} = m, Z, H_t = 0)$$

2. Given $\tilde{p}_{t+1}^{z;m,n}$, calculate the partial conditional marginal density

$$\tilde{\pi}_{t+1}^{z;n} = \sum_{m=0}^{M_X-1} \tilde{p}_{t+1}^{z;m,n}$$

The sampling algorithm is:

1. Draw the state at the end of the sequence (T) from a multinomial distribution with probability mass function $P(X_T = k | Z) = \tilde{\pi}_T^{z;k}$
2. Given the state of the next sequence position ($t + 1$), draw the current state from from a multinomial distribution with probability mass function $P(X_t = m | Z = z) = \frac{\tilde{p}_{t+1}^{z;m,n}}{\sum_{n=0}^{M_X-1} \tilde{p}_{t+1}^{z;m,n}}$ given $X_{t+1} = n$ of learner type Z .

2.3.2 Learner Type Augmentation

As long as the conditional likelihood of the observed data can be computed, augment the learner type is trivial, because it merely samples from a multinomial distribution with probability mass function:

$$P(Z = z | O, \Theta) = \frac{\alpha_z P(O | \mathbf{A} Z = z, \Theta)}{\sum_{w=1}^{M_Z} (\alpha_w P(O | \mathbf{A}, Z = w, \Theta))}$$

2.4 Parameter Update

The parameters are updated by the Gibbs sampler. Other than parameters of the parametric hazard model, the conjugate posterior distribution for all parameters is the Dirichlet

distribution, from which it is easy to sample. Although the parameters of the parametric hazard model cannot be sampled easily, it can still be drawn from the marginal conditional distribution by the adaptive rejection sampling.

2.4.1 Conjugate Posterior

The mixture density is drawn from the posterior distribution $Dir(a_{Z_1}, \dots, a_{Z_{M_Z}})$ where $a_{Z_k} = 1 + \sum_{i=1}^N I(Z_1^i = k)$.

The initial state density of learner type z is drawn from the posterior distribution $Dir(a_{X_0;z}, \dots, a_{X_{M_X-1};z})$ where $a_{X_k;z} = 1 + \sum_{i=1}^N I(X_1^i = k, Z^i = z)$.

The effort rates conditional on the latent state mastery ($X = k$) of learner type z are drawn from the posterior distribution $Beta(a_{E_0;z}, a_{E_1;z})$ where $a_{E_e;z} = 1 + \sum_{t=1}^T \sum_{i=1}^N I(E_t = e, X_t^i = k, Z^i = z)$.

The response rates conditional on the latent state mastery ($X = k$) are drawn from the posterior distribution $Dir(a_{Y_0}, a_{Y_1})$ where $a_{Y_r} = 1 + \sum_{t=1}^T \sum_{i=1}^N I(Y_t^i = r, X_t^i = k, E_t = 1)$.

When the hazard rate curve is specified as non-parametric, the hazard rate conditional on the latent mastery k of learner type z is drawn from the posterior distribution $Beta(a_{H_{t,0}^{z;k}}, a_{H_{t,1}^{z;k}})$ where $a_{H_{t,h}^{z;k}} = 1 + \sum_{i=1}^N I(X_t^i = k, Z^i = z, H_t = h)$. Similarly, the hazard rate conditional on the response r is drawn from $Beta(a_{H_{t,0}^r}, a_{H_{t,1}^r})$ where $a_{H_{t,h}^r} = 1 + \sum_{i=1}^N I(Y_t^i = r, H_t = h)$.

2.4.2 Adaptive Rejection Sampling (ARS)

It is expensive to sample from the posterior distribution when the model is the discrete-time proportional hazard model with time-varying covariates. This problem is first solved by Dellaportas and Smith(1993) with the Adaptive Rejection Sampling algorithm(ARS)(Gilks and Wild, 1992). As long as the target likelihood function is log-concave, one can sample

the posterior distribution by drawing from the interval of two piecewise spline functions. By constructing an upper hull and a lower hull to sandwich the true posterior distribution, the ARS algorithm reduces the computational cost to draw from a non-standard distribution, compared to the standard rejection method.

Here is a short description of the ARS algorithm.

1. Choose a few values x_j from the domain. Construct the upper hull and the lower hull of the target distribution function $f(x)$ by piecewise linear functions of

$$u(x) = f(x_j) + f'(x_j)(x - x_j)$$

$$l(x) = \frac{(x_{j+1} - x)f(x_j) + (x - x_j)f(x_{j+1})}{x_{j+1} - x_j}$$

defined over intervals $x \in (z_{j-1}, z_j)$ where

$$z_j = \frac{f(x_j) - f(x_{j+1}) - x_{j+1}f'(x_{j+1}) + x_jf'(x_j)}{f'(x_j) - f'(x_{j+1})}$$

2. Sample new value of x^* by the probability of $s(x)$ where

$$s(x) = \frac{\exp(u(x))}{\int_{D_x} \exp(u(x)) dx}$$

3. Sample w independently from $\text{uniform}(0,1)$. Accept the new value x^* if

$$w \leq e^{l(x^*) - u(x^*)}$$

Otherwise, accept the new value x^* if

$$w \leq e^{f(x^*) - u(x^*)}$$

Otherwise reject x^* and draw again.

4. If x^* is accepted, add to the list of x_j for the next draw.

Because the target distribution function is concave, it follows that $f'(x_1) > 0$ and $f'(x_J) < 0$. The initial value thus cannot be sampled randomly.

The following theorem proves that the augmented data likelihood is log-concave. Therefore, the adaptive rejection sampling algorithm can be applied to update the parameter.

Theorem 2.1. *The full conditional likelihood function is log-concave*

Proof. For λ

$$\begin{aligned}\frac{\partial \ell}{\partial \lambda_{z,k}} &= \sum_{i=1}^N \sum_{t=1}^{T_i} I(X_t^i = k) \left[-\frac{(1 - H_{i,t})e^{\beta_{z,k}t}}{1 - \lambda_{z,k}e^{\beta_{z,k}t}} + \frac{H_{i,t}}{\lambda_{z,k}} \right] \\ \frac{\partial^2 \ell}{\partial \lambda_{z,k}^2} &= \sum_{i=1}^N \sum_{t=1}^{T_i} I(X_t^i = k) \left[-\frac{(1 - H_{i,t})e^{2\beta_{z,k}t}}{(1 - \lambda_{z,k}e^{\beta_{z,k}t})^2} - \frac{H_{i,t}}{\lambda_{z,k}^2} \right]\end{aligned}$$

Because $H_{i,t} \geq 0$, $1 - H_{i,t} \geq 0$, $e^{\beta_{z,k}} \geq 0$. Therefore, $\frac{\partial^2 \ell}{\partial \lambda_{z,k}^2} < 0$.

For $\beta_{z,k}$

$$\begin{aligned}\frac{\partial \ell}{\partial \beta_{z,k}} &= \sum_{i=1}^N \sum_{t=1}^{T_i} I(X_t^i = k, Z^i = z) \left[-\frac{(1 - H_{i,t})\lambda_{z,k}e^{\beta_{z,k}}}{1 - \lambda_{z,k}e^{\beta_{z,k}}} + H_{i,t} \right] t \\ \frac{\partial^2 \ell}{\partial \beta_{z,k}^2} &= \sum_{i=1}^N \sum_{t=1}^{T_i} -I(X_t^i = k, Z^i = z) \left[\frac{1}{1 - \lambda_{z,k}e^{\beta_{z,k}t}} + \frac{e^{\beta_{z,k}t}\lambda_{z,k}}{(1 - \lambda_{z,k}e^{\beta_{z,k}t})^2} \right] e^{\beta_{z,k}t} t^2 (1 - H_{i,t}) \lambda_{z,k}\end{aligned}$$

Because $\lambda e^{\beta_{z,k}t} < 1$ by definition and $\lambda_{z,k} \geq 0$, $\frac{1}{1 - \lambda e^{\beta_{z,k}t}} + \frac{e^{\beta_{z,k}t}\lambda_{z,k}}{(1 - \lambda e^{\beta_{z,k}t})^2} > 0$. Furthermore, $1 - H_{i,t} > 0$ for some i , $\frac{\partial^2 \ell}{\partial \beta_{z,k}^2} < 0$. \square

However, there are two additional issues. First, the range of $\lambda_{z,k}$ and $\beta_{z,k}$ is constrained because the hazard rate is strictly less than 1 for sequence max to T . Given T and $\lambda_{z,k}$, $\beta_{z,k} \in (-\infty, \frac{\log(\lambda_{z,k})}{T})$; Given T , $\lambda_{z,k} \in (0, \frac{1}{e^{\beta_{z,k}T}})$. The range is then passed into the ARS algorithm to ensure that parameters drawn are always valid. Second, when the number of observations grows, the algorithm may experience numerical overflow. To prevent this, scale

the log likelihood to be larger than -3000.

The prior distribution of the parameters is chosen to be the uniform distribution to facilitate the posterior draw and justify assigning zero mass to a certain interval of the parameter space to enforce the range constraints. Unfortunately, the lower bound only exists for the $\lambda_{z,k}$. Set the lower bound of $\beta_{z,k}$ to be -1. Start the draw from the ϵ from the boundary, where $\epsilon = 0.01$. If the initial draws produce a numerical overflow, symmetrically narrow the bound by a step size of 0.1 until a valid draw occurred.

2.5 Label Switching

Label switching is a common problem in the mixture model. Consider a general mixture model with K components (C_1, \dots, C_K) , each associated with a parameter set $\Theta_{C_1}, \dots, \Theta_{C_K}$. Because the labels of the components are arbitrary, permutation of the labels produces identical likelihood.

Rank order condition is one solution to the label switching problem. The non-regressive state assumption is not sufficient to prevent label switching, even for the two-state case. When $M_X = 2, M_Y = 2$, it is conventional (Corbett and Anderson, 1994) to assume that the correct rate is positively correlated with the mastery.

$$P(Y_t = 1|X_t = 0) < P(Y_t = 1|X_t = 1) \leftrightarrow c_j^0 < c_j^1$$

When $M_Y = 2$, the previous rank order conditions can be generalized as

$$P(Y_t = 1|X_t = m) < P(Y_t = 1|X_t = n) \quad \forall m < n$$

As for the learner type, assume that the type with higher value also has higher initial mastery.

To wit:

$$P(X_1 = 1|Z = w) < P(X_1 = 1|Z = v) \quad \forall w < v$$

2.6 Posterior Inference

After the parameters are estimated ($\hat{\Theta}$), The major interest of posterior inference of the LTP model is the posterior distribution of the type (α_t^z) and the latent mastery (π_t^z) of ONE learner at sequence t given the posterior distribution of learner's type (α_{t-1}^z) and latent mastery (π_{t-1}^z) at sequence $t - 1$, the observed response (Y_t), the stop decision(H_t), and the effort decision(E_t). This problem can be solved by the following iterative formula through tracking the posterior distribution of latent mastery for each learner type:

- (1) For given learner type z , the posterior distribution of latent mastery after at least one observation is

$$\pi_t^{z;k} = \frac{P(X_t = k, Y_t, E_t, H_t | \hat{\Theta}, \pi_{t-1}^z, Z = z,)}{\sum_{m=0}^{M_X-1} P(X_t = m, Y_t, E_t, H_t | \hat{\Theta}, \pi_{t-1}^z, Z = z)}$$

- (2) Given the updated posterior distribution of latent mastery, update the posterior distribution of learner type

$$\alpha_t^z = \frac{\alpha_{t-1}^z P(Y_t, E_t, H_t | \hat{\Theta}, \pi_t^z, Z = z)}{\sum_{w=1}^{M_Z} \alpha_{t-1}^w P(Y_t, E_t, H_t | \hat{\Theta}, \pi_t^w, Z = w)}$$

The other possible interest of posterior inference is the observed learning curve $P(Y_t = r | H_{t-1} = 0)$. This problem is more difficult than the posterior inference on one learner because it needs to marginalize out learner type, preceding responses and preceding effort choices. The problem can be solved by the following iterative formula:

Let $\xi_{r,e,q,h} = P(Y_t = r, E_t = e, H_{t-1} = q, H_t = h | \hat{\Theta}, X_t, X_{t-1})$, the posterior inference can be computed by the following recursive formula

$$\begin{aligned}
P(X_t = n | H_{t-1} = 0, Z) &= \frac{\sum_r \sum_e \sum_m \sum_h \xi_{r,e,0,h} \hat{\ell}^{z;m,n} P(X_{t-1} = m | Z, H_{t-2} = 0)}{\sum_r \sum_e \sum_m \sum_h \sum_q \xi_{r,e,q,h} \hat{\ell}^{z;m,n} P(X_{t-1} = m | Z, H_{t-2} = 0)} \\
P(X_t = n | H_{t-1} = 0) &= \sum_z \alpha_z P(X_t = n | H_{t-1} = 0, Z) \\
P(Y_t = r | H_{t-1} = 0) &= \sum_k P(Y = r | X = k) P(X_t = k | H_{t-1} = 0)
\end{aligned}$$

2.7 Simulation

This section provides a simulation study to show the convergence property of the MCMC algorithm. The MCMC algorithm works well with the single learner type, but not so well with multiple types of learners.

2.7.1 Single Learner Type

To demonstrate the model's ability to identify beyond the binary states, the simulation set $M_X = 3$ and $M_Y = 3$. The initial learner mastery heavily clustered in the lower state.

To demonstrate the multiple-item efficacy identification, the simulation has two items $J = 2$. The first item has a low transition rate from low mastery ($X = 0$) to high mastery ($X = 2$) with $\ell_1^{0,2} = 0.3$ but high transition rate from medium mastery ($X = 1$) to high mastery with $\ell_1^{1,2} = 0.6$. The second item has the reverse pattern with $\ell_2^{0,2} = 0.5$ and $\ell_2^{1,2} = 0.3$. The first item appears 30% of the time.

The first item has a low half-correct rate in the low mastery and high half-correct rate in the medium mastery. The second item has the reverse pattern. Both effort decision and exit decision are present in the simulation. The effort rate is positively correlated with mastery state for the two items. The hazard rate is mastery-dependent. The baseline hazard rate

is 0.1 for all states but the hazard rate grows faster for the low mastery than for the high mastery. The values of all simulation parameters are in Appendix B.1

Table 2.1 only reports the estimated pedagogical efficacy. Both parametric and non-parametric specifications are fit to the model. Both specifications report reasonably good point estimation. The parametric model has slightly tight 95% credible interval, confirming that it is more efficient when the model specification is right.

Table 2.1: Estimated Pedagogical Efficacy of the LTP Models

Specification	Parameter	Item	True	Point Est.	95%CI(L)	95%CI(H)
non-parametric hazard	l01	1	0.3	0.320	0.164	0.482
non-parametric hazard	l02	1	0.2	0.303	0.207	0.397
non-parametric hazard	l12	1	0.6	0.616	0.529	0.709
non-parametric hazard	l01	2	0.3	0.368	0.190	0.562
non-parametric hazard	l02	2	0.5	0.467	0.326	0.601
non-parametric hazard	l12	2	0.3	0.298	0.153	0.423
proportional hazard	l01	1	0.3	0.289	0.165	0.431
proportional hazard	l02	1	0.2	0.308	0.214	0.391
proportional hazard	l12	1	0.6	0.615	0.522	0.702
proportional hazard	l01	2	0.3	0.344	0.170	0.502
proportional hazard	l02	2	0.5	0.457	0.329	0.581
proportional hazard	l12	2	0.3	0.286	0.163	0.412

2.7.2 Multiple Learner Types

The current rank order conditions are insufficient for the MCMC algorithm to converge to the true parameters in a mixture model with multiple types of learners. The second simulation demonstrates that the point estimation of the MCMC algorithm fails to converge because of the bimodal posterior distribution of the parameters.

The second simulation does not include the learner engagement. The number of states of latent mastery and the observed response is both two ($M_X = M_Y = 2$). There are two items and two learner types. For all learner types, the second item has much higher efficacy than the first item ($\ell_2^{z;0,1} > \ell_1^{z;0,1}$). The first item has higher efficacy for the type 1 learner ($\ell_1^{1;0,1} > \ell_1^{2;0,1}$) while the second item has higher efficacy for the type 2 learner ($\ell_2^{2;0,1} > \ell_2^{1;0,1}$). Both items are a relatively clean measure of the latent mastery. The numeric values of all simulation parameters are in Appendix B.2.

Table 2.2 reports the global point estimations. The efficacies to the type 1 learner are estimated with precision. However, the efficacies to the type 2 learner and the mixture density are not precisely measured. The point estimation of the mixture density is close to the true value by coincidence.

Table 2.2: Estimated Parameter of the Mixture Model
(Global)

Parameter	True	Point Estimation	95%CI(L)	95%CI(H)
Item 1 Efficacy Type 1	0.3	0.289	0.138	0.755
Item 2 Efficacy Type 1	0.7	0.701	0.460	0.882
Item 1 Efficacy Type 2	0.2	0.457	0.098	0.934
Item 2 Efficacy Type 2	0.8	0.668	0.192	0.969
mixture density	0.6	0.597	0.032	0.962

The posterior distribution of the mixture density is bimodal. The higher mode centers around 0.7. The lower mode centers around 0.2. Table 2.3 and 2.4 separately report the point estimations of the parameters for each mode. The posterior distribution of the efficacies of the type 2 learner is almost flat in the higher mode. In contrast, the posterior distribution of the efficacies of the type 1 learner is almost flat in the lower mode while that of the type 2 learner are better estimated. A closer look that posterior distribution mix shows that the higher mode comes from three chains while the lower mode comes from one chain. It appears that MCMC algorithm is trapped in a local optimal where one type is precisely estimated while the other type remains obscure.

