Control Theory Forecasts of Optimal Training Dosage to Facilitate Children's Arithmetic Learning in a Digital Educational Application

Sy-Miin Chow[a], Jungmin Lee[a], Abe D. Hofman[b], Han L. J. van der Maas[b], Dennis K. Pearl[a], & Peter C. M. Molenaar[a]

[a]Pennsylvania State University [b] University of Amsterdam

Citation:

Chow, S-M., Lee, J, Hofman, A., van der Maas, H. L. J., Pearl, D. K., & Molenaar, P. C. M. (2022). Control theory forecasts of optimal training dosage to facilitate children's arithmetic learning in a digital educational application. Special Issue on "Models for Intensive Longitudinal Data and Forecasting," *Psychometrika*. DOI: 10.1007/s11336-021-09829-3

Abstract

Education can be viewed as a control theory problem in which students seek ongoing exogenous input — either through traditional classroom teaching or other alternative training resources — to minimize the discrepancies between their actual and target (reference) performance levels. Using illustrative data from $n = 784$ Dutch elementary school students as measured using the Math Garden, a web-based computer adaptive practice and monitoring system, we simulate and evaluate the outcomes of using off-line and finite memory linear quadratic controllers with constraints to forecast students' optimal training durations. By integrating population standards with each student's own latent change information, we demonstrate that adoption of the control theory-guided, person- and time-specific training dosages could yield increased training benefits at reduced costs compared to students' actual observed training durations, and a fixed-duration training scheme. The control theory approach also outperforms a linear scheme that provides training recommendations based on observed scores under noisy and the presence of missing data. Design-related issues such as ways to determine the penalty cost of input administration and the size of the control horizon window are addressed through a series of illustrative and empirically (Math Garden) motivated simulations.

Control Theory Forecasts of Optimal Training Dosage to Facilitate Children's Arithmetic Learning in a Digital Educational Application

Mastery of arithmetic competence is a dynamic process that reflects the integration of idiographic (i.e., individual-specific) learning characteristics and nomothetic (general) educational practices that target the students at large (Biddlecomb, 2002; Dowker, 2015; Hackenberg & Lee, 2015; Lo & Watanabe, 1997). The advent of modern technology in recent years has led to increased tendency for schools to use digital educational applications (apps), particularly in the area of arithmetic training, to supplement traditional classroom teaching (e.g., Khan Academy Academy, 2017). Despite the appeal of digital educational apps in personalizing learning pace, there are other ways in which their designs can, and should be further optimized. For instance, most educational apps provide "one-size-fits-all" guidelines (e.g., practice arithmetic problems for 15 minutes a week) that assume that students are homogeneous, and any idiosyncratic differences or within-student variations in learning over time can be safely ignored. In reality, such guidelines are far from adequate (Rose, 2016), and often result in missed opportunities to forecast and deliver training when improvements are most needed.

Challenges in Traditional Arithmetic Training

Researchers and educators have long been interested in understanding children's arithmetic competencies from preschool years through higher education. Compared with arithmetic operations such as addition and subtraction, the arithmetic operation of division is particularly challenging as students advance to higher grade levels because division requires understanding and use of a sequence of operations that even teachers disagree on what constitutes the best principles (Hadass & Bransky, 1991). For example, solving long division problems (e.g. 200 / 3 = ?) requires students to recursively carry out a sequence of computations such as "divide–multiply–subtract–bring it down" and they frequently show systematic errors due to misunderstanding of the place value system (Lee, 2007). Solving division with remainder word problems is another type of problem that students

often find difficult (Li & Silver, 2000) because mastery requires correct interpretations of the questions as well as successful performance of the computational procedures.