Table 2.3: Estimated Parameter of the Mixture Model
(Higher Mode)

Parameter	True	Point Estimation	95%CI(L)	95%CI(H)
Item 1 Efficacy Type 1	0.3	0.243	0.169	0.316
Item 2 Efficacy Type 1	0.7	0.708	0.630	0.777
Item 1 Efficacy Type 2	0.2	0.524	0.084	0.950
Item 2 Efficacy Type 2	0.8	0.656	0.185	0.974
mixture density	0.6	0.759	0.524	0.972

Table 2.4: Estimated Parameter of the Mixture Model
(Lower Mode)

Parameter	True	Point Estimation	95%CI(L)	95%CI(H)
Item 1 Efficacy Type 1	0.3	0.426	0.026	0.959
Item 2 Efficacy Type 1	0.7	0.683	0.223	0.945
Item 1 Efficacy Type 2	0.2	0.255	0.195	0.308
Item 2 Efficacy Type 2	0.8	0.704	0.645	0.765
mixture density	0.6	0.110	0.012	0.245

Chapter 3

Model Identification

The previous chapter outlines how to obtain the posterior distribution for the parameters of the Learning Through Practice Model (LTP) by the Markov Chain Monte Carlo algorithm. Practically, point estimation is usually preferred to posterior distribution because the point estimation is computationally less expensive to use in an online service and intuitively more straightforward to interpret. It then begs the question when it is OK to reduce the posterior distribution of parameters to a point estimation? If the posterior parameter distribution is unimodal, then a proper point estimation makes sense. If the posterior distribution is multimodal, a point estimation is an inappropriate summary statistic of the posterior distribution. More specifically, if the point estimator is Maximum A Priori (MAP), $\hat{\theta}_{MAP}(x) = \arg \max_{\theta} f(x|\theta)g(\theta)$, unimodality is equivalent to the concept of identification in the frequentist view. Therefore, understand the identification of the LTP model is important for the practical usage.

This chapter is a preliminary study of the identification problem of the LTP model from the perspective of the generalized method of moments. It provides a necessary condition for the identification of the LTP model: If a globally optimal solution exists, the number of parameters is smaller than the number of first moments of the joint distribution of the

observed data and the Jacobian matrix has full column rank evaluated at the solution. This condition puts an upper limit of the model complexity with respect to the latent states. However, for reasonably simple latent state structure, this condition does not guarantee a unique identification. However, although a sufficient identification condition cannot be given to the general model, a sufficient identification condition is found for the Bayesian Knowledge Tracing model (BKT), the simplest possible structure for latent variables. The BKT model is identified as long as the practice sequence is longer than three periods, the efficacy is not zero, and the guess rate does not equal to the slip rate.

The chapter is organized as the following. The first subsection reviews the existing work of model identification for the HMM model family and motivates the perspective of generalized method of moments as a new alternative. The second subsection describes the first moments of the observed data, which turns out to be sufficient statistics. The third subsection derives the necessary identification condition for the LTP model. The fourth subsection derives the sufficient identification condition of the BKT model.

3.1 The Literature Review

Because the LTP model is an application of the hidden Markov model (HMM), if identification conditions for the HMM are established, identification conditions of the LTP model can be derived from it. Ephraim and Merhav (2002) provided a literature review of researches on model identification, from which emerges two promising perspectives. The first perspective of proof comes from Leroux's observation (1992) that the likelihood of HMM can be viewed as a finite mixture of the product densities. He proved that the HMM is identified under loose conditions if the observed response is drawn from Poisson distribution, Gaussian distribution with fixed variance, Exponential distribution, and negative Exponential distribution. The second perspective of proof comes from Ito et al(1992), who took an algebraic approach

to study the equivalence of different discrete Markov processes. Their approach requires the mapping between the latent variable and the observed variable to be a proper function, i.e. one latent state can only map to one value of the observed variable. Because observed responses of the LTP model follow the multinomial distribution, Leroux’s theorem does not apply. Because the mapping between the latent mastery and the observed response of the LTP model is not a proper function, Ito’s result does not apply. In short, both perspectives are not conducive to identification proof of the LTP model.

As for the literature on the Bayesian Knowledge Tracing model, the identification (or lack thereof) is an important topic. Beck and Chang (2007) first raised the issue of model identification after correctly observing that multiple parameter sets with very different interpretations of the learning process lead to an identical learning curve. They subsequently (incorrectly) concluded that the BKT model is not identified. This conclusion was later echoed by Pardos and Heffernan (2010a) and Van De Sande (2013). The “lack” of identifiability is one of the reasons for the literature to move toward individualized model parameters (Baker et al., 2008; Pardos and Heffernan, 2010b). However, the identification conditions of the individualization strategy are never formally proved. In fact, it is a counter-intuitive strategy because more, not fewer, parameters are introduced to solve the identifiability problem with the same dataset and essentially the same structural representation. Rai et al (2009) introduce the Dirichlet prior to confining the EM algorithm within a certain parameter space. This is a reasonable strategy if the prior is correctly specified, which is not easy to guarantee. In short, the identification of the BKT model is not properly addressed in the learning analytics literature.

This chapter attempts to study the identification of the LTP model from a new perspective by using the moments of the joint distribution of item assignments, observed responses, stop decisions, and effort decisions. This identification strategy is inspired by Blackwell and Koopmans’s work (1957) to prove the identification for a special hidden Markov model. In

addition, the perspective connects the identification of the LTP model to the literature on the identification of Generalized Method of Moments (Hansen, 1982).

3.2 Sufficient Statistics of the LTP Model

This section reparametrizes data generating process assumed by the LTP model as a multinomial distribution. The sample frequencies of the joint variable are sufficient statistics for the multinomial distribution. By drawing the analogy, it is can be proven that the sample frequencies of the joint distribution of item assignments, observed responses, stop decisions, and effort decisions are the sufficient statistics of the LTP model.

To review, \mathbf{A} , is the joint item assignments of the sequence. \mathbf{Y} is the joint responses. \mathbf{H} is the joint stop decisions. \mathbf{E} is the joint effort decisions. $\mathbf{a}, \mathbf{y}, \mathbf{h}, \mathbf{e}$ are the realized value. Because $\mathbf{Y}, \mathbf{A}, \mathbf{E}, \mathbf{H}$ are all discrete variables, the sample space of their joint distribution is countable. Define a mapping function $G(\mathbf{y}, \mathbf{a}, \mathbf{e}, \mathbf{h}) = \omega$ that projects each distinct combinations of $\mathbf{Y}, \mathbf{A}, \mathbf{E}, \mathbf{H}$ to a positive integer ω . The value of ω is arbitrary as long as each combinations are designated a distinct value.

For example, if only two periods are observed and only responses different in the two periods, the mapping function can be defined as $G(\mathbf{A}, \mathbf{Y}, \mathbf{E}, \mathbf{H}) = Y_1 + 2Y_2$ and $\omega \in \{0, 1, 2, 3\}$. Alternatively, if the sequence has two items ($j \in \{0, 1\}$), then the mapping function can be defined as $G(\mathbf{A}, \mathbf{Y}, \mathbf{E}, \mathbf{H}) = Y_1 + 2Y_2 + 4A_1 + 8A_2$ and $\omega \in \{X | 0 \leq X \leq 15, X \in \mathbb{Z}\}$.

Given the mapping function, define a new random variable Ω whose sample space is the collection of ω . It is obvious that Ω follows a multinomial distribution whose probability mass function is

$$P(\Omega = \omega) = P(\mathbf{A} = \mathbf{a}, \mathbf{Y} = \mathbf{y}, \mathbf{E} = \mathbf{e}, \mathbf{H} = \mathbf{h})$$

Let N_Ω be the cardinality of the sample space of Ω . Define $n_\omega = \sum_{i=1}^N I(G(\mathbf{y}^i, \mathbf{a}^i, \mathbf{e}^i, \mathbf{h}^i) = \omega)$ where i is the learner id and N is the number of learners.

Theorem 3.1. $\{n_1, \dots, n_{N_\Omega}\}$ are sufficient statistics of the joint distribution $(\mathbf{Y}, \mathbf{A}, \mathbf{E}, \mathbf{H})$

Proof. The total likelihood function is

$$L = \prod_{i=1}^N P(\mathbf{Y} = \mathbf{y}^i, \mathbf{A} = \mathbf{a}^i, \mathbf{E} = \mathbf{e}^i, \mathbf{H} = \mathbf{h}^i)$$

Apply the mapping function, the total likelihood is equivalent to

$$\begin{aligned} L &= \prod_{i=1}^N P(\Omega = G(\mathbf{y}^i, \mathbf{a}^i, \mathbf{e}^i, \mathbf{h}^i)) \\ &= \prod_{\omega=1}^{N_\Omega} P(\Omega = \omega)^{\sum_{i=1}^N I(G(\mathbf{y}^i, \mathbf{a}^i, \mathbf{e}^i, \mathbf{h}^i) = \omega)} \\ &= \prod_{\omega=1}^{N_\Omega} P(\Omega = \omega)^{n_\omega} \end{aligned}$$

Since the likelihood function is expressed as a multinomial distribution, it is easy to see that the sample frequencies of the joint distribution of item assignments, observed responses, stop decisions and effort decisions $(\{n_1, \dots, n_{N_\Omega}\})$ are sufficient statistics. \square

3.3 Identification of the Learning Through Practice Model

In the reparametrized the likelihood, $P(\Omega = \omega)$ is a function of the original model parameters. For example, let $G(\mathbf{A}, \mathbf{Y}, \mathbf{E}, \mathbf{H}) = Y_1 + 2Y_2$, then

$$\begin{aligned} P(\Omega = 2) &= P(Y_1 = 0, Y_2 = 1) \\ &= (1 - \pi)(1 - \ell)(1 - c^0)c^0 + (1 - \pi)\ell(1 - c^0)c^1 + \pi c^1(1 - c^1) \end{aligned}$$

Because all other combinations of item assignments, observed responses, stop decisions and effort decisions are a collection of elements in Ω , all other first order moments are linear combinations of $\{n_1, \dots, n_{N_\Omega}\}$. Therefore, Theorem 3.1 provides $N_\Omega - 1$ moment conditions that are nonlinear functions of the parameters. To prove the LTP model is uniquely identified is equivalent to prove that there exists a global unique solution θ^* to the system of nonlinear equations. Although the global uniqueness cannot be proved for the nonlinear equations with multiple variables, it is possible to prove that a unique local solution exists.

Theorem 3.2. *The parameters are locally identified only if the number of parameters is smaller than or equal to $N_\Omega - 1$ and the Jacobian matrix of the moment conditions evaluated at the local optimal solution has full column rank.*

Proof. If the number of parameters is larger than $N_\Omega - 1$, there are more variables than equations. The system has no unique solution.

If Jacobian matrix evaluated at the optimal solution does not full column rank, it means that for at least one such θ_j^* that $\frac{\partial P(\Omega=\omega, \theta^*)}{\partial \theta_j^*} = 0 \quad \forall \omega$. Therefore, θ_j has multiple roots and the solution is not unique. \square

3.3.1 The Practical Implication

Theorem 3.2 puts an upper limit on the maximum number of parameters based on the number of sufficient statistics rather than the number of observations. The number of observations affects the precision of the estimated parameter but not the identification of the parameter. Assume a data set captures three responses for each learner. The analyst tries to set up an LTP model without learner engagement. If there is only one item, there are seven moment conditions. A LTP model with $M_X = 2$ and $M_y = 2$ (BKT) can be identified because it has four free parameters. A LTP model with $M_X = 2$ and $M_y = 3$ can be identified because it has six free parameters. A LTP model with $M_x = 3$ and $M_y = 2$ cannot be identified because it

has eight free parameters. If there are two items, but the item sequence is fixed, the number of moment conditions is still seven. Only the BKT model can be identified because it has six free parameters in this case. In contrast, if the sequence order is not fixed for the two item, the number of moment conditions increase sharply to $63!$ This theorem is useful in practice because the analyst can design the sequence structure needed for identification.

Another practical implication is that the learning curve is not the sufficient statistics. It should not be the basis of statistical inference. For example, Beck and Chang (2007) give three model specifications that generate the same learning curve but different distributions of join responses. Have they looked at the correct sufficient statistics, they would not have concluded that the Bayesian Knowledge Tracing model is not identified. For another example, Murray et al (2013) are puzzled by an observation that each individual learning curve divided by practice sequence length is upward sloping but the aggregated learning curve is flat. Had they approached the problem from the perspective of this chapter, they would never look at the aggregate learning curve for the evidence of learning in the first place.

Besides its fundamental importance to identification, the sufficient statistics of the LTP model may also serve as a useful tool for diagnostics. Because the LTP model offers a great latitude in model specifications, a comparison of the predicted sufficient statistics generated by the fitted parameters with the sample sufficient statistics may be revealing about the adequacy of the specification. The author has observed that the BKT model often under predicts the proportion of learners who get all or most of the items wrong. Such pattern suggests that increasing the number of states in the latent mastery would better capture the tail behavior. However, it could be hard to visualize the sufficient statistics in a complex dataset because of its large number of moments, a statistical test of model fitness needs to be developed to perform formal diagnostics.

3.4 Identification of the Bayesian Knowledge Tracing Model

Although sufficient identification conditions cannot be obtained for the LTP model in general, they can be obtained for the Bayesian Knowledge Tracing (BKT) model by solving the system of nonlinear equations analytically.

Theorem 3.1 implies that the BKT model cannot be identified with a sequence of two responses because there are four parameters but only three moment conditions. When the sequence length is longer than or equal to three, there are more moment conditions than variables. Start by the simplest case of a sequence of length three.

Theorem 3.3. *The Bayesian Knowledge Tracing model is identified on practice sequences with length three if $\pi \neq 1$, $0 \leq \ell < 1$ and $c^0 \neq c^1$.*

Proof. Let $p_{ijk} = P(Y_1 = i, Y_2 = j, Y_3 = k)$, $p_{i,j} = P(Y_1 = i, Y_2 = j)$, and $p_i = P(Y_1 = i)$.

Excluding $p_{0,0,0}$, the rest seven moment conditions are:

$$\begin{aligned}
p_{111} &= (1 - \pi)(1 - \ell)(c^0)^3 + (1 - \pi)(1 - \ell)\ell(c^0)^2c^1 \\
&\quad + (1 - \pi)\ell c^0(c^1)^2 + \pi(c^1)^3 \\
p_{110} &= (1 - \pi)(1 - \ell)(c^0)^2(1 - c^0) + (1 - \pi)(1 - \ell)\ell(c^0)^2(1 - c^1) \\
&\quad + (1 - \pi)\ell c^0c^1(1 - c^1) + \pi(c^1)^2(1 - c^1) \\
p_{101} &= (1 - \pi)(1 - \ell)(c^0)^2(1 - c^0) + (1 - \pi)(1 - \ell)\ell c^0(1 - c^0)c^1 \\
&\quad + (1 - \pi)\ell c^0(1 - c^1)c^1 + \pi(c^1)^2(1 - c^1) \\
p_{011} &= (1 - \pi)(1 - \ell)(c^0)^2(1 - c^0) + (1 - \pi)(1 - \ell)\ell(1 - c^0)c^0c^1 \\
&\quad + (1 - \pi)\ell(1 - c^0)(c^1)^2 + \pi(c^1)^2(1 - c^1)
\end{aligned}$$

$$\begin{aligned}
p_{100} &= (1 - \pi)(1 - \ell)c^0(1 - c^0)^2 + (1 - \pi)(1 - \ell)\ell c^0(1 - c^0)(1 - c^1) \\
&\quad + (1 - \pi)\ell c^0(1 - c^1)^2 + \pi c^1(1 - c^1)^2 \\
p_{010} &= (1 - \pi)(1 - \ell)c^0(1 - c^0)^2 + (1 - \pi)(1 - \ell)\ell(1 - c^0)c^0(1 - c^1) \\
&\quad + (1 - \pi)\ell(1 - c^0)(1 - c^1)c^1 + \pi c^1(1 - c^1)^2 \\
p_{001} &= (1 - \pi)(1 - \ell)c^0(1 - c^0)^2 + (1 - \pi)(1 - \ell)\ell(1 - c^0)^2 c^1 \\
&\quad + (1 - \pi)\ell(1 - c^0)(1 - c^1)c^1 + \pi c^1(1 - c^1)^2
\end{aligned}$$

From these base moments, derive the following moments by marginalizing over nuisance periods. For example $p_{11} = p_{111} + p_{110}$.

$$\begin{aligned}
p_{11} &= (1 - \pi)(1 - \ell)(c^0)^2 + (1 - \pi)\ell c^0 c^1 + \pi(c^1)^2 \\
p_{01} &= (1 - \pi)(1 - \ell)(1 - c^0)c^0 + (1 - \pi)\ell(1 - c^0)c^1 + \pi(1 - c^1)c^1 \\
p_{10} &= (1 - \pi)(1 - \ell)c^0(1 - c^0) + (1 - \pi)\ell c^0(1 - c^1) + \pi c^1(1 - c^1) \\
p_1 &= (1 - \pi)c^0 + \pi c^1
\end{aligned}$$

With some algebra, it is easy to show that if $\pi \neq 1$, $0 \leq \ell < 1$ and $c^1 \neq c^0$,

$$\begin{aligned}
c^1 &= \frac{p_{101} - p_{011}}{p_{01} - p_{10}} \\
c^0 &= \frac{p_{110} - p_{101}}{p_{110} - p_{101} + p_{001} - p_{010}}
\end{aligned}$$

Plug c^1 and c^0 into p_1 to solve for π

$$\pi = \frac{p_{10} + p_{01} - \frac{p_{110} - p_{101}}{p_{110} - p_{101} + p_{001} - p_{010}}}{\frac{p_{101} - p_{011}}{p_{01} - p_{10}} - \frac{p_{110} - p_{101}}{p_{110} - p_{101} + p_{001} - p_{010}}}$$

Plug c^1 , c^0 and π into any of the equations above to solve for ℓ . This proof chooses $p_{01} - p_{10}$.

$$\ell = \frac{p_{01} - p_{10}}{(1 - \pi)(c^1 - c^0)} = \frac{p_{01} - p_{10}}{\frac{p_{101} - p_{011}}{p_{01} - p_{10}} - (p_{11} + p_{10})}$$

Now that one solution to the system is found, it is necessary to prove that it is the only solution. c^1 and c^0 are both solutions to a linear equation with one unknown, therefore they are unique. π is also the unique solution to a linear equation with one unknown when c^1 and c^0 are plugged in. When c^1 , c^0 and π are solved, ℓ is the unique solution to a linear equation in any equations. In sum, the solution is unique although the representation of the solution is not.

Now consider the special cases when the model is not identified.

If $\pi = 1$ but $0 < \ell < 1$, ℓ and c^0 are not identified because they are never observed.

If $\ell = 1$ but $0 < \pi < 1$, π and c^0 are not identified because p_1 is a linear equation with two unknowns.

If $c^1 = c^0$, π and ℓ are not uniquely identified because the latent variable collapses to one state. □

With Theorem 3.3, it is possible to prove

Theorem 3.4. *The Bayesian Knowledge Tracing model is identified if at least three periods of response are observed, $\pi \neq 1$, $0 \leq \ell < 1$ and $c^{1,0} \neq c^{1,1}$.*

Proof. The equivalent representation of Theorem 3.4 is that the BKT model is identified if the three-response sequence BKT model is identified.

Assume the BKT model based on sequences of length three is identified, but the model on sequences of length T is not identified. Let m ($m \geq 2$) be the size of the observation equivalent parameter sets. The parameters sets are denoted as $\Theta_1, \dots, \Theta_m$. Because that the parameter space is the same for the BKT model on sequences with length three and that with length T , $\Theta_1, \dots, \Theta_m$ also generates the observation for the BKT model on sequences with length three. However, it is uniquely identified, therefore $\Theta_1 = \dots = \Theta_m = \Theta$, and the BKT model on sequences with length T is identified. □

3.4.1 Revisit Beck&Chang

A revisit of the example in Beck and Chang (2007) illustrates the argument in this section. The parameter sets are listed in Table 3.1:

Table 3.1: BKT Parameters in Beck&Chang (2007)

model	pi	l	g	s
Knowledge	0.56	0.1	0.00	0.05
Guess	0.36	0.1	0.30	0.05
Read Tutor	0.01	0.1	0.53	0.05

Figure 3.1 plots the learning curves. The x-axis is the sequence position while the y-axis is the average success rate at that position ($P(Y_t = 1)$). It confirms Beck and Chang’s observation that three parameter sets generate essentially the same learning curve.

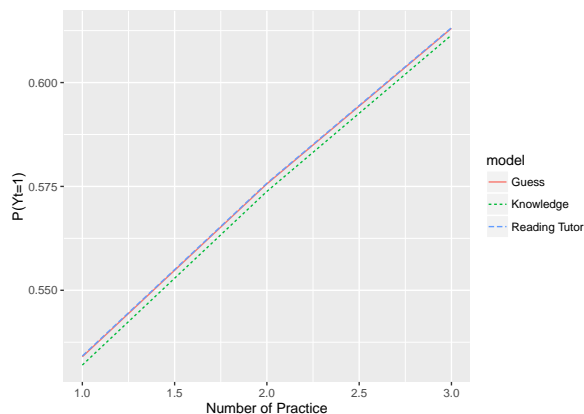


Figure 3.1: The Learning Curves (T=3)

Figure 3.2 shows the sufficient statistics generated by three parameter sets. The x-axis is the joint distribution of the response ($Y_1 = i, Y_2 = j, Y_3 = k$ as i, j, k), the y-axis is the sample proportions ($P(Y_1 = i, Y_2 = j, Y_3 = k)$). Figure 3.2 clearly shows that three models generate

distinct sufficient statistics and thus have distinct estimated parameters.

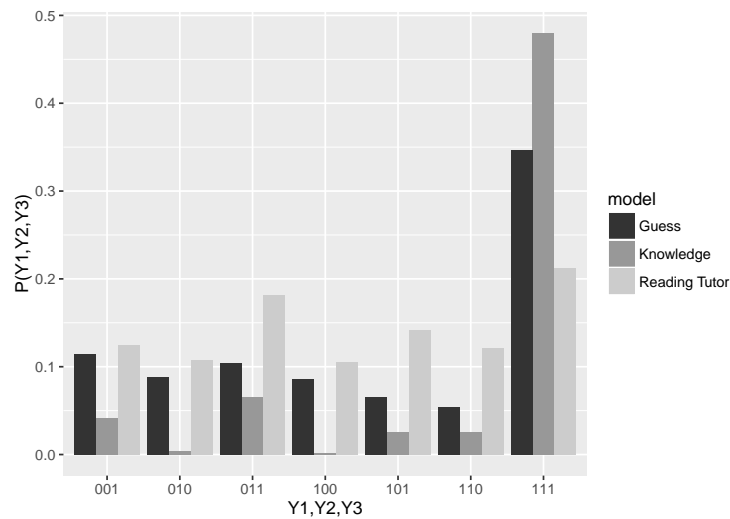


Figure 3.2: The Sufficient Statistics (T=3)

Table 3.2 shows the fitted pedagogical efficacy from simulation by the Monte Carlo Markov Chain algorithm introduced in the next chapter. The MCMC algorithm initiates four chains from random starting points so as to show that convergence does not depend on the initial guess. The tables report the mean and the 95% credible interval of the posterior parameter distribution.

The “Knowledge” model and the “Guess” model converge to the true value when the sequence length is only three, disproving Beck&Chang’s claim that the BKT model cannot distinguish between the two. The “Read Tutor” model fails to converge until sequences of length five. Because the initial mastery probability is close to zero, the realized number of observations who have mastery at $t = 1$ can be far from the expected value in one simulation due to the large variance. It results in slow bayesian learning when the degree of freedom is small because of the short sequence length.

Table 3.2: Estimated Efficacy

Model	Sequence length	True	Mean	95% CI(L)	95% CI(H)
Guess	3	0.1	0.094	0.061	0.129
Knowledge	3	0.1	0.099	0.079	0.119
Read Tutor	3	0.1	0.405	0.241	0.583
Read Tutor	4	0.1	0.249	0.098	0.577
Read Tutor	5	0.1	0.117	0.080	0.161

Table 3.3,3.4,3.5 show the estimated parameters for the initial mastery probability (π), the slip rate(g) and the slip rate (s). The parameters of the “Knowledge” model and the “Guess” model are well identified with a sequence of length three, but the “Read Tutor” model still does not converge to the true value with a sequence of length five.

Table 3.3: Estimated Initial Mastery

Model	Sequence length	True	Mean	95% CI(L)	95% CI(H)
Guess	3	0.36	0.372	0.314	0.439
Knowledge	3	0.56	0.521	0.494	0.548
Read Tutor	3	0.01	0.360	0.144	0.568
Read Tutor	4	0.01	0.227	0.024	0.524
Read Tutor	5	0.01	0.079	0.007	0.205

Table 3.4: Estimated Guess Rate

Model	Sequence length	True	Mean	95% CI(L)	95% CI(H)
Guess	3	0.30	0.309	0.269	0.344
Knowledge	3	0.00	0.005	0.001	0.012
Read Tutor	3	0.53	0.418	0.333	0.490
Read Tutor	4	0.53	0.493	0.417	0.543
Read Tutor	5	0.53	0.500	0.452	0.531

Table 3.5: Estimated Slip Rate

Model	Sequence length	True	Mean	95% CI(L)	95% CI(H)
Guess	3	0.05	0.063	0.028	0.102
Knowledge	3	0.05	0.047	0.038	0.059
Read Tutor	3	0.05	0.315	0.254	0.375
Read Tutor	4	0.05	0.261	0.140	0.363
Read Tutor	5	0.05	0.109	0.021	0.192

Chapter 4

Dynamic Selection Bias of Sample Attrition

4.1 Introduction

One way to improve the intelligent tutoring system is to identify knowledge components that student are not learning even though they are practicing. An intuitive plan is to examine the slope of the average success rate at each practice opportunity (the observed learning curve) and identify those with a flat slope as not pedagogically effective (Ritter et al., 2007).

This may not be a good plan when learners can stop practicing at different times. Even when the true pedagogical efficacy is zero, it is still possible to generate positive or negative “learning” in the observed learning curve. For example, if a learner is forced to stop after accumulating X number of wrong answers, learners with low mastery drop out earlier than learners with high mastery because they are more likely to make errors. Consequently, the proportion of learners with high mastery is inflated. The observed learning curve slope upward and spurious positive learning is generated. Similarly, if a learner is forced to stop after accumulating X number of right answers, learners with high mastery exit earlier than

learners with low mastery. As a consequence, the observed learning curve slopes downward. (Murray et al., 2013)

Pelánek et al (2016) point out that the learning curve is not a reliable metric of pedagogical efficacy if learners have differential attrition rates. Without rigorous proof, they proceed to argue that pedagogical efficacy estimated by the Bayesian Knowledge Tracing (BKT) model is inconsistent as a result of selective attrition. This chapter shows that differential sample attrition is not a sufficient condition of inconsistent parameter estimation for the BKT model. The mechanism of the stop decision matters. If learners stop by rule, the BKT model can still consistently estimate the pedagogical efficacy. If learners stop by choice, the BKT model is likely to produce inconsistent parameter estimation and the learning through practice (LTP) model can consistently estimate the pedagogical efficacy by correcting the selective sample attrition with a hazard model.

This chapter is organized as the following: The first section characterizes the relationship between hazard rate dependence structure and selection bias. The second section briefly reviews how the LTP model correct for the selection bias, supported by a simulation exercise. The last section applies the model to a quiz dataset of math learning in third grade. The selection bias accounts for a large share of the “learning gain” as defined by the slope of the observed learning curve.

4.2 Sample Attrition and Dynamic Selection Bias

4.2.1 Dynamic Selection Bias

Dynamic selection bias is a well-studied inference problem in econometrics (Heckman and Singer, 1984) and program evaluation (Ham and LaLonde, 1996). The population under study is categorized into latent types, with a different set of parameters for each type. If

the population has a changing composition in latent types, a model would be biased with an implicit assumption that the composition of latent types is static. Such bias is called as dynamic selection bias. Pelánek et al (2016) make a similar argument for why sample attrition is a cause for concern in the Bayesian Knowledge Tracing (BKT) model. In the context of the BKT model, the mastery is the latent type. Assume learners with mastery drop out earlier than those without, the differential sample attrition leads to a lower proportion of learners with mastery over time. Since the BKT model assumes either no sample attrition or equal sample attrition, the change in the proportion of learners with mastery due to dynamic selection is attributed to the learning process, the two processes being observational equivalent. Consequently, the pedagogical efficacy is under-estimated.

Although Pelánek et al are correct in reasoning that the change in the mastery trajectory due to selective attrition is the source of the bias of the BKT model, they fail to realize that it is only a necessary condition, but not sufficient. They fail to distinguish two types of selective sample attrition: response-dependent sample attrition and mastery-dependent sample attrition. Only the latter type leads to the dynamic selection bias in the parameter learning. Their mistake, similar to that made by Beck and Chang, is to base their inference of parameter estimation on the shape of the observed learning curve. The observed learning curve and the parameter learning, as well as the posterior inference on learner mastery, condition on different information sets. Each point on the observed learning curve conditions only on continuing at the previous practice opportunity. In contrast, the parameter learning conditions on all observed response in addition to survival at previous practice opportunities. If the sample attrition is response-dependent, it does not affect parameter learning because the inference is already conditional on the response. Further conditioning on a function of responses does not bring in any new information. However, the observed learning curve is changed because conditioning on a function of responses does change the posterior inference of the latent mastery and thus the success rate. In short, although both types of sample attrition alter the observed learning curve, only the mastery-dependent sample attrition can

change the posterior distribution of latent mastery conditional on the observed response, thus change the estimation of pedagogical efficacies.

In sum, although the dynamic selection process explains the intuition behind the parameter bias caused by selective sample attrition, the composition change of latent mastery unconditional on the observed response is only a necessary condition of estimation bias in the BKT model. The next subsection develops a formal mathematical proof for the above intuitive reasoning.

4.2.2 Dynamic Selection bias and Dependence Structure

Let t be the practice sequence id, Y_t be the observed response at sequence t , X_t be the latent mastery, and H_t be the stop decision where $H_t = 1$ denotes attrition at t^{th} attempt. For simplicity, let $\mathbf{Y}_{\mathbf{t}_1, \mathbf{t}_2} = \{Y_{t_1}, \dots, Y_{t_2}\}$. When referencing the whole practice sequence ($t = T$), denote $\mathbf{Y}_{\mathbf{1}, \mathbf{t}} = \mathbf{Y}$ for short.

Theorem 4.1. *The pedagogical efficacy is consistently estimated by the Bayesian Knowledge Tracing Model only if*

$$P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2} | \mathbf{Y}, H_{t-1} = 0) = P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2} | \mathbf{Y}) \quad \forall \quad 1 < t_1 < t_2 \leq T$$

Proof. If the true parameter set Θ is stationary in the iterative estimation procedure of the BKT model, which means the estimated parameter set of iteration $s + 1$ is the same as that of iteration s , the BKT model converges to the true parameter. Take all parameters other than the pedagogical efficacy from iteration s as given, update the pedagogical efficacy parameter

($\ell^{mn} = P(X_t = n | X_{t-1} = m)$) for iteration $s + 1$:

$$\begin{aligned}\hat{\ell}_{BKT}^{mn} &= \frac{\sum_{t=2}^T \sum_{i=1}^N I(X_t^i = n, X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)}{\sum_{t=2}^T \sum_{i=1}^N I(X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)} \\ &= \frac{\sum_{t=2}^T \frac{\sum_{i=1}^N I(X_t^i = n, X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)}{N}}{\sum_{t=2}^T \frac{\sum_{i=1}^N I(X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)}{N}}\end{aligned}$$

By law of large number,

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N I(X_t^i = n, X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)}{N} &\rightarrow P(X_t = n, X_{t-1} = m | \mathbf{Y}) \\ \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N I(X_{t-1}^i = m | \Theta_s, \mathbf{Y}^i)}{N} &\rightarrow P(X_{t-1} = m | \Theta_s, \mathbf{Y}) \\ \lim_{N \rightarrow \infty} \hat{\ell}_{s+1}^{mn} &\rightarrow \frac{\sum_{t=2}^T P(X_t = n, X_{t-1} = m | \mathbf{Y})}{\sum_{t=2}^T P(X_{t-1} = m | \Theta_s, \mathbf{Y})} \\ &= \frac{\sum_{t=2}^T \ell^{mn} P(X_{t-1} = m | \Theta_s, \mathbf{Y})}{\sum_{t=2}^T P(X_{t-1} = m | \Theta_s, \mathbf{Y})} \\ &= \ell^{mn}\end{aligned}$$

□

In order for the BKT model to be consistent, Theorem 4.1 claims that the stop decision should be independent of the latent mastery conditional on the response. Therefore, the sample attrition essentially only a function of responses. This observation is captured by the two following lemmas

Lemma 4.1. *If the stop decision depends only on the observed response, $P(H_t = 1) = f(Y_1, \dots, Y_t)$, the Bayesian Knowledge Tracing model consistently estimates the pedagogical efficacy.*

Proof. Notice that

$$\begin{aligned}
P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2} | \mathbf{Y}, H_{t-1} = 0) &= \frac{P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2}, \mathbf{Y}, H_{t-1} = 0)}{P(\mathbf{Y}, H_{t-1} = 0)} \\
&= \frac{P(H_{t-1} = 0 | \mathbf{Y}) P(\mathbf{Y} | \mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2}) P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2})}{\sum_{X_{t_1}} \cdots \sum_{X_{t_2}} P(H_{t-1} = 0 | \mathbf{Y}) P(\mathbf{Y} | \mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2}) P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2})} \\
&= \frac{P(\mathbf{Y} | \mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2}) P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2})}{\sum_{X_{t_1}} \cdots \sum_{X_{t_2}} P(\mathbf{Y} | \mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2}) P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2})} \\
&= P(\mathbf{X}_{\mathbf{t}_1, \mathbf{t}_2} | \mathbf{Y})
\end{aligned}$$

The results follow by Theorem 4.1. □

Lemma 4.2. *If the stop decision depends on the latent mastery, $P(H_t = 1) = f(Y_1, \dots, Y_t, X_1, \dots, X_t)$, the Bayesian Knowledge Tracing model consistently estimates the pedagogical efficacy only if $f(Y_1, \dots, Y_t, X_1 = x_1^1, \dots, X_t = x_t^1) = f(Y_1, \dots, Y_t, X_1 = x_1^2, \dots, X_t = x_t^2) \quad \forall x_1^1, x_1^2, \dots, x_t^1, x_t^2$.*

Proof. Let η_t^k denote conditional hazard rate $P(H_t = 1 | X_t = k)$. The the posterior distribution of the marginal distribution is

$$\begin{aligned}
P(X_t = m | \mathbf{Y}, H_{t-1} = 0) &= \frac{\sum_{k=1}^m P(X_t = m | X_{t-1} = k) (1 - \eta_{t-1}^k) P(X_{t-1} = k | \mathbf{Y})}{\sum_{n=1}^{M_x} [\sum_{l=1}^n P(X_t = n | X_{t-1} = l) (1 - \eta_{t-1}^l) P(X_{t-1} = l | \mathbf{Y})]} \\
P(X_t = m | \mathbf{Y}) &= \frac{\sum_{k=1}^m P(X_t = m | X_{t-1} = k) P(X_{t-1} = k | \mathbf{Y})}{\sum_{n=1}^{M_x} [\sum_{l=1}^n P(X_t = n | X_{t-1} = l) P(X_{t-1} = l | \mathbf{Y})]}
\end{aligned}$$

Let $P(X_t = m | X_{t-1} = n)$ be ℓ^{nm} , $P(X_{t-1} = k | \mathbf{Y}) = \pi_k$, $P(H_{t-1} = 0 | X_{t-1} = k) = h^k$.

$$P(X_t = m | \mathbf{Y}, H_{t-1} = 0) = P(X_t = m | \mathbf{Y}) \rightarrow \sum_{n=1}^{M_x} \sum_{l=1}^n \sum_{k=1}^m \ell^{km} \ell^{ln} \pi_k \pi_l (h_l - h_k) = 0$$

If $h_l = h_k \quad \forall l, k$, it is obvious that the equality stands. If the equality stands, but $h_l \neq h_k$ for some l, k , then it must be true that $\ell^{km} \ell^{ln} = \ell^{lm} \ell^{kn}$. Since no conditions are imposed on the pedagogical efficacy, it must not be the case. Therefore, the sufficient and necessary condition is that the conditional hazard rates are equal across states. □

The contrapositive statement of lemma 4.2 essentially states that BKT model inconsistently estimates the pedagogical efficacy if the stop decision depends on the latent mastery in a meaningful way.