Mulligan and Mitchelmore (1997) found that students in the second and third grades typically use four main strategies in performing division tasks: direct counting, repeated addition, repeated subtraction, and multiplicative operation. Their study suggested that the strategy of choice varies by developmental stages and reflects the mathematical structure imposed by the student on the problem at hand. Indeed, increasing evidence has suggested that some of the current hurdles in mathematical education stem from individual learning gaps and misconceptions that are difficult to correct in large-group settings. Graeber and Tirosh (1990) showed that misconceptions about multiplication and division were pervasive among fourth and fifth graders, one example of which was the misconception that division always makes something smaller. Such misconceptions interfere with accurate understanding of multiplication and division by decimals. Other researchers concurred with the ubiquitous but idiosyncratic nature of such misconceptions (Biddlecomb, 2002; Hackenberg & Lee, 2015; Lo & Watanabe, 1997), and called for early intervention (e.g., in third grade; Blanton et al., 2015) to help students understand mathematical equivalence and arithmetic generalization.

The critical "take-home message" from these studies is that arithmetic — and specifically, division — training often spans multiple grade levels. Early identification of training deficiencies can yield tremendous returns in the long run, but requires ongoing monitoring and tailored training that may not be feasible in real-world educational settings. In the present article, we demonstrate, using a series of illustrative and empirically motivated simulations, that idiographic and nomothetic learning information can be integrated and used within a control theory approach to forecast the optimal "training dosages" for arithmetic training in school-aged students. In particular, the current work is the first at adapting constrained controllers that are typically designed to drive systems toward stationarity or stability (i.e., time-invariant means, variances, and covariances;

Lütkepohl, 2005) to an educational context in which the target itself is time-varying (e.g., the expectations for children's arithmetic performance naturally differ across grades). Even though a group-based model is used as the operating model to circumvent estimation issues due to the finite number of time points available from each individual, some individual differences are built into the model by including selected individual-specific parameters as latent variables. The person- and time-specific model-implied trajectories are integrated, in turn, with population norms to define the target levels toward which individuals' performance is driven. Design-related considerations and adaptations made to the proposed controllers for the educational application at hand are discussed and demonstrated via a series of illustrative and empirically motivated illustrations.

Education as a Control Theory Problem

In engineering, control theory is routinely used to steer a system to stay as close as possible to a desired reference state (Åström & Murray, 2008; Bellman, 1964; Goodwin, Seron, & de Don, 2005; Kwon & Han, 2005; Liu, Wang, Golnaraghi, & Kubica, 2010; Wang et al., 2014), an application of which is the cruise control system of a car. In this case, the car with the cruise control is the "system"; the controller (the cruise control) determines the external input — namely, the engine's throttle position, which governs the power delivered by the engine to minimize the car's deviations from the desired (reference or target) speed. In a similar vein, education can be viewed as a control theory problem in which students seek ongoing input, such as training in the forms of classes and online resources, to minimize the discrepancies between their actual and target performance levels (Savi, van der Maas, & Maris, 2015). Thus, the optimal amounts of training dosages can be deduced and tailored to individuals' performance gaps in times when such training is most needed.

With few exceptions, most applications of control theory principles in the social, behavioral and health sciences have been limited thus far to theoretical conceptualizations (Carver & Scheier, 1982). The limited exceptions include the work of Wang et al. (2014), who designed and implemented real-world control theory applications to compute the

optimal insulin input to control the glucose levels of diabetic patients. Molenaar (2010) demonstrated via numerical simulations the plausibility of using control theory computation to optimize psychotherapy dosages to maintain desired levels of treatment effectiveness. Rivera, Pew, and Collins (2007) utilized computer simulation to investigate the anticipated impact of intervention design choices in developing adaptive interventions. Still, no studies to date have shown the utility of using such control theory principles with real-world education data, the forecasting of which requires successful integration of population- and group-level norms with individual learning characteristics, strengths, limitations, as well as practical constraints.

The control theory approach and associated novel adaptions proposed in this article were motivated by the need to optimize the arithmetic performance of $n = 784$ Dutch elementary school students as measured using the Math Garden (Klinkenberg, Straatemeier, & van der Maas, 2011), a web-based computer adaptive practice and monitoring system available at https://www.prowise.com/en/learn (or its original Dutch version called Rekentuin.nl). Math Garden was designed to bolster mathematical training in elementary education by providing more time for students to practice and maintain basic mathematical skills, and a more efficient and effective way of measuring as well as using the measurement results to improve the ability of individual students in educational settings. Student data from answering one particular type of arithmetic problem, namely, division, are used. Training dosage is operationalized in the context of the Math Garden data as each student's weekly activity time on the website.