4.2.3 Stop-by-Rule and Stop-by-Choice

Lemma 4.1 and Lemma 4.2 highlight the importance of dependence structure in determining the existence of dynamic selection bias in the BKT model. The response-dependent stop decision and the mastery dependent stop decision are quite abstract and far away from practical application. This subsection links the dependence structure to learner behavior.

There are two types of stop decision: Stop-by-rule and Stop-by-choice. Stop-by-rule refers to the exit from practice due to a system rule, which is usually the consequence of system design. Conditional on the observed response, the stop decision is deterministic. The “X-strike” rule is the most common example of the stop-by-rule behavior. For instance, in the old version of Duolingo, a learner is forced to quit the level after three (at most four) errors. The case study in later of the chapter uses the X-strike rule in both ways. If a learner accumulates three or four correct answers, she is forced to quit; if a learner accumulates two or three incorrect answers, she is forced to quit. Stop-by-choice refers to the exit from practice due to learner’s own initiative when she could have continued to practice, which is usually the consequence of learner’s non-cognitive skill. For example, if a learner has mastery, continuing the practice is boring; if a learner has no mastery, continuing the practice is frustrating. Both boredom and frustration lead to an exit from practice sequence, at different rates (Baker et al., 2010).

In the rest of the chapter, the response-dependent stop decision is equivalent to stop-by-rule while the mastery-dependent stop decision is equivalent to stop-by-choice. This generalization of the stop decision mechanism has exceptions. The proficiency rule is one such exception. In the new version of Duolingo, a learner is forced to quit the level after accumulating enough correct answers to reach the proficiency threshold. The cognitive tutor developed by Carnegie

Learning LLC has a similar “optimal-stop” decision system (Murray et al., 2013) where the practice is terminated once the learner reaches 90% mastery likelihood in the posterior inference. The stop decision depends on latent mastery but it is implemented by a system rule based on responses. However, such exceptions are rare because only very few education service providers have the capability to deliver mastery inference in real time. In most cases, stop-by-rule is response-dependent stop decision and stop-by-choice is mastery dependent stop decision.

4.2.4 Sign of Dynamic Selection Bias in Stop-by-Choice

Lemma 4.1 and Lemma 4.2 establishes that only stop-by-choice, the mastery-dependent stop decision, results in inconsistent parameter estimation. This subsection further studies the properties of the dynamic selection bias. Although a rigorous proof is not established, an educated guess is made, supported by simulation evidence, on the sign of the selection bias in the stop-by-choice decision system.

If learners with mastery have a lower attrition rate than those without, the increase in the proportion of learners with mastery stem from differential attrition is falsely attributed to the pedagogical efficacy of the practices by the BKT model. Therefore, the estimated pedagogical efficacy is biased upwards. Such reasoning leads to Proposition 4.1.

Proposition 4.1. If the data generating process is $P(H_t = 1|H_{t-1} = 0) = f(X_t)$ and $h_t^{X=1} < h_t^{X=0}$, the BKT model over-estimates the pedagogical efficacy.

In contrast, if learners with mastery have a higher attrition rate than those without, the decrease in the proportion of learners with mastery stem from differential attrition is also falsely attributed to the pedagogical efficacy of the practices by the BKT model. Therefore, the estimated pedagogical efficacy is biased downwards. Such reasoning leads to Proposition 4.2.

Proposition 4.2. If the data generating process is $P(H_t = 1|H_{t-1} = 0) = f(X_t)$ and

$h_t^{X=1} > h_t^{X=0}$, the BKT model under-estimates the pedagogical efficacy.

However, Proposition 4.1 and 4.2 are built on the premise that the selective sample attrition is driven only by stop-by-choice. If the selective sample attrition is driven by both stop-by-rule and stop-by-choice, the learning through practice (LTP) model that considers only stop-by-choice is also biased. In a mixture stop decision system, neither the BKT model nor the LTP model estimates the pedagogical efficacy consistently, but they may serve as bounds on the true pedagogical efficacy. The key assumption for the BKT-LTP bound is that response-dependent and mastery dependent sample attrition biased in the same direction. If the response-dependent sample attrition favors learners with lower mastery, the mastery-dependent sample attrition needs to favor learners with lower mastery as well. If the assumption stands, the observed difference in hazard rates reflect the data generating process and the BKT-LTP bound can be established, which leads to the following proposition:

Proposition 4.3. If the selective sample attrition is both mastery-dependent and response-dependent, if both processes have the same direction of bias of the attrition rate, the pedagogical efficacies estimated by the BKT model and the LTP model bound the true pedagogical efficacy. When $P(H_t = 1|Y = 1) > P(H_t = 1|Y = 0)$, the BKT estimator is the lower bound and the LTP estimator is the upper bound. When $P(H_t = 1|Y = 1) < P(H_t = 1|Y = 0)$, the BKT estimator is the upper bound the LTP estimator is the lower bound.

In reality, the mixture stop decision process is likely to be true data generating process for the intelligent tutoring systems. Most tutoring systems have some kind of stop-by-rule decision rule so that learners are not trapped in practices they don't need when they have mastery or they don't want when they have no mastery. However, learners can quit by their own initiative by just shutting down the program. Proposition 3 argues that both BKT model and LTP model needs to be estimated in order to better understand the true pedagogical efficacy

4.3 Correct Dynamic Selection Bias

The previous section discusses the condition of inconsistent efficacy estimation as a result of selective sample attrition and the property of such inconsistency. This section briefly reviews how the learning through practice (LTP) model introduced in Chapter One corrects the dynamic selection bias. This section demonstrates the bias correction procedure with a simulation study of stop-by-rule and stop-by-choice decision process.

4.3.1 The LTP Model

The LTP model corrects the selection bias by modeling sample attrition explicitly. The dependence structure has two key components, the dependent factor (observed response or latent mastery) and the memory lag of dependence (concurrent, first order Markov, etc). The LTP model assumes the stop decision only depends on the concurrent data, thus limit the discussion to only the dependent factor. For simplicity, this chapter only considers one learner type. Thus the learner type is omitted from the following discussion. Also, the effort decision is omitted.

For the response-dependent stop decision, the join likelihood of $\{X_t, Y_t, H_t\}$ conditional on the learner continues at previous practice opportunity ($H_{t-1} = 0$) is

$$P(X_t, Y_t, H_t | H_{t-1} = 0, \mathbf{Y}_{1,t-1}) = P(H_t | H_{t-1} = 0, Y_t) P(Y_t | X_t) P(X_t | H_{t-1} = 0, \mathbf{Y}_{1,t-1})$$

Similarly, for the mastery-dependent stop decision, the conditional joint likelihood of $\{X_t, Y_t, H_t\}$ is

$$P(X_t, Y_t, H_t | H_{t-1} = 0, \mathbf{Y}_{1,t-1}) = P(H_t | H_{t-1} = 0, X_t) P(Y_t | X_t) P(X_t | H_{t-1} = 0, \mathbf{Y}_{1,t-1})$$

In the statistics literature, $P(H_t = 1 | H_{t-1} = 0)$ is called the hazard rate, meaning the probability to exit at this period given surviving the last period. If the functional form of the conditional hazard rate is specified, the conditional likelihood can be computed. The LTP model offers both parametric and non-parametric specification.

The parametric model is the discrete time version of the proportional hazard model. For example, if the hazard function depends on the latent mastery, it is specified as

$$P(H_t = 1 | X_t = k, H_{t-1} = 0) = \lambda_k e^{\beta_k t}$$

λ is the baseline hazard rate. β is the proportional growth of hazard rate over time. λ must be positive but β can be both positive and negative.

The non-parametric model imposes no structure on the hazard function and assumes the conditional hazard rates are drawn from independent Bernoulli distributions with different means. For example, if the exit function depends on the latent mastery and there is one learner type, the hazard rate is specified as

$$P(H_t = 1 | X_t = k, H_{t-1} = 0) = h_t^k$$

If the true data generating process is a proportional hazard model, both specifications are consistent but the parametric specification is more efficient because it only needs to estimate $2M_X$ parameters while the non-parametric specification needs to estimate $2M_X T$ parameters. However, if the data generating process is not a proportional hazard model, the parametric specification is not consistent while the non-parametric specification is.

4.3.2 Simulation

This subsection supports the previous discussion by simulations. It first shows the consistency of the BKT estimator and the inconsistency of the LTP estimator if the stop decision is stop-by-rule. It then demonstrates the inconsistency of the BKT estimator and the consistency of the LTP estimator if the stop decision is stop-by-choice. The values of all simulation parameters are in Appendix B.3.

Stop-by-rule (response-dependent) sample attrition is illustrated by the “two-strike” rule: The learners are forced to stop if they accumulate two errors. The learners only stop practicing when they make a mistake (except for the first period) and never quit when they answer correctly. The conditional hazard rates are listed in Table 4.1.

Table 4.1: Hazard Rates Conditional on Observed Response in Stop-by-Rule

Practice Sequence	Hazard Rate($Y=0$)	Hazard Rate($Y=1$)
1	0.000	0
2	0.714	0
3	0.686	0
4	0.565	0

Because the hazard rate curve conditioning on the wrong response is not monotonic in time, the LTP model uses the nonparametric specification of the hazard function. Table 4.2 shows the point estimation of the pedagogical efficacy obtained from the BKT model and the LTP model. The BKT estimator produces a point estimation very close to the true value of pedagogical efficacy, in contrast to the substantial downward bias of the LTP estimator.

Table 4.2: Estimated Pedagogical Efficacy of Stop-by-Rule
Sample Attrition

Model	Estimates
True	0.300
BKT	0.293
LTP	0.199

To keep the bias direction the same as the previous stop-by-rule example, the simulation assumes that learners with mastery have lower hazard rate than those without mastery. The hazard rates are generated by a proportional hazard function. The learners with mastery has lower baseline hazard rate ($\lambda_0 = 0.5$ Vs. $\lambda_1 = 0.2$) but higher growth rate ($\beta_0 = \log(1.1)$ V.S. $\beta_1 = \log(1.2)$). However, at the end of the practice sequence, learners with mastery are still less likely to stop than those without mastery. The two hazard rate curves are listed in Table 4.3.

Table 4.3: Conditional Hazard Rate of Stop-by-Choice
Sample Attrition

Practice Sequence	Hazard Rate(X=0)	Hazard Rate(X=1)
1	0.500	0.200
2	0.550	0.240
3	0.605	0.288
4	0.666	0.346

Table 4.4 shows the point estimation of the pedagogical efficacy obtained from the BKT model and the LTP model. The estimated pedagogical efficacy of the LTP model is close to

the true value, while the BKT model suffers from an upward bias.

Table 4.4: Estimated Pedagogical Efficacy of Stop-by-Choice Sample Attrition

Model	Estimates
True	0.300
BKT	0.394
LTP	0.283

Figure 4.1 visualizes the posterior distribution of the estimated pedagogical efficacy by the BKT and the LTP model in the stop-by-rule and the stop-by-choice scenario.

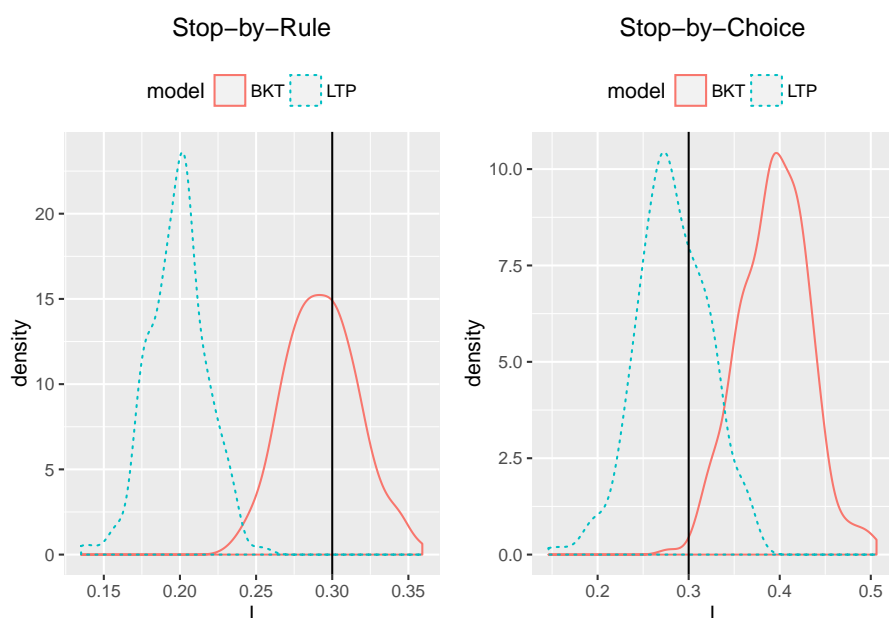


Figure 4.1: Posterior Distribution of Pedagogical Efficacy:Simulation

4.4 Case Study

4.4.1 The Learning Environment

The data were collected on a Chinese online learning platform from December 2015 to January 2016. The learning product embeds the math practices in a turn-based role-playing game. Figure 4.2 shows a screenshot of the product interface. The learning game is supplemental learning material that students practice on their own initiative in addition to the homework. The target learners are K-6, mainly from first grade to third grade. The demographics of the learner population are unknown beyond the grade distribution.



Figure 4.2: The Interface of the Learning Game

To complete a level in the game, a learner must defeat the enemy by doing practice questions. Each turn, the learner is challenged with a practice problem. If she answers correctly, the health point of the AI avatar is randomly reduced. If she answers incorrectly, the health point of the learner avatar is randomly reduced. If the health point of either avatar drops to zero, the practice session stops. In expectation, the learner fails the level if he accumulates

two or three errors and clears the level if he scores three or four hits. At any time, the learner is free to quit before she clears or fails the level. There is no punishment for early exit and the learner can try the same level as often as she wants.

For each level, the algorithm only recommends practice questions on one knowledge point. For example, if the knowledge point is “reducing fractions”, a question may be “ $5:8=x:32$ ” (as in Figure 4.2) or “ $3:7=9:x$ ”. Because questions are highly similar by design, the parameters of items belonging to the same knowledge point are considered identical across items. This chapter chooses two representative knowledge points out of more than 200 candidates: two-digit multiplication (grade 2) and long division (grade 3). The two items represent “slow-learning” and “fast-learning” judged by the observed learning curve, which is plotted in Figure 4.3.

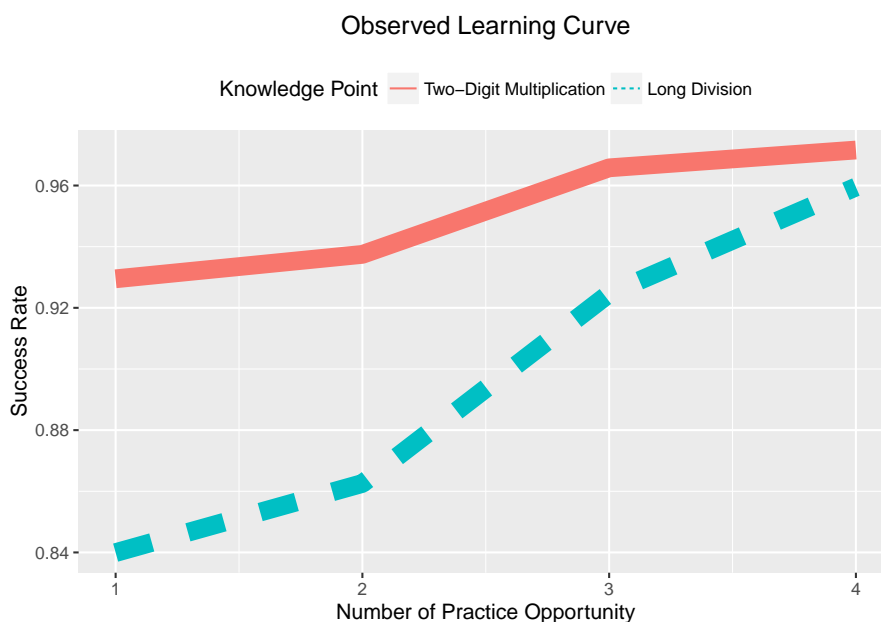


Figure 4.3: Observed Learning Curves of Knowledge Points

4.4.2 Data Cleaning Procedure

This chapter collected more than 68 million exercise logs. Only serious learners, defined as those have more than 50 log entries, are retained. Serious learners are only 20% of the total learners but they generate 62% of the total logs (about 42 million).

The recommendation algorithm changes knowledge points between different practice sequences so that the learner is not bored. Therefore, there are two sequence ranks one can compute. The global sequence rank ignores the interval and continues the practice rank count over the whole sample period. The local sequence rank only counts the practice rank within each sequence. Neither sequence ranking is perfect. The global sequence rank completely ignores the dropout; while the local sequence rank violates the assumption of homogeneous initial mastery. For the purpose of this chapter, learner heterogeneity is a lesser evil so the local sequence rank is chosen. If the learner has multiple practice sequences, choose the first one by calendar time.

The two knowledge points all have more than 200,000 practice logs. For two-digit multiplication, 95% of the practice sequences have a life span smaller than 4 periods, 80% for the long division. Because data are quite sparse past the fifth practice, the max practice sequence is 4.

For the remaining valid data, take 40% random sample of training data and 20% random sample as out sample forecast. The analysis limits the size because the adaptive rejection sampling algorithm has trouble fitting a large dataset.

4.4.3 Empirical Hazard Rate Curve

The hazard rate curve is plotted in Figure 4.4. The X-axis is the practice sequence number, the Y-axis is the hazard rate conditional on the response ($P(H_t = 1 | H_{t-1} = 0, Y_t = k)$). The solid line represents the correct answer while the dotted line represents the incorrect answer. 95% confidence interval is plotted as the error bar. Both knowledge points exhibit differential

conditional hazard rates that incorrect answer is correlated with a higher hazard rate. The difference between two conditional learning curves is smaller for the two-digit multiplication than the long division.

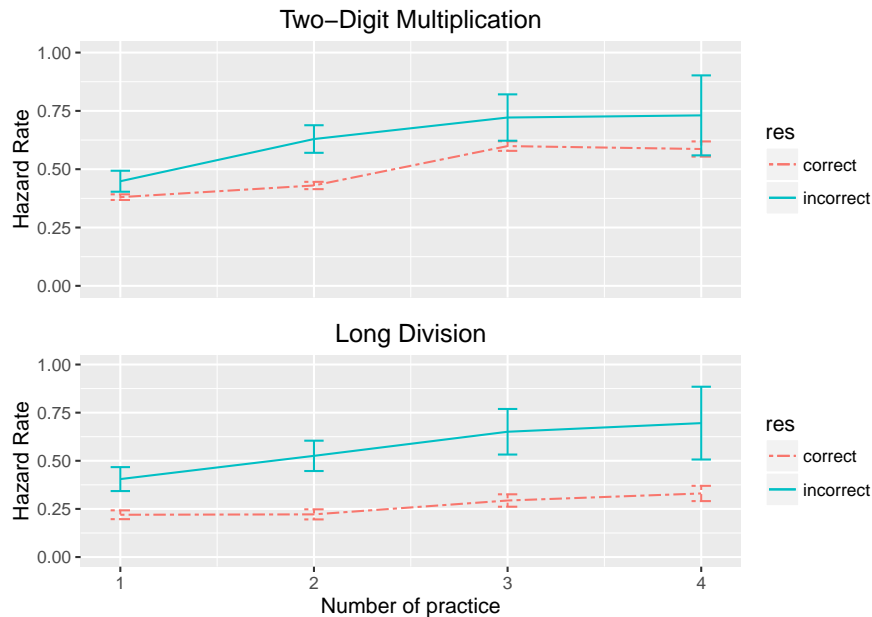


Figure 4.4: Empirical Hazard Rates of Different Knowledge Points

Because the learning game uses the X-strike rule, part of the stop decision is stop-by-rule (response dependent sample attrition). The hazard rate increases sharply in the third practice and flattens on the fourth practice is partial evidence of the X-strike rule. However, not all the stop decisions are driven by the X-strike rule. Otherwise, the hazard rate conditional on a correct response should be zero, as well as the hazard rate of the first period. The slow increase of the hazard rate curve conditional on the correct response is partial evidence of stop-by-choice (mastery dependent sample attrition). Therefore, it is reasonable to assume that the stop decision is a hybrid of stop-by-rule and stop-by-choice. That said, Proposition 4.3 argues that the BKT estimator and the LTP estimator bounds the true pedagogical efficacy.

Both parametric and nonparametric specification are fitted for the LTP model. Figure 4.5

shows the fitness of hazard rates by the parametric and the nonparametric specification. The solid line is the fitted hazard curve while the dotted line is the empirical hazard rate curve. The 95% credible interval derived from the posterior parameters are plotted as the error bar. Overall, both parametric specification and nonparametric specification fit the empirical hazard rate curve reasonably well. The parametric specification overfits the hazard rate at 4th sequence position for the two-digit multiplication.

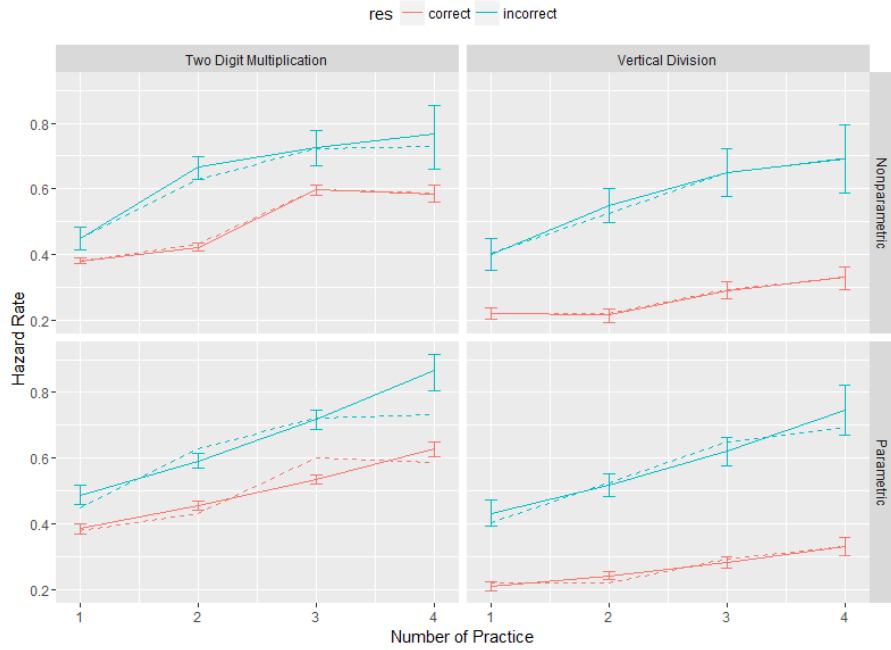


Figure 4.5: Conditional Hazard Curve of the LTP model

4.4.4 Estimated Pedagogical Efficacy

Since the selective sample attrition is largely accounted for, it is time to compare the pedagogical efficacy estimated by the BKT model and the LTP model. Table 4.5 shows the estimated pedagogical efficacy under the parametric specification. The LTP model produces significantly lower efficacy estimates. Because the long division has a larger difference in the conditional hazard rates, it also has a larger reduction in estimated pedagogical efficacy: While the pedagogical efficacy of the two-digit multiplication shrinks by a factor of 5, that of

the long division shrinks by a factor of 7.

Table 4.5: Estimated Pedagogical Efficacy of the Parametric Specification

Knowledge Point	Dependence	Estimates	95% CI(L)	95% CI(H)
Long Division	BKT	0.063	0.017	0.113
Long Division	LTP	0.009	0.000	0.026
Two-Digit Multiplication	BKT	0.121	0.070	0.174
Two-Digit Multiplication	LTP	0.024	0.003	0.056

Table 4.6 shows the estimated pedagogical efficacy of the nonparametric model. Since the two specifications have a similar fit of the hazard rate curve, they also have a similar estimation of pedagogical efficacy. Again, the LTP model produces significantly lower estimated practice efficacy.

Table 4.6: Estimated Pedagogical Efficacy of the Non-parametric Specification

Knowledge Point	Dependence	Estimates	95% CI(L)	95% CI(H)
Long Division	BKT	0.065	0.021	0.114
Long Division	LTP	0.009	0.001	0.026
Two-Digit Multiplication	BKT	0.130	0.076	0.178
Two-Digit Multiplication	LTP	0.026	0.003	0.062

4.4.5 The Fallacy of “Learning Gain” from Learning Curve

The introduction of this chapter argues that observed learning curve is not a reliable source of inference for learning gain. This subsection verifies this claim with empirical data. A large share of the observed “learning gain” in the learning curve may be the result of sample attrition rather than practice efficacy.

To isolate the effect of sample attrition from practice, a counterfactual learning curve under no sample attrition needs to be estimated. This can be done by using the parameter from the last section while setting the hazard rate at 0. Theorem 4.1 implies that LTP model with response-dependent sample attrition and the BKT model have the same parameter estimation other than the hazard rates. Therefore, the hazard rates associated with the BKT model is estimated from the LTP model with response-dependent sample attrition. Furthermore, because the parametric and non-parametric specifications have a similar estimation of pedagogical efficacy but the non-parametric specification has a slight advantage in fitting the hazard rate curve, this subsection only reports the result from the non-parametric hazard model. Figure 4.6 shows the imputed counterfactual learning curve in solid line, with an error bar of 95% credible interval. The dotted line is the observed learning curve from data. The counterfactual learning curve is almost flat because the estimated efficacy is small.

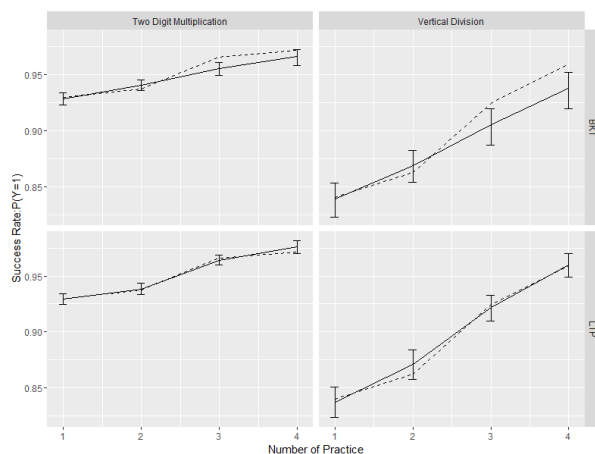


Figure 4.6: Counterfactual Learning Curve: No Sample Attrition

In sharp contrast, Figure 4.7 shows the fitted learning curve with both practice efficacy and sample attrition in solid line, with the error bar of 95% credible interval. Again, the dotted line is the observed learning curve from data. Both BKT and LTP model is able to fit the observed learning curve quite well.

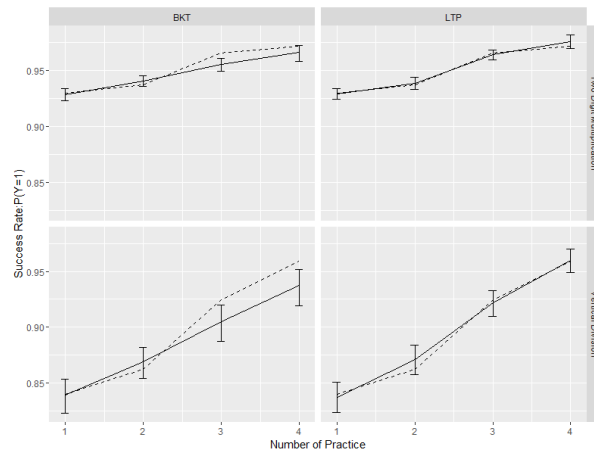


Figure 4.7: Counterfactual Learning Curve: With Sample Attrition

The comparison between Figure 4.6 and Figure 4.7 implies that the learning gain inferred from the observed learning curve is largely spurious because it is the result of sample attrition rather than practice efficacy. Table 4.7 calculated the “learning gain” from the counterfactual learning curve ($P(Y_4 = 1) - P(Y_1 = 1)$) as a percentage of the “learning gain” from the observed learning curve ($P(Y_4 = 1|H_3 = 0) - P(Y_1 = 1)$). The analysis shows that the practice efficacy only accounts for a small portion of observed “learning gain”. Under the BKT model, practice efficacy contributes about 50% in the two-digit multiplication and 20% in the long division. Under the LTP model, practice efficacy contributes about 10% in the two-digit multiplication and 5% in the long division! The true contribution of practice efficacy is likely to be between the two models. However, even take the upper bound of BKT model, practice efficacy only accounts for at most half of the observed “learning gain”.

Table 4.7: Percentage of Learning Gain Attributed to the
Practice Efficacy

Knowledge Point	BKT Efficacy Contribution(%)	LTP Efficacy Contribution(%)
Two-Digit Multiplication	52.6	11.73
Long Division	23.8	3.53

Chapter 5

Effort Choice and Efficacy Ranking

5.1 Introduction: Efficacy Versus Effectiveness

Eric Taylor (2015) surveyed the evaluation of new education technology and concluded that education technologies are generally pedagogically ineffective in raising test scores. There are two competing hypotheses in explaining this observation:

1. New technology does not improve teaching/learning productivity even under ideal conditions.
2. New technology improves productivity under ideal conditions but fails to do so in practical settings.

The medical literature makes a clear distinction between the two hypotheses (Flay, 1986). Efficacy is “whether a technology, treatment, procedure, or program does more good than harm when delivered under optimum conditions”. Effectiveness is “whether a technology, treatment, procedure, intervention, or program does more good than harm when delivered under real-world conditions”. Following this terminology, the first hypothesis is a test on efficacy, while the second hypothesis is a test on effectiveness.

In medical literature, the difference between efficacy and effectiveness are often illustrated by the observation that patients often do not adhere to dosage and timing prescribed by a doctor. Similarly, in a low stake learning environment, learners have considerable latitude in when and how they do learning activities assigned by an instructor. If the learner does not exert effort in learning, even efficacious intervention would be ineffective. As the proverb “No pain no gain” suggests, some efficacious pedagogical method may elicit the least amount of effort if the learner is not willing to put in the work.

The evaluation of intelligent tutoring systems usually does not make the distinction between efficacy and effectiveness. This makes sense because the program evaluation is usually performed in the dissemination stage (Flay et al., 2005) during which the effectiveness is the key measurement of success. However, from the perspective of product development via rapid iterations, the distinction is important. If the proposed pedagogical technology is not efficacious, a developer should stop experimenting to cut losses. If the proposed pedagogical technology is efficacious but ineffective, the developer can improve the effectiveness by adjusting the technology to user behavior.

To this end, this chapter proposes to rank efficacies of different pedagogies in a Randomized Control Trial (RCT) by the Learning Through Practice (LTP) model, instead of the Difference in Difference (DID) regression, if such RCT is conducted in a low stake learning environment. If learners choose different levels of effort in the treatment group as compared to the control group, the DID estimator does not guarantee to correctly rank the efficacies of the tested pedagogies. In contrast, the LTP model ranks efficacies correctly by accounting for different effort rates with a structural model of effort choice.

This chapter is organized as the following: The first section discusses how learner’s effort choice creates the difference between efficacy and effectiveness. Even when the RCT procedure grants no selection bias under the ideal condition for evaluation, learners’ choice can still be differentially affected by the group assignment, resulting in different discrepancies between

efficacy and effectiveness for different groups. The second section describes how the LTP model controls for the effort decisions to recover the correct relative efficacy ranking. The third section shows the DID estimator in general is not able to recover the right efficacy ranking in the presence of differential effort choices. The fourth section supports the analysis in the second and the third section with a simulation study. The fifth section applies both the LTP model and the DID estimator to an RCT dataset on geometry learning in third grade. It shows evidence of differential effort choices. Because of such differential choices, the DID estimator may fail to detect the statistically significant difference between two pedagogical efficacies while the LTP model successfully captures it by modeling the effort decision.

5.2 Effort Choice in Randomized Control Trials

The introduction argues that the evaluation of ITS development should distinguish between effectiveness and efficacy, but what is the ideal implementation condition that allows efficacy to be evaluated? Ericsson et al (1993; 2016) and Duckworth et al(2007) argues that passion and grit are essential for practice to be effective. Following this school of thought, this chapter regards the ideal condition for implementation as learners exert their best effort no matter how challenging or boring the learning task is.

The learning analytics literature has been aware that not all learners are paragons of grit in low stake learning for a decade. Ryan Baker(2004b; 2004a) first documented the off-task behavior in using an intelligent tutor system, which later he developed into the “Baker-Rodrigo Observation Method Protocol”. Following his research, there has been a few studies on the so-called “system gaming” behavior where learners avoid thinking by abusing the auxiliary instruction or hints (Koedinger and Aleven, 2007; Wixon et al., 2012). The same protocol is also used to study the correlation between daily effort and final test performance (Pardos et al., 2013). The education economics literature also starts to document the lack of full effort

in low stake testing (Levitt et al., 2013) and even in high stake testing (Metcalfe et al., 2011).

Whether or not the ideal condition for implementation is met has significant implication for the inference of the practice efficacy. Under the ideal condition, observed responses to practice problems are sufficient for ranking their relative efficacies. Given the same assessment questions, the problem that generates the highest positive difference in success rates between the pre-test and the post-test, aka effectiveness, must have the highest efficacy. However, under an imperfect implementation condition, the gain in success rate is the composite effect of the efficacy and the effort appeal of the problem. A problem with high efficacy but low effort appeal may generate less observed increase in success rate than a problem with low efficacy but high effort appeal. Consequently, the the ranking of effectiveness no longer mirrors the ranking of efficacy.

A Randomized Control Trial (RCT) does not preclude the problem of effort choice. A properly executed RCT requires random assignment of the treatment status and random sample attrition to justify the assumption that the treatment and the control group are similar in observed and unobserved characteristics. Under such ideal condition for evaluation, the estimated average treatment effect is free of selection bias. However, RCT does not, and need not, guarantee that learners always apply their best effort in the experiment. If the treatment pedagogy and the control pedagogy elicit different levels of learner effort, the difference between efficacy and effectiveness can also be different between the two groups. If the analysis only looks at the relative effectiveness based on the response data, such as the DID estimator, its inference on efficacy ranking may be incorrect.