Through a series of simulations, we demonstrate and evaluate the efficacy of the proposed control theory approaches in forecasting the optimal weekly training durations that most efficiently reduce the ongoing discrepancies between individuals' current latent and target performance levels. Such control theory-based approaches utilize a state-space model consisting of a measurement model and a dynamic model of latent variables to guide the decisions on training dosages. Thus, compared to an "observed variable" approach of

recommending a particular amount of increment in training duration as proportionate to every unit of under-performance in observed score, the dynamic model and associated estimation procedures offer recommendations in anticipation of some of the changes that may unfold at the latent level even and especially on occasions when the observed scores are missing. In addition, we propose several novel adaptions to standard control theory procedures by: (1) demonstrating the need and a possible way to incorporate both population norms and individual change information in constructing person-specific target levels in the Math Garden application; (2) proposing and investigating ways to improve ongoing estimation of individuals' latent performance levels based on all the data collected within a future moving window; and (3) presenting a way to quantify the costs and benefits associated with alternative training schemes. Insights and recommendations on ways to enhance future versions of Math Garden, and possible future adaptations of the proposed controllers to better tailor to the needs of educational and other applications in the social and behavioral sciences are discussed.

Modeling Framework

State-Space Model

The state-space model (Durbin & Koopman, 2001; Harvey, 2001; Shumway, 2000) is a longitudinal model formulated at discrete, equally spaced time intervals that consists of a measurement and a dynamic model. The measurement model serves to relate the observed variables to the latent variables (also known as "state variables"), and may take the form, for example, of a factor analytic model. The dynamic model is used to delineate the evolution of the latent variables over time as related to values of the latent variables at previous time points.

The particular form of state-space model considered in this paper is a linear discrete-time time-invariant model. That is, the model is characterized by time increments in the form of integers (e.g., t, $t+1$, $t+2$), has measurement and dynamic functions that are linear in form, and consists of person- and time-invariant parameters. The

measurement and dynamic models for the linear state-space model considered are expressed respectively as:

$$\boldsymbol{y}_{it} = \boldsymbol{\Lambda}\boldsymbol{\eta}_{it} + \boldsymbol{\epsilon}_{it}, \quad \boldsymbol{\epsilon}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_{\epsilon}), \tag{1}$$

$$\boldsymbol{\eta}_{it} = \mathbf{B}\boldsymbol{\eta}_{i,t-1} + \mathbf{G}\boldsymbol{u}_{i,t-1} + \boldsymbol{\zeta}_{it}, \quad \boldsymbol{\zeta}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_{\zeta}), \quad \boldsymbol{\eta}_{i1} \sim \mathcal{N}(\boldsymbol{a}_0, \mathbf{P}_0), \tag{2}$$

where $\boldsymbol{y}_{it}$ is a $p \times 1$ vector of observed variables for individual i at time t ($t = 1, \ldots, T$; $i = 1, \ldots n$); $\boldsymbol{\eta}_{it}$ is a $w \times 1$ vector of latent variables, also known as *state* variables; $\boldsymbol{\Lambda}$ is a matrix of factor loadings, $\boldsymbol{\epsilon}_{it}$ is a vector of measurement errors; $\boldsymbol{\zeta}_{it}$ is a vector of process noises or disturbances; and $\mathbf{B}$ is a $w \times w$ matrix of regression effects among the latent variables. $\boldsymbol{u}_{i,t-1}$ is a $r \times 1$ vector of exogenous input or predictor variables at time t-1 that affect the latent variables through the matrix of regression coefficients, $\mathbf{G}$. Critically, the values of these exogenous variables are what we seek to manipulate —or control — to drive values of the latent variables to a desired range. One example of such controllable input variables is shown in Wang et al. (2014), in which control theory algorithm was used to determine the optimal amount of insulin to be administered to diabetic patients to minimize deviations of the patients' glucose levels from a desired target level.