In the next two sections, this chapter discusses the efficacy ranking within the context of RCT. For the LTP model, the RCT design ensures that the treatment group and the control group can be described by the same set of LTP parameters. For the DID regression, the RCT design ensures that the treatment group and the control group are valid counterfactuals. In addition, this chapter assumes that the pre-test and post-test items are the same for both

groups, so that effort rates on the training problems are the only source of difference effort choices.

5.3 The LTP model

This section reviews the effort choice in the Learning Through Practice (LTP) model.

5.3.1 Key Assumptions

The critical assumptions on how effort affects the learning process are:

Assumption 1 (No pain no gain): If learners do not exert effort at practice, they learn with probability 0.

Assumption 2 : If learners do not exert effort at practice, they produce a completely incorrect answer with probability 1, regardless of their latent mastery.

Assumption 3: The effort choice is independent and identically distributed, conditional on learner's latent mastery.

Assumption 1 is the source of difference between efficacy and effectiveness. Because this chapter adopts a loose definition of effort in the empirical case study, learners who exert no effort literally made no effort in understanding the training question, let alone solving it, and thus is very unlikely to learn anything from it.

Assumption 2 is not valid for all question types. For example, if the assessment item is a multiple choice problem, it is possible to make a lucky guess without any thinking. However, if the assessment item is a fill-in-the-blank problem, as in the empirical study, the probability of a blind correct guess is almost zero.

Assumption 3 imposes a very strong assumption on grit, as grit is the deciding factor behind

the effort choice in the LTP model. It assumes that the grit level is a characteristic of states of latent mastery, rather than that of a learner: When the learner switches from low mastery to high mastery, her grit level also changes. It is not compatible with the conventional understanding that grit is a personal trait that changes slowly. This chapter justifies Assumption 3 by arguing that it reflects two different aspects of grit. In low mastery, it is the resilience against failures. In high mastery, it is the resilience against boredom.

5.3.2 Consistent Estimation of Relative Efficacy

Theorem 5.1. *The point estimation of the efficacy of the LTP model is a consistent estimator of the true efficacy, if the LTP model correctly specifies the learning process.*

Proof. The point estimation of the efficacy of the LTP model is

$$\begin{aligned}\hat{\ell} &= \frac{\sum_t \sum_i I(X_{i,t} = 1, X_{i,t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T) + \alpha}{\sum_t \sum_i I(X_{i,t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T) + \alpha + \beta} \\ &= \frac{\sum_t \frac{\sum_i I(X_{i,t}=1, X_{i,t-1}=0 | Y_1, \dots, Y_T, E_1, \dots, E_T)}{N} + \frac{\alpha}{N}}{\sum_t \frac{\sum_i I(X_{i,t-1}=0 | Y_1, \dots, Y_T, E_1, \dots, E_T)}{N} + \frac{\alpha+\beta}{N}}\end{aligned}$$

In large sample, the probability limit of the point estimation is

$$\begin{aligned}\lim_{N \rightarrow \infty} \hat{\ell} &\rightarrow \frac{\sum_{t=1}^T P(X_t = 1, X_{t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T)}{\sum_{t=1}^T P(X_{t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T)} \\ &= \frac{\sum_{t=1}^T \ell P(X_{t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T)}{\sum_{t=1}^T P(X_{t-1} = 0 | Y_1, \dots, Y_T, E_1, \dots, E_T)} \\ &= \ell\end{aligned}$$

□

Lemma 5.1. *In a randomized control trial, the LTP model correctly ranks the efficacies, if the LTP model correctly specifies the learning process.*

Proof. The RCT implies that the treatment group and the control group share the same parameter set, except for the efficacies and effort rates on training questions that differ between the two groups.

Let D be the group assignment. $D = T$ is the treatment group, $D = C$ the control group. Let ℓ_D denote the efficacy of training question in group D . By Theorem 6.1,

$$\begin{aligned}\lim_{N \rightarrow \infty} \hat{\ell}_{D=T} &\rightarrow \ell_{D=T} \\ \lim_{N \rightarrow \infty} \hat{\ell}_{D=C} &\rightarrow \ell_{D=C}\end{aligned}$$

Therefore,

$$\lim_{N \rightarrow \infty} \hat{\ell}_{D=T} - \hat{\ell}_{D=C} \rightarrow \ell_{D=T} - \ell_{D=C}$$

□

5.4 The Difference in Difference Estimator

Econometricians use Difference in Difference (DID) regression to estimate the average treatment effect from an RCT. Let Y be the response. $Y = 1$ is a correct response. $Y = 0$ is the incorrect response. Let t marks the experiment phase. $t = 0$ is the pre-test phase and $t = 1$ is the post-test phase. Under random assignment of an RCT, Average Treatment Effect (ATE) can be expressed as

$$ATE = [E(Y_{t=1}|D = T) - E(Y_{t=1}|D = C)] - [E(Y_{t=0}|D = T) - E(Y_{t=0}|D = C)] \quad (5.1)$$

The DID estimator accounts for common growth from pre-test to post-test period (β_t). It also controls for potential difference in the correct rate between the treatment and control

groups in the pre-test period (β_D). Let β_0 be the average correct rate for the control group at the pre-test period. Let γ denote the ATE. The four components in equation ?? can be expressed as:

$$\begin{aligned} E(Y|D = C, t = 0) &= \beta_0 \\ E(Y|D = T, t = 0) &= \beta_0 + \beta_D \\ E(Y|D = C, t = 1) &= \beta_0 + \beta_t \\ E(Y|D = T, t = 1) &= \beta_0 + \beta_t + \beta_D + \gamma \end{aligned}$$

Let i be the learner id, D_i be the group assignment for the learner and $Y_{i,t}$ be the response of the learner in either pre-test or post-test period. $\hat{\gamma}$ from the following regression is the Difference in Difference estimator of ATE.

$$Y_{i,t} = \beta_0 + \beta_d D_i + \beta_t t + \gamma D_i t + \epsilon_{i,t}$$

5.4.1 Inferring Relative Efficacy from Relative Effectiveness

In a RCT, a qualitative judgment on the efficacy ranking is more important than a quantitative measurement on how large the efficacy difference is. Even though the DID estimator measures the effectiveness rather than the efficacy because it does not take effort choice into account, it is a valid inference tool for ranking if relative efficacy can be inferred from relative effectiveness. Although it is possible to do so if the control group has null efficacy, it is not always feasible if the control group has either positive or negative efficacy.

If the control group has a null efficacy, the placebo group for medical research for instance, the estimated effectiveness changes the magnitude of the efficacy but not its direction, thus relative efficacy can be inferred from relative effectiveness. When the efficacy of the treatment is positive but fails to elicit effort from participants, the estimated effectiveness is biased toward the null effect, but still positive. Similarly, if the efficacy of the treatment is negative,

the estimated effectiveness is biased toward the null effect, but still negative. Therefore, when the control group has no efficacy, the effort choice is essentially a power problem so far as the ranking inference is concerned.

If the control group has positive (or negative) efficacy, the order of efficacies may not be the same as the order of effectiveness. Consider the special case where the two pedagogies have identical positive efficacy, thus a true relative efficacy of ZERO. Differential effort choice can lead to either a positive or a negative relative effectiveness. If the treatment elicits full effort while the control practice elicits no effort, there is positive learning for treatment group, an increase in success rate of the post-test item, and a positive relative effectiveness. By the same token, if the treatment practice elicits no effort while the control practice elicits full effort, the relative effectiveness is negative. Therefore, when the control group has non-zero efficacy, the effort choice is a bias problem, as well as a power problem, for ranking inference.

5.4.2 Efficacy Ranking by the DID Estimator

This subsection develops the properties of efficacy ranking by the DID estimator under the assumption that the LTP model correctly specifies the data generating process. To be compatible with the DID estimator, the observed response has only two states ($M_Y = 2$). To simplify the discussion, assume the latent mastery also has only two states ($M_X = 2$) and there is only one learner type ($M_Z = 1$). The learner engagement include the effort decision but not the stop decision.

5.4.2.1 Notation

Let J denote the item sequence each group received. In the context of random control trials, the item sequence has the generic form of (pre-test, training, post-test). Let the pre-test be 0, the post-test be 1, the treatment training be T , and the control training be C . In short,

the treatment group receives the item sequence $J_T = (0, T, 1)$ and the control group receives the item sequence $J_C = (0, C, 1)$.

Let the initial mastery probability be π . Let the correct rate of item j conditional on the latent mastery of state k be c_j^k . Let the effort rate of the item j conditional on the latent mastery of state k be e_j^k . Let the pedagogical efficacy of the item j be ℓ_j . Thus the treatment's relative efficacy to the control can be expressed as $\Delta\ell = \ell_T - \ell_C$.

Let O_t be the observed response in sequence position t . The following definition describes the mapping between the observed response (O_t), the learner's true response (Y_t) and the learner's effort (E_t):

$$O_t = \begin{cases} 0 & \text{otherwise} \\ 1 & \text{if } Y_t = 1 \& E_t = 1 \end{cases}$$

5.4.2.2 Rank Ordering Without Effort Choice

When learners always exert effort, it can be proved that the DID estimator makes the right efficacy ranking in a large sample.

Theorem 5.2. *Without effort choice, the DID estimator correctly ranks pedagogical efficacies if there are learners without mastery before the experiment ($\pi < 1$), the pre-test item does not convert all learners without mastery to having mastery ($\ell_0 < 1$), and the post-test item discriminates between two latent states ($c_1^1 > c_1^0$).*

Proof. Let δ be the estimated relative effectiveness and $\Delta\ell$ be relative efficacy.

For the estimated effectiveness to have the same rank order as the efficacy, δ and $\Delta\ell$ need to satisfy the following properties.

1. If $\Delta\ell = 0$, $\delta = 0$.
2. If $\Delta\ell \neq 0$, $\delta\Delta\ell > 0$.

When there is no effort choice, the first difference within each group is

$$\begin{aligned}
\delta_D &= P(Y_1 = 1, Y_0 = 1|D) - P(Y_1 = 1, Y_0 = 1|D) \\
&= (1 - \pi)(1 - \ell_0)(1 - \ell_D)(c_1^0 - c_0^{1,0}) + ((1 - \pi)[(1 - \ell_0)\ell_D + \ell_0])(c_1^1 - c_0^{1,0}) \\
&\quad + (1 - \pi)(c_1^1 - c_0^{1,1})
\end{aligned}$$

The second difference is

$$\delta = \delta_T - \delta_C = \Delta\ell(1 - \pi)(1 - \ell_0)(c_1^1 - c_1^0)$$

If $\ell_0 < 1$, $\pi < 1$, and $c_1^1 > c_1^0$, it is easy to verify that rank order properties are satisfied. \square

5.4.2.3 Rank Order with Effort Choice

However, if the learning process involves the effort decision, the correct efficacy ranking requires a much stronger assumption that effort rates of learners without mastery are the same for the training question.

Theorem 5.3. *With effort choice, the DID estimator correctly ranks pedagogical efficacies if $\pi < 1$ and $\ell_0 < 1$, $c_1^1 e_1^1 - c_1^0 e_1^0$ and $e_T^0 = e_C^0$*

Proof. The first difference is

$$\begin{aligned}
\tilde{\delta}_D &= P(O_0 = 0, O_1 = 1|D) - P(O_0 = 1, O_1 = 0|D) \\
&= P(E_0 = 0, E_1 = 1, Y_1 = 1|D) + P(E_0 = 1, E_1 = 1, Y_0 = 0, Y_1 = 1|D) \\
&\quad - P(E_0 = 1, E_1 = 0, Y_0 = 1|D) - P(E_0 = 1, E_1 = 1, Y_0 = 1, Y_1 = 0|D) \\
&= (1 - \pi)(1 - \ell_0 e_0^0)(1 - \ell_D e_D)(c_1^0 e_1^0 - c_0^{1,0} e_0^0) \\
&\quad + ((1 - \pi)[(1 - \ell_0 e_0^0)\ell_D e_D^0 + \ell_0 e_0^0])(c_1^1 e_1^1 - c_0^{1,0} e_0^0) + (1 - \pi)(c_1^1 e_1^1 - c_0^{1,1} e_0^0)
\end{aligned}$$

The second difference is

$$\tilde{\delta} = \tilde{\delta}_T - \tilde{\delta}_C = (\ell_T e_T^0 - \ell_C e_C^0)(1 - \pi)(1 - \ell_0 e_0^0)(c_1^1 e_1^1 - c_1^0 e_1^0)$$

When $\Delta\ell = 0$, $\tilde{\delta} = 0$ only when $e_T^0 = e_C^0$.

If $e_T^0 = e_C^0$, when $\Delta\ell \neq 0$, $\tilde{\delta}\Delta\ell > 0$ requires $c_1^1 e_1^1 - c_1^0 e_1^0 > 0$. □

Theorem 5.3 essentially says the DID estimator correctly recovers the rank order only when the effort choices of learners without mastery are the same for the treatment pedagogy and the control pedagogy. It implies that the DID estimator does not preserve the rank order of efficacies when the effort choices differ by practice questions, which is more realistic in a low stake learning environment.

5.4.2.4 Rank Ordering Conditional on Effort

One intuitive correction technique is to condition the analysis on learners who exert efforts on the training question. However, the conditional DID estimator does not guarantee a correct efficacy rank either.

Lemma 5.2. *Conditioning on the effort on the training question is not a sufficient condition of the DID estimator being a monotone function of the pedagogical efficacy difference.*

Proof. The analysis is almost the same as Theorem 6.3 except for the initial mastery probability is $P(X_0 = 1|E_D = 1)$. It can be shown that

$$\tilde{\pi}_D = P(X_0 = 1|E_D = 1) = \frac{\pi e_D^1}{[(1 - \pi)(1 - \ell_0)e_0 + (1 - \pi)(1 - e_0)]e_D^0 + [(1 - \pi)\ell_0 e_0 + \pi]e_D^1}$$

Plug it in the DID estimator

$$\tilde{\delta} = \tilde{\delta}_T - \tilde{\delta}_C = [\ell_T(1 - \tilde{\pi}_T) - \ell_C(1 - \tilde{\pi}_C)](1 - \ell_0 e_0^0)(c_1^1 e_1^1 - c_1^0 e_1^0) + (\tilde{\pi}_T - \tilde{\pi}_C)(c_1^1 - c_0^1)e_1^1$$

Even under the favorable condition that $c_1^1 = c_0^1$, the correct rank order requires

$$(\ell_T - \ell_C)\left(\frac{\ell_T}{\ell_C} - \frac{1 - \tilde{\pi}_T}{1 - \tilde{\pi}_C}\right) > 0$$

In general, the inequality is not true. □

The key insight from Lemma 5.1 is that conditioning on the effort is a selection process. Such selection process breaks the valid counterfactual assumption granted by the random assignment of the RCT design. Therefore the estimated effectiveness of each group is biased. Although it can be argued that a biased estimator of the effectiveness is not sufficient to conclude that the difference of two biased estimators does not preserve the rank order of pedagogical efficacies. The author wishes to point out that it is easy to construct a counterexample where such bias leads to a wrong rank order.

5.5 Simulation Study

The simulation demonstrates Theorem 5.2 and Theorem 5.3 in the previous section. The true efficacy difference is set to be 0.4. Without effort choice, the true effectiveness difference for the DID estimator is 0.176. With effort choice, the true effectiveness difference for the DID estimator is 0, where the differential effort choices offset the efficacy difference entirely. The values of all simulation parameters are in Appendix B.4.

Use the 95% credible interval as the decision rule for ranking. If the interval is to the right of 0, the treatment ranks first. If the interval is to the left of the 0, the control ranks first. The

simulation study is interested in whether the DID estimator and the LTP estimator make the right rank decision.

Table 5.1 shows the result of the DID estimator. When the data generating process does not involve the effort choice, the DID estimator consistently estimates the relative effectiveness and thus correctly ranks the two groups. However, when the data generating process involves the effort choice, the 95% credible interval contains zero, thus the DID estimator fails to rank the treatment higher than the control.

Table 5.1: Simulation Result of the DID estimator

D.G.P.	Effectiveness	Est.Effectiveness	95% CI(L)	95% CI(H)
Without Effort	0.176	0.147	0.097	0.197
With Effort	0.176	0.029	-0.021	0.079

In contrast, Table 5.2 shows that the LTP estimator always recovers the correct rank order whether the data generating process has the effort choice or not.

Table 5.2: Simulation Result of the LTP estimator

D.G.P.	Efficacy	Est.Efficacy	95% CI(L)	95% CI(H)
Without Effort	0.4	0.426	0.300	0.563
With Effort	0.4	0.458	0.361	0.555

5.6 Case Study

This section applies the LTP model to a randomized control trial. The section first describes the data collection process of the experiment, then describes the identification of the effort

choice, and finally shows different rankings of the DID estimator and the LTP model.

5.6.1 The Learning Environment

The experiment is carried out in a paid self-learning product offered by a Chinese online learning service provider. The product is used after school, rather than in the classroom. The product is framed as a role-playing game where a learner clears a level to claim some reward by succeeding in practice questions. The screen shots of the selection, the practice and the completion of a level are shown in Figure 5.1.



Figure 5.1: Stages of a Level in the Learning Game

The product is in a low stake learning environment because it offers a small monetary incentive for good performance and no direct punishment for poor performance. If a learner correctly answers the question, she receives *xuedou*, an in-game virtual currency, that can be used to buy in-game gears or real world gifts. During the experiment, the reward for each correct response is valued at about a tenth of ¥\$0.01 (or \$0.00014). If the learner answers the

question incorrectly, she receives nothing and loses nothing on the spot, but is offered a second chance to try again later for a smaller reward. Because the stake in this learning environment is low, it does not inspire learners to struggle when challenged with difficult questions. In addition, because each correct answer is worth the same regardless of its difficulty, some students develop a strategy of picking easy items and skipping hard items. Both low incentive and the system design flaw keep learners from exerting their full effort in practice.