State-space models have been compared to structural equation models, and their equivalence has been established in cases involving cross-sectional models with $T = 1$, and panel data extensions in which T is small relative to n, and special constraints have been imposed to ensure equivalence in the initial distribution of the latent variables at time 1 when data just become available (Chow, Ho, Hamaker, & Dolan, 2010).

Bivariate Dual Change Score Model with Exogenous Input (BDCM-X)

In the current study, we examine the coupling relations between individuals' weekly latent ability and reaction time (RT) under the influence of an exogenous input variable, namely, weekly training duration. A special case of the state-space model, the bivariate dual change score model with exogenous variables (BDCM-X), was used. Univariate and bivariate versions of the dual change model with no exogenous input were proposed

originally by McArdle and Hamagami (2001) to represent latent processes that unfold following sigmoid-shaped curves — consonant with the general time trends observed in the Math Garden data. Variations of the BDCM with no exogenous variable have also been employed in similar contexts to represent students' arithmetic growth (e.g., Chow, Grimm, Guillaume, Dolan, & McArdle, 2013).

The vector of observed variables, $\boldsymbol{y}_{it}$, consists in the present example of two observed indicators, $y_{1,it}$ and $y_{2,it}$, corresponding, respectively, to individual i's average estimated division ability at week t obtained from the Math Garden, and the corresponding average reaction time across all the division items attempted that week. They are used to identify $\boldsymbol{\eta}^1_{it} = [\eta_{1,it}\ \eta_{2,it}]'$, two latent variables that represent individual i's underlying true division ability and reaction time, respectively, that are separated from their measurement error counterparts (McArdle & Hamagami, 2001). Even though the "observed" ability scores available from Math Garden directly are, in a sense, latent ability estimates produced by the Elo system (Klinkenberg et al., 2011), these Elo estimates may still include sources of occasion- and person-specific variability that would be regarded as "measurement errors" in the classical test theory sense (Lord & Novick, 1968). The latent ability included in the vector $\boldsymbol{\eta}^1_{it}$, in contrast, refers to the latent (i.e., measurement error-free) portion of the Elo scores that shows systematic interindividual differences in dynamics over time as captured by the BDCM-X model.

The BDCM-X essentially posits that the latent changes (i.e., changes free of measurement errors) in the two latent processes of interest depend on the levels of these processes at time t-1, their respective latent person-specific intercepts, a_{1i} and a_{2i}, and a vector of lag-1 exogenous input variables at time t, $\boldsymbol{u}_{i,t-1}$, as:

$$\Delta\boldsymbol{\eta}^1_{it} = \boldsymbol{a}_i + \mathbf{B}^1\boldsymbol{\eta}^1_{i,t-1} + \mathbf{G}^1\boldsymbol{u}_{i,t-1} + \boldsymbol{\zeta}^1_{it} \tag{3}$$

where $\Delta[.]$ on the left-hand-side of the equation represents latent changes in the values of the components enclosed in brackets over one time unit, or in Equation 3, $\Delta\boldsymbol{\eta}^1_{it} =$

$\boldsymbol{\eta}^1_{it} - \boldsymbol{\eta}^1_{i,t-1}$. $\mathbf{G}^1$ is a matrix of regression coefficients relating the latent ability and reaction time variables to the exogenous input variables. $\mathbf{B}^1$ is a matrix of coefficients relating the latent variables at time t to their values at time t-1. In particular,

$$\mathbf{B}^1 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

The diagonal entries, b_{11} and b_{22}, represent the auto-proportion effects of the processes from time $t-1$ on time t. Adding a unity constant to the auto-proportion parameters b_{11} and b_{22} gives rise to what is typically referred to as the autoregression parameters (Chow et al., 2010), for reason that will become clear shortly. The reciprocal coupling effects between the latent processes are captured by b_{12} and b_{21}, including influence in the direction from $\eta_{2,it}$ to $\eta_{1,it}$, and from $\eta_{1,it}$ to $\eta_{2,it}$, respectively.

The person-specific intercepts impose a constant amount of change on each process's latent change at each time point. When $\mathbf{B}^1 = \mathbf{G}^1 = \mathbf{0}$, a_{1i} and a_{2i} may be conceived as individual-specific linear slopes. To allow these constant slope terms to have a random component, a_{1i} and a_{2i} have to be included as part of a larger latent variable vector, $\boldsymbol{\eta}_{it} = [\boldsymbol{\eta}^{1\prime}_{it}\ a_{1it}\ a_{2it}\ ]'$, and explicitly constrained to be invariant over time, or in other words, showing *no latent changes*. In other words, $\boldsymbol{\eta}^1_{it}$ is only a subvector of the full latent variable vector, $\boldsymbol{\eta}_{it}$, as the latter also contains the person-specific intercepts, $\boldsymbol{a}_i = [a_{1i}\ a_{2i}]'$, as additional latent variables. In addition, using $u_{i,t-1}$ to denote an individual's Math Garden training duration in the previous week as the sole exogenous input variable in our application, we obtain a dynamic model of the form:

$$\Delta \begin{bmatrix} \eta_{1it} \\ \eta_{2it} \\ a_{1it} \\ a_{2it} \end{bmatrix} \triangleq \boldsymbol{\eta}_{it} - \boldsymbol{\eta}_{i,t-1} = \begin{bmatrix} b_{11} & b_{12} & 1 & 0 \\ b_{21} & b_{22} & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{1i,t-1} \\ \eta_{2i,t-1} \\ a_{1i,t-1} \\ a_{2i,t-1} \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ 0 \\ 0 \end{bmatrix} [u_{i,t-1}] + \begin{bmatrix} \zeta_{1it} \\ \zeta_{2it} \\ 0 \\ 0 \end{bmatrix}. \qquad (4)$$

Then, substituting Equation 4 into $\boldsymbol{\eta}_{it} = \boldsymbol{\eta}_{i,t-1} + \Delta\boldsymbol{\eta}_{it}$ yields:

$$\begin{aligned} \boldsymbol{\eta}_{it} &= (\mathbf{I} + \mathbf{B})\,\boldsymbol{\eta}_{i,t-1} + \mathbf{G}u_{i,t-1} + \boldsymbol{\zeta}_{it}, \qquad \boldsymbol{\zeta}_{it} \sim N\Big(\mathbf{0}, \mathrm{Cov}(\boldsymbol{\zeta}_{it})\Big) \\ &= \left(\mathbf{I} + \begin{bmatrix} b_{11} & b_{12} & 1 & 0 \\ b_{21} & b_{22} & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) \boldsymbol{\eta}_{i,t-1} + \begin{bmatrix} g_1 \\ g_2 \\ 0 \\ 0 \end{bmatrix} u_{i,t-1} + \begin{bmatrix} \zeta_{1it} \\ \zeta_{2it} \\ 0 \\ 0 \end{bmatrix}, \end{aligned} \tag{5}$$

that is, a dynamic model of the form of Equation 2 with:

$$\mathbf{B} = \begin{bmatrix} 1 + b_{11} & b_{12} & 1 & 0 \\ b_{21} & 1 + b_{22} & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; \mathbf{G} = \begin{bmatrix} g_1 \\ g_2 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and } \boldsymbol{\zeta}_{it} = \begin{bmatrix} \zeta_{1it} \\ \zeta_{2it} \\ 0 \\ 0 \end{bmatrix}, \tag{6}$$

where we assume that $\boldsymbol{\zeta}_{it}$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix, $\mathrm{Cov}(\boldsymbol{\zeta}_{it}) = \mathrm{diag}[\psi_{11}\ \psi_{22}\ c\ c]$, where c is a small constant to ensure the positive definiteness of this covariance matrix. Due to the explicit upper time limit Math Garden allows for each division item (20 seconds), and the inherent limits in individuals' learning capacity, we expect b_{11} and b_{22} to be negative. That is, higher levels on these constructs are expected to yield reduced latent changes, or specifically, latent growth, thereby driving these constructs toward their asymptotes, which are made person-specific by the constant slopes, a_{1i} and a_{2i}. In addition, the signs of the coupling (or cross-regression) parameters, b_{12} and b_{21}, can help shed light on the interplay between an individual's latent ability and reaction time as the individual acquires and further solidifies his/her division skills over time.