5.6.2 Raw data

The log data collected for the experiment include the following fields:

1. learner id
2. question id
3. Submit time
4. time spent on the question (seconds)
5. the grade (0-1)
6. the actual answer in the text

The time spent on the item is defined as the time elapsed between the timestamp the server sends out the question to the learner’s device and the timestamp the learner’s answer received by the server. The transmission time in the network is negligible, usually in the magnitude of 10 milliseconds. However, the time spent does not distinguish how much time spent on each sub question. In addition, it is not a clean measure of student’s active learning time. What student did between when the question is presented and when the answer is submitted is not observed. Although there are 4% of the observations logged response time longer than 2 minutes, the majority of the response time falls within a reasonable time range.

5.6.3 The Design and the Implementation of The Experiment

The experiment is administrated from June 9th to June 10th 2016 to third-grade students whose parents paid for the learning product. By then, all learners should have learned the pre-requisite knowledge in the school: calculating circumference and area of a rectangle.

Each group receives three items, the pre-test item, the training item and the post-test item, as shown in Figure 5.2. All items ask the learner to calculate the circumference and the area of two identical rectangles joined by one side given the length and width of the small rectangle. The pre-test item and the training item are rectangles joined by length but have different parameter values. The pre-test and post-test item share the same parameter value but whose small rectangles are joined by different sides. The highly similar yet not identical test items aim to increase the measurement validity while preventing students from memorizing the answer.

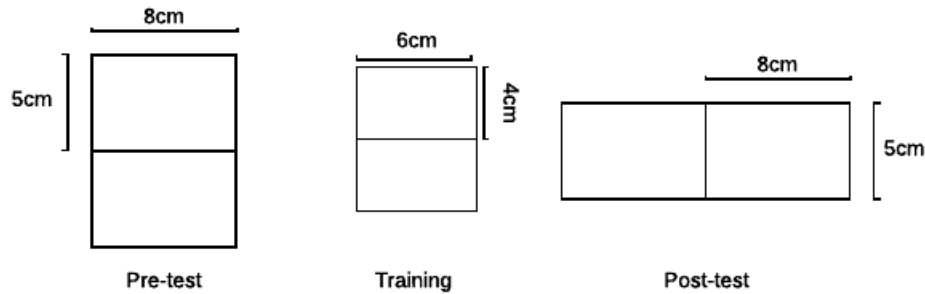


Figure 5.2: Descriptions of Experiment Problems

The learners are randomly assigned to the treatment group and the control group by user id, which is randomly generated. The treatment and the control group differ in the pedagogical methods. Both the control group and the treatment group receives the training question, while the treatment group receives an additional link to an animated video tutorial containing the following instructions:

-
1. Calculate the circumference and the area of the small rectangle
 2. To get the circumference of the large rectangle, multiply the circumference of the small rectangle by two and subtract two times of the length of the joined side.
 3. To get the area of the large rectangle, multiply the area of the small rectangle by two

The learner can choose to skip the video tutorial. Unfortunately, the system does not track the learners' watch history by their user ids. On aggregate, the video was played about 800 times. If assume each view is a different learner, the maximum exposure to the treatment is about 30%. The video tutorial is 67 seconds long. The average play time is about 47 seconds.

5.6.4 Differential Effort Choice

The identification of the LTP model hinges on the identification of the effort level. The effort level is not directly observed but has to be inferred from the log data.

5.6.4.1 Effort Classification

This chapter identifies the effort from the actual text response student submitted. A correct response or an honest error is labeled as “effort” while a blank answer or a non-sensical answer is labeled as “no-effort”/“give-up”. This classification method is also used by Zamarro et al (2016) to validating measures of non-cognitive skills by analyzing respondents's answer behavior to a psychological survey.

In the context of this experiment, there are only four honest mistakes:

1. Wrong shape: The learner calculates correctly either the circumference or the area of the small rectangle
2. Wrong circumference: The learner calculates correctly the area but not the circumference
3. Wrong area: The learner calculates correctly the circumference but not the area

-
4. Slip: The answer is correctly calculated but the learner inputs in a wrong way (for example an extra zero at the end)

Table 5.3 shows a breakdown of the answer classification. The classification is discussed in details in Appendix C.

Table 5.3: Answer Composition - All Wrong

Group	Task	No-Effort(%)	S.E.	Honest Error (%)	S.E.	Correct(%)	S.E.
Control	pre	24.9	0.870	28.4	0.908	46.8	1.00
Treatment	pre	26.9	0.950	27.4	0.954	45.7	1.07
Control	train	30.5	0.927	22.5	0.842	47.0	1.00
Treatment	train	33.9	1.013	20.9	0.870	45.3	1.07
Control	post	35.4	0.963	15.9	0.736	48.7	1.01
Treatment	post	35.6	1.025	16.1	0.787	48.3	1.07

5.6.4.2 Robustness Check of the Effort Classification

Time spent on the item is suggestive of the learner’s level of effort. Figure 5.3 shows its distribution. While the response time distribution of the honest error is very similar to that of the correct answer, the response time distribution of the no-effort answer is very different to the other two. The median time spent on an honest error or a correct answer is about 47 seconds while that on a no-effort answer is about 10 seconds. The first quartile of the response time of an honest error or a correct answer is about 34 seconds, which is higher than the third quartile of the response time of a no-effort answer, 29 seconds. The similarity between the honest error and the correct answer and the difference of the two from the no-effort answer validates the effort classification.

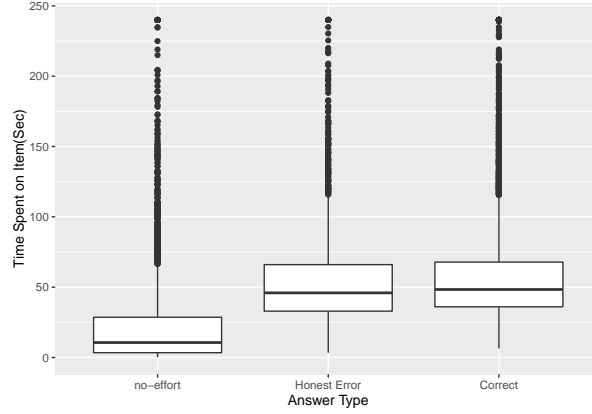


Figure 5.3: Boxplot of Time Spent by Answer Type

It may be argued that the classification strategy underestimates the level of effort because it is possible that learners do think hard on these problems despite putting in a non-sensical answer or leaving it blank. Although it may be an issue for the pre-test item, such miscoding is less problematic for the training item and the post-test item. Figure 5.4 shows the distribution of response time when submitting a no-effort answer. Compared to the pre-test item, the response time distribution of the training item and the post-test item shifts to the left and skews toward 0. The learners are spending less time on the problem, so little that it may not them to read the question carefully. Therefore, the effort identification is likely to be more accurate in the next two items.

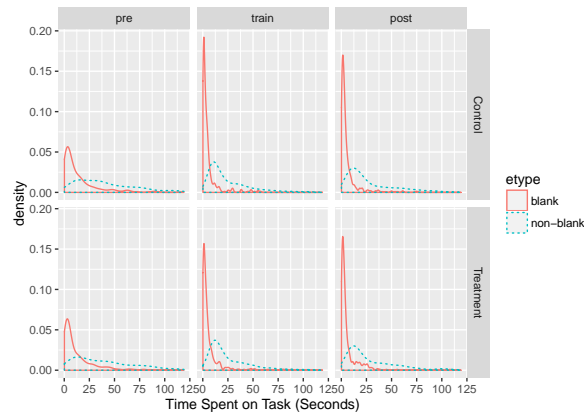


Figure 5.4: Distribution of Time Spent on Item without Effort by Answer Classification

In contrast, answers with actual effort input do not demonstrate the same shift. Figure 5.5 shows the response time distribution when submitting a valid answer. A honest error or a correct answer is taxing on one’s mind and takes time, while a no-effort error made on the fly. The comparison of response time distribution between effort exertion and the lack of it further substantiates the classification.

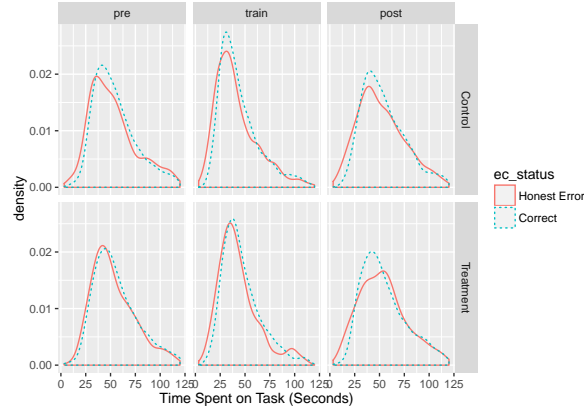


Figure 5.5: Distribution of Time Spent on Item With Effort by Answer Classification

5.6.4.3 Differential Effort Level Choice

Table 5.4 lists the summary statistics of the effort rate group by the experiment sequence and the group assignment. The treatment group exerts less effort compared to the control group in the pre-test period and the training period, the difference of which is statistically significant. The effort gap narrows in the post-test period. Ideally, the two groups should have the same effort rate in the pre-test period but different effort rates in the training session. The different effort rates in the pre-test period suggest that the randomization is not perfect. However, the effort rate gap widens from 2.3% in the pre-test period to 3.4% in the training session. Although the widening of the effort gap is not statistically significant, it suggests that treatment pedagogy elicits less effort from the learners.

Table 5.4: Effort Rate Choices - All

Period	Group	Mean Effort Rate	Standard Error
pre	Control	0.756	0.004
pre	Treatment	0.733	0.004
train	Control	0.696	0.004
train	Treatment	0.662	0.005
post	Control	0.646	0.005
post	Treatment	0.645	0.005

Further, this chapter wants to measure the effort gap of the learners with mastery and those without. From the data, one can only observe the effort rate conditional the response, but not on the state of latent mastery. That said, the state of latent mastery can be partially approximated by the response in the pre-test item. Because the pre-test item is a fill-in-the-blanks question, the probability of making a correct guess when the learner has no mastery is small enough to be negligible. Therefore, if a learner correctly answers the question, she must have mastered the knowledge point. If the learner incorrectly answers the question, it is unclear whether it is because she has no mastery or because she does not want to exert effort. Because the learners who fail the pre-test is a mixture of learners with and without mastery, if the effort gap of learners who correctly answers the pre-test is smaller than that of learners who incorrectly answers the pre-test, the latter becomes the lower bound of the effort gap of learners without mastery.

Table 5.5 reports the effort rates for the learners partitioned by the response in the pre-test item. The effort rate gap is 1.6% for the learners who (presumably) have mastery because they correctly answer the pre-test question. It is lower than the effort gap of learners who incorrectly answer the pre-test question, which is around 4%. Therefore, the effort rate gap

for the learners without mastery must be larger than 4%. Even at the lower bound, the effort rate gap of learners without mastery is about 2 times larger than those with mastery.

Table 5.5: Effort Rate Choice - By Response in the Pre-test

Period	Group	Mean Effort Rate(Y=0)	S.E.	Mean Effort Rate(Y=1)	S.E.
pre	Control	0.542	0.007	1.000	0.000
pre	Treatment	0.508	0.007	1.000	0.000
train	Control	0.483	0.007	0.938	0.002
train	Treatment	0.443	0.007	0.922	0.002
post	Control	0.427	0.007	0.895	0.003
post	Treatment	0.430	0.007	0.900	0.003

The previous analysis suggests that the treatment group exerts less effort than the control group and the difference comes mainly from the learners without mastery. However, readers may worry about selection bias because the effort rates differ in the pre-test item, suggesting that the RCT is not properly executed. The author acknowledges this defect and wishes to conduct a better-executed RCT in future research.

5.6.5 The result

This section compares the DID estimator to the LTP estimator. It demonstrates how differential effort rates can lead to a qualitatively different ranking of the pedagogical efficacy. The consensus among pedagogical experts is that the presence of a video instruction should increase the pedagogical efficacy of the treatment training question. Especially, because the circumference of the large rectangle is not simply double that of the small rectangle, they expect the special trick in the video to produce a larger gain in calculating the circumference.

The LTP model confirms both prior expectations of the domain experts, while the DID estimator confirms none.

5.6.5.1 Relative Effectiveness and Relative Efficacy in Aggregate

Both the DID estimator and the LTP estimator are applied to the aggregate score. Table 5.6 shows the result. The point estimation of relative effectiveness given by the DID estimator is close to 0. Because its 95% confidence interval contains 0, the DID estimator suggests that the efficacy ranking of the two questions cannot be determined. In contrast, the point estimation of relative efficacy given by the LTP model is positive. Because its 95% credible interval excludes 0, the LTP estimator suggests that the training question with video instructions has better efficacy than the question without.

Table 5.6: Estimated Relative Effectiveness and Relative Efficacy in Aggregate Score

Model	Est. Rel Effective	Est. Rel Efficacy	95% CI(L)	95% CI(H)	Treatment Win
DID	0.007	NA	-0.034	0.048	N
LTP	NA	0.119	0.014	0.228	Y

5.6.5.2 Relative Effectiveness and Relative Efficacy by Knowledge Components

To correctly solve the problem, a learner must be proficient in three knowledge components: shape identification, circumference calculation, and area calculation. Therefore, it is possible to score each knowledge components and analyze the respective pedagogical efficacy ranking. The effort choice is inherited from the aggregate score. For example, if the learner leaves a blank for circumference but fills in the area, the response is considered a valid effort for all three components, rather than treating it as a no-effort for the circumference due to the

blank answer.

According to the prior expectation of the pedagogical experts, the video scaffolding is likely to be more effective in teaching how to calculate the circumference, because it introduces a new method to calculate the circumference that is less likely to cause confusion. Table 5.7 reports the estimated relative effectiveness and relative efficacy by these components. The breakdown analysis indeed shows that the question with video instruction has higher pedagogical efficacy in teaching circumference, but not in teaching the other two components. This finding supports the plausibility of the claim that the LTP estimator is truly estimating the pedagogical efficacy.

Table 5.7: Estimated Relative Effectiveness and Relative Efficacy by Knowledge Components

Knowledge	Model	Est. Rel Effect	Est. Rel Efficacy	95% CI(L)	95% CI(H)
Area	LTP	NA	0.029	-0.124	0.187
Area	DID	0.018	NA	-0.021	0.058
Circumference	LTP	NA	0.234	0.087	0.397
Circumference	DID	0.008	NA	-0.032	0.048
Shape	LTP	NA	0.023	-0.110	0.154
Shape	DID	0.014	NA	-0.025	0.053

Appendix A

Notation

i : Index for learner. The count starts from 1.

j : Index for practice item. The cardinality of the item id is J . The count starts from 1.

k, m, n : States/levels of the latent mastery, usually $m < n$. The count starts from 0.

t : Index for practice sequence. The cardinality of the sequence id is T .

r : States/levels of the observed response. The count starts from 0.

z : The type of a learner

A : The assignment function that maps the items to the practice sequence. $A_t = j$ means assign item j to sequence t . \mathbf{A}_{t_1, t_2} is mapping from sequence position t_1 to t_2 .

H : The exit decision. Binary variable. If the learner stops at sequence t , $H_t = 1$; otherwise, $H_t = 0$. \mathbf{H}_{t_1, t_2} is the joint exit decision $(H_{t_1}, \dots, H_{t_2})$.

E : The effort decision. Binary variable. If the learner exerts effort at sequence t , $E_t = 1$; otherwise, $E_t = 0$. \mathbf{E}_{t_1, t_2} is the joint effort decision $(E_{t_1}, \dots, E_{t_2})$.

X : The latent mastery at position t . Discrete variable. The cardinality of the possible state is M_X . \mathbf{X}_{t_1, t_2} is the joint latent mastery $(X_{t_1}, \dots, X_{t_2})$.

Y : The observed response at position t . Discrete variable. The cardinality of the possible state is M_Y . \mathbf{Y}_{t_1, t_2} is the joint responses $(Y_{t_1}, \dots, Y_{t_2})$.

Z : The learner's type. Discrete variable. The cardinality of the possible learner type is M_Z .

Ω : The sample space of the re-parametrized joint distribution of observed response, stop decision and effort decision. Its cardinality is N_{Ω} .

α_z : The probability that the learner is type Z .

$\beta_{z; k}$: The proportion hazard rate of one additional practice opportunity given latent mastery k of learner type z in the proportional hazard model

$c_j^{r, k}$: The probability of observing response r given item j and latent mastery level k .

$h_t^{z;k}$: The hazard rate at sequence t given latent mastery k of learner type z in the non-parametric hazard model.

h_t^r : The hazard rate at sequence t given observed response r in the non-parametric hazard model.

$\ell_j^{z;m,n}$: The pedagogical efficacy. The upper label describes the transition states (m, n) of the latent mastery for learner type z while the lower label describes the transition context, item id (j) .

$\pi^{z;k}$: The probability of latent mastery k for learner type z at the beginning of the practice.

$\gamma_j^{z;k}$: The probability of exerting effort given item j and latent mastery level k .

$\lambda_{z;k}$: The baseline hazard rate given latent mastery k of learner type z in the proportional hazard model

Appendix B

Simulation Parameters

B.1 MCMC Algorithm Convergence Demo: Single Type

The initial state density is:

$$P(X) = \begin{cases} 0.6 & \text{if } X = 1 \\ 0.3 & \text{if } X = 2 \\ 0.1 & \text{if } X = 3 \end{cases}$$

The state transtion matrix for item 1 is

$$\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0 & 0.4 & 0.6 \\ 0 & 0 & 1 \end{bmatrix}$$

The state transtion matrix for item 2 is

$$\begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0 & 0.7 & 0.3 \\ 0 & 0 & 1 \end{bmatrix}$$

The observation matrix for item 1 is

$$\begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.2 & 0.6 & 0.2 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

The observation matrix for item 2 is

$$\begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0 & 0.1 & 0.9 \end{bmatrix}$$

The effort matrix for item 1 is

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \\ 0.1 & 0.9 \end{bmatrix}$$

The effort matrix for item 2 is

$$\begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \\ 0.01 & 0.99 \end{bmatrix}$$

The baseline hazard rate is $\lambda = 0.1$ for all states.

The duration dependence is

$$\beta = \begin{cases} 1.2 & \text{if } X = 1 \\ 1.1 & \text{if } X = 2 \\ 1 & \text{if } X = 3 \end{cases}$$

B.2 MCMC Algorithm Convergence Demo: Multiple Types

The first type is 60% of the sample ($\alpha_1 = 0.6$). The first item occurs with 50% probability ($P(A_t = 1) = 0.5$).

For the type 1 learner, the initial mastery is 0.1 ($\pi_1 = 0.1$). The pedagogical efficacy of the first item is 0.3 ($\ell_{1,1} = 0.3$), that of the second item is 0.7 ($\ell_{2,1} = 0.7$).

For the type 2 learner, the initial mastery is 0.7 ($\pi_2 = 0.7$). The pedagogical efficacy of the first item is 0.2 ($\ell_{1,2} = 0.2$), that of the second item is 0.8 ($\ell_{2,2} = 0.8$).

For the both learner types, The observation matrix for item 1 is

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

The observation matrix for item 2 is

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

1000 learners are simulated. The practice sequence is 5.

B.3 The Dynamic Selection Bias Demo

$$\begin{aligned}
M_x &= 2 \\
M_y &= 2 \\
P(X_1 = 1) &= 0.4 \\
P(X_t = 1|X_{t-1} = 0) &= 0.3 \\
P(Y_t = 1|X_t = 1) &= 0.9 \\
P(Y_t = 1|X_t = 0) &= 0.2
\end{aligned}$$

B.4 The Effort-induced Measurement Error Demo

$$\begin{aligned}
P(X_0 = 1) &= \pi = 0.4 \\
P(X_t = 1|X_{t-1} = 0, 0) &= \ell_0 = 0.1 \\
P(X_t = 1|X_{t-1} = 0, T) &= \ell_T = 0.7 \\
P(X_t = 1|X_{t-1} = 0, C) &= \ell_C = 0.3 \\
P(Y_1 = 1|X_t = 1) &= c_1^1 = 0.9 \\
P(Y_1 = 1|X_t = 0) &= c_1^0 = 0.1
\end{aligned}$$

In the effort choice model, the effort rate is specified to offset the difference in the pedagogical efficacy. All effort rates are 1 except for

$$\begin{aligned}
P(E_T = 1|X_t = 0) &= e_T^0 = 0.3 \\
P(E_C = 1|X_t = 0) &= e_C^0 = 0.7
\end{aligned}$$

3000 learners are simulated. Half of the learner receives item sequence (0,T,1) and the other receives item sequence (0,C,1).

Appendix C

Answer Classification

The answers are initially classified into six categories:

- (1) Blank answer: The learner submits nothing on the circumference and the area
- (2) Non-blank wrong answer: Neither circumference nor area is correctly calculated and not includes in the slip or the wrong shape category
- (3) Slip: Both circumference and area are correctly calculated but the learner type in the wrong way. For example, switch the two blanks or add the unit of measurements
- (4) Wrong Shape: The learner correctly calculates either the circumference or the area but for the small rectangle.
- (5) wrong circumference: The learner correctly calculates the area of the large rectangle but not the circumference
- (6) wrong area: The learner correctly calculates the circumference of the large rectangle but not the area
- (7) Correct Answer: Both circumference and area of the large rectangle are correctly calculated

Table C.1 and Table C.2 show the summary statistics of different groups' answer composition at a different stage of the experiment. There is a steady increase in the blank answer as the item sequence progresses. The nonblank answer increases most in the training session.

Table C.1: Answer Composition (Percentage) - All Wrong

Group	Task	Blank Ans	Non Blank Wrong Ans
Control	pre	10.7	13.6
Treatment	pre	11.4	15.3
Control	train	14.6	15.9
Treatment	train	16.3	17.6
Control	post	18.4	17.0
Treatment	post	18.2	17.3

For learners who make a valid effort but still do not get all blanks correct, the percentage of the wrong shape declines steadily, but not the wrong area and the wrong circumference. In fact, there is a large increase in the wrong area in the post-test item and a large decrease in the wrong circumference. This observation is consistent with the step-wise efficacy analysis.

Table C.2: Answer Composition (Percentage) - Partial Correct

Group	Task	Wrong Shape	Wrong Area	Wrong Circumference	Slip
Control	pre	8.80	8.84	11.23	1.095
Treatment	pre	7.70	8.89	11.00	0.733
Control	train	5.07	5.35	12.16	0.324
Treatment	train	3.53	4.90	12.47	0.367
Control	post	1.91	7.54	6.49	4.420
Treatment	post	1.60	7.10	7.47	4.812

The reader may also be interested in the detailed breakdown of the answer composition. Other than the non-blank wrong answer, all four error categories have clustered answer patterns: The top 5 answers cover over 50% of the answers. The category of the non-blank answer has a wide dispersion. The top non-blank answer for all but the pre-test item is the correct answer to the pre-test question, which occupies about 30% of the non-blank answer.

As a reminder, the right answer for the pre-test item is (36,80), for the training item (28,48), for the post-test item (42,80).

For the non-blank answer, the most common error of the training item and the post-test item is (38,80), which is the correct answer to the pre-test item. The second common error for the post-test is (28,48), which is the correct answer to the training item. The learners just memorized the answer key from the previous question, which signals that they are not making effort.

Table C.3: Answer Breakdown: Nonblank Answer

qtype	raw_ans	n	pct	cum_pct
pre	[[“80”, “40”]]	33	0.049	0.049
pre	[[“80”, “”]]	17	0.025	0.075
pre	[[“13”, “40”]]	12	0.018	0.093
pre	[[“23”, “40”]]	12	0.018	0.110
pre	[[“52”, “160”]]	12	0.018	0.128
train	[[“36”, “80”]]	256	0.331	0.331
train	[[“16”, “16”]]	27	0.035	0.366
train	[[“48”, “24”]]	23	0.030	0.395
train	[[“20”, “16”]]	14	0.018	0.413
train	[[“36”, “40”]]	9	0.012	0.425
post	[[“36”, “80”]]	197	0.247	0.247
post	[[“28”, “48”]]	58	0.073	0.320
post	[[“36”, “40”]]	16	0.020	0.340
post	[[“80”, “40”]]	14	0.018	0.358
post	[[“32”, “40”]]	7	0.009	0.367

The top category of the wrong shape is the right circumference and the right shape of the small rectangle, which is not surprising.

Table C.4: Answer Breakdown: Wrong Shape

qtype	raw_ans	n	pct	cum_pct
pre	[[“26”, “40”]]	316	0.821	0.821
pre	[[“26”, “”]]	32	0.083	0.904
pre	[[“40”, “40”]]	14	0.036	0.940
pre	[[“40”, “”]]	9	0.023	0.964
pre	[[“40”, “26”]]	8	0.021	0.984
train	[[“20”, “24”]]	160	0.792	0.792
train	[[“20”, “”]]	12	0.059	0.851
train	[[“24”, “24”]]	11	0.054	0.906
train	[[“24”, “”]]	7	0.035	0.941
train	[[“20”, “20”]]	5	0.025	0.965
post	[[“26”, “40”]]	67	0.817	0.817
post	[[“26”, “”]]	4	0.049	0.866
post	[[“40”, “40”]]	4	0.049	0.915
post	[[“”, “40”]]	3	0.037	0.951
post	[[“40”, “”]]	2	0.024	0.976

It is not surprising that either the circumference or the double of the circumference of the small rectangle is in top 3 wrong answer category of the wrong circumference. In all three

items, substitute the circumference with a (wrong) area is also a common mistake. For example, (160,80) in the pre-test and (24,48) in the training.

Table C.5: Answer Breakdown: Wrong Circumference

qtype	raw_ans	n	pct	cum_pct
pre	[[“26”, “80”]]	128	0.248	0.248
pre	[[“160”, “80”]]	43	0.083	0.331
pre	[[“52”, “80”]]	41	0.079	0.410
pre	[[“32”, “80”]]	36	0.070	0.480
pre	[[“96”, “80”]]	33	0.064	0.544
train	[[“24”, “48”]]	133	0.233	0.233
train	[[“20”, “48”]]	115	0.201	0.434
train	[[“30”, “48”]]	42	0.073	0.507
train	[[“40”, “48”]]	41	0.072	0.579
train	[[“48”, “48”]]	33	0.058	0.636
post	[[“26”, “80”]]	56	0.173	0.173
post	[[“52”, “80”]]	31	0.096	0.269
post	[[“32”, “80”]]	23	0.071	0.341
post	[[“46”, “80”]]	20	0.062	0.402
post	[[“40”, “80”]]	20	0.062	0.464

The most common mistake of the wrong area is using the area of the small rectangle.

Table C.6: Answer Breakdown: Wrong Area

qtype	raw_ans	n	pct	cum_pct
pre	[[“36”, “40”]]	126	0.306	0.306
pre	[[“36”, “”]]	126	0.306	0.612
pre	[[“36”, “18”]]	20	0.049	0.660
pre	[[“36”, “160”]]	19	0.046	0.706
pre	[[“36”, “36”]]	11	0.027	0.733
train	[[“28”, “24”]]	74	0.310	0.310
train	[[“28”, “”]]	41	0.172	0.481
train	[[“28”, “32”]]	21	0.088	0.569
train	[[“28”, “60”]]	8	0.033	0.603
train	[[“28”, “28”]]	7	0.029	0.632
post	[[“42”, “40”]]	64	0.188	0.188
post	[[“42”, “”]]	49	0.144	0.331
post	[[“42”, “60”]]	35	0.103	0.434
post	[[“42”, “160”]]	10	0.029	0.463
post	[[“42”, “70”]]	10	0.029	0.493

Slip mainly involves adding 1 before or 0 after the right answer.

Table C.7: Answer Breakdown: Slip

qtype	raw_ans	n	pct	cum_pct
pre	[[“36”, “180”]]	16	0.372	0.372
pre	[[“36”, “800”]]	6	0.140	0.512
pre	[[“36cm”, “80cm2”]]	4	0.093	0.605
pre	[[“3680”, “”]]	3	0.070	0.674
pre	[[“360”, “80”]]	2	0.047	0.721
train	[[“2848”, “”]]	10	0.625	0.625
train	[[“48”, “28”]]	3	0.188	0.812
train	[[“128”, “48”]]	1	0.062	0.875
train	[[“28”, “448”]]	1	0.062	0.938
train	[[“28”, “488”]]	1	0.062	1.000
post	[[“42”, “90”]]	108	0.505	0.505
post	[[“41”, “80”]]	53	0.248	0.752
post	[[“42”, “08”]]	18	0.084	0.836
post	[[“4280”, “”]]	16	0.075	0.911
post	[[“42”, “800”]]	4	0.019	0.930

Bibliography

- Aleven, V. A. and Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive science*, 26(2):147–179.
- Anderson, J. R. (2000). *Learning and memory*. John Wiley New York.
- Anderson, J. R. (2013). *Language, memory, and thought*. Psychology Press.
- Baker, R. S., Corbett, A. T., and Aleven, V. (2008). More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In *International Conference on Intelligent Tutoring Systems*, pages 406–415. Springer.
- Baker, R. S., Corbett, A. T., and Koedinger, K. R. (2004a). Detecting student misuse of intelligent tutoring systems. In *International Conference on Intelligent Tutoring Systems*, pages 531–540. Springer.
- Baker, R. S., Corbett, A. T., Koedinger, K. R., and Wagner, A. Z. (2004b). Off-task behavior in the cognitive tutor classroom: when students game the system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 383–390. ACM.
- Baker, R. S., D’Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners’ cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68(4):223–241.
- Beck, J. E. and Chang, K.-m. (2007). Identifiability: A fundamental problem of student modeling. In *International Conference on User Modeling*, pages 137–146. Springer.
- Blackwell, D. and Koopmans, L. (1957). On the identifiability problem for functions of finite markov chains. *The Annals of Mathematical Statistics*, pages 1011–1015.
- Carlson, J. E. (2013). Item response theory. <https://www.ets.org/Media/Research/pdf/RR-13-28.pdf>.
- Corbett, A. T. and Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 4(4):253–278.
- Dellaportas, P. and Smith, A. F. (1993). Bayesian inference for generalized linear and proportional hazards models via gibbs sampling. *Applied Statistics*, pages 443–459.

-
- Doignon, J.-P. and Falmagne, J.-C. (2012). *Knowledge spaces*. Springer Science & Business Media.
- Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, 92(6):1087.
- Ephraim, Y. and Merhav, N. (2002). Hidden markov processes. *IEEE Transactions on information theory*, 48(6):1518–1569.
- Ericsson, A. (2016). *Peak: Secrets from the New Science of Expertise*. Random House.
- Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363.
- Faus, M. (2015). Systems and methods for split testing educational videos. US Patent 20,150,310,753.
- Feng, M., Heffernan, N., and Koedinger, K. (2009). Addressing the assessment challenge with an online system that tutors as it assesses. *User Modeling and User-Adapted Interaction*, 19(3):243–266.
- Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive medicine*, 15(5):451–474.
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., Mościcki, E. K., Schinke, S., Valentine, J. C., and Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention science*, 6(3):151–175.
- Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for gibbs sampling. *Applied Statistics*, pages 337–348.
- Glasgow, R. E., Lichtenstein, E., and Marcus, A. C. (2003). Why don’t we see more translation of health promotion research to practice? rethinking the efficacy-to-effectiveness transition. *American journal of public health*, 93(8):1261–1267.
- Ham, J. C. and LaLonde, R. J. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica: Journal of the Econometric Society*, pages 175–205.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.
- Heckman, J. and Singer, B. (1984). The identifiability of the proportional hazard model. *The Review of Economic Studies*, 51(2):231–241.
- Holroyd, C. B. and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, 109(4):679.

-
- Ito, H., Amari, S.-I., and Kobayashi, K. (1992). Identifiability of hidden markov information sources and their minimum degrees of freedom. *IEEE transactions on information theory*, 38(2):324–333.
- Koedinger, K. R. and Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19(3):239–264.
- Koedinger, K. R. and Corbett, A. (2006). Cognitive tutors. *The Cambridge handbook of the learning sciences*, pages 61–77.
- Lee, J. I. and Brunskill, E. (2012). The impact on individualizing student models on necessary practice opportunities. *International Educational Data Mining Society*.
- Leroux, B. G. (1992). Maximum-likelihood estimation for hidden markov models. *Stochastic processes and their applications*, 40(1):127–143.
- Levitt, S., List, J., Neckermann, S., and Sadoff, S. (2013). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *Framed Field Experiments*, 379.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.
- Metcalfe, R., Burgess, S., Proud, S., et al. (2011). *Student effort and educational attainment: Using the England football team to identify the education production function*. Centre for Market and Public Organisation Bristol.
- Murray, R. C., Ritter, S., Nixon, T., Schwiebert, R., Hausmann, R. G., Towle, B., Fancsali, S. E., and Vuong, A. (2013). Revealing the learning in learning curves. In *International Conference on Artificial Intelligence in Education*, pages 473–482. Springer.
- Pardos, Z. and Heffernan, N. (2010a). Navigating the parameter space of bayesian knowledge tracing models: Visualizations of the convergence of the expectation maximization algorithm. In *Educational Data Mining 2010*.
- Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., and Gowda, S. M. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 117–124. ACM.
- Pardos, Z. A. and Heffernan, N. T. (2010b). Modeling individualization in a bayesian networks implementation of knowledge tracing. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 255–266. Springer.
- Pavlik Jr, P. I., Cen, H., and Koedinger, K. R. (2009). Performance factors analysis—a new alternative to knowledge tracing. *Online Submission*.
- Pelánek, R., Rihák, J., and Papoušek, J. (2016). Impact of data collection on interpretation and evaluation of student models. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pages 40–47. ACM.

-
- Rai, D., Gong, Y., and Beck, J. (2009). Using dirichlet priors to improve model parameter plausibility. In *Educational Data Mining 2009*.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. ERIC.
- Ritter, S., Anderson, J. R., Koedinger, K. R., and Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic bulletin & review*, 14(2):249–255.
- Scott, S. L. (2002). Bayesian methods for hidden markov models: Recursive computing in the 21st century. *Journal of the American Statistical Association*.
- Taylor, E. S. (2015). New technology and teacher productivity.
- van De Sande, B. (2013). Properties of the bayesian knowledge tracing model. *JEDM-Journal of Educational Data Mining*, 5(2):1–10.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, 23(3):34–41.
- Wixon, M., d Baker, R. S., Gobert, J. D., Ocumpaugh, J., and Bachmann, M. (2012). Wtf? detecting students who are conducting inquiry without thinking fastidiously. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 286–296. Springer.
- Yudelson, M. V., Koedinger, K. R., and Gordon, G. J. (2013). Individualized bayesian knowledge tracing models. In *International Conference on Artificial Intelligence in Education*, pages 171–180. Springer.
- Zamarro, G., Cheng, A., Shakeel, M., and Hitt, C. (2016). Comparing and validating measures of character skills: Findings from a nationally representative sample (edre working paper 2016-08). *Fayetteville, AR: Department of Education Reform, University of Arkansas*.