As alluded to earlier, the BDCM-X is a special extension of the vector autoregressive model with exogenous variables (VAR-X). Equation (6) highlights that after the algebraic re-arrangements of the terms from Equation (4) to Equation (6), the lag-1 autoregression parameters for latent ability and reaction time are given respectively by $1 + b_{11}$ and $1 + b_{22}$,

with cross-regression parameters, b_{12} and b_{21} (Chow et al., 2010). Convergence of the two latent processes toward stable equilibrium levels requires that b_{11}, b_{22}, b_{12}, and b_{21} take on values that render the VAR portion (namely, $\boldsymbol{\eta}_{it}$ when $a_{1i} = a_{2i} = 0$ for all participants) of the model stable or stationary (e.g., does not show changes in means or variances over time; Hamilton, 1994; Lütkepohl, 2005). Still, even if the auto- and cross-regression parameters fall within this stationary range, the BDCM-X as a whole would still be non-stationary in most instances because the constant slopes, a_{1i} and a_{2i}, would typically lead to over-time changes in means, $E(\boldsymbol{\eta}_{it})$ and $E(\boldsymbol{y}_{it})$, thus violating the stationarity assumptions.

As with all sequentially dependent longitudinal processes, the latent processes have to be "started up" at time $t = 1$. The initial values of these latent variables, commonly known as the initial conditions of the latent processes, are modeled as:

$$\begin{bmatrix} \eta_{1i1} \\ \eta_{2i1} \\ a_{1i} \\ a_{2i} \end{bmatrix} = \begin{bmatrix} \mu_{\eta_1} \\ \mu_{\eta_2} \\ \mu_{a_1} \\ \mu_{a_2} \end{bmatrix} + \begin{bmatrix} v_{\eta_{1,i}} \\ v_{\eta_{2,i}} \\ v_{a_{1,i}} \\ v_{a_{2,i}} \end{bmatrix}, \quad \begin{bmatrix} v_{\eta_{1,i}} \\ v_{\eta_{1,i}} \\ v_{a_{1,i}} \\ v_{a_{2,i}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{v_{\eta_1}} & & & \\ \sigma_{v_{\eta_1},v_{\eta_2}} & \sigma^2_{v_{\eta_2}} & & \\ \sigma_{v_{\eta_1},v_{a_1}} & \sigma_{v_{\eta_2},v_{a_1}} & \sigma^2_{v_{a_1}} & \\ \sigma_{v_{\eta_1},v_{a_2}} & \sigma_{v_{\eta a},v_{\eta_2}} & \sigma_{v_{a_1},v_{a_2}} & \sigma^2_{v_{\eta_2}} \end{bmatrix} \right). \tag{7}$$

Here, the initial levels of the two latent processes, η_{1i1} and η_{2i1}, are composed of group average initial levels, μ_{η_1} and μ_{η_2}, and person-specific deviations from them, $v_{\eta_{1,i}}$ and $v_{\eta_{2,i}}$. The model also allows the person-specific intercepts, a_{1i} and a_{2i}, to be a function of the group average slopes, μ_{a_1} and μ_{a_2}, along with person-specific deviations, $v_{a_{1i}}$ and $v_{a_{2i}}$.

The true scores of division ability and reaction time are, in turn, linked to individual i's "observed" Elo division ability score and reaction time at time t, denoted respectively as y_{1it} and y_{2it}, as:

$$\begin{bmatrix} y_{1it} \\ y_{2it} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{1it} \\ \eta_{2it} \\ a_{1it} \\ a_{2it} \end{bmatrix} + \begin{bmatrix} \epsilon_{1it} \\ \epsilon_{2it} \end{bmatrix}, \quad \begin{bmatrix} \epsilon_{1it} \\ \epsilon_{2it} \end{bmatrix} \sim N\Big(\mathbf{0}, \text{diag}\begin{bmatrix} \sigma^2_{\epsilon_1} & \sigma^2_{\epsilon_2} \end{bmatrix}\Big). \tag{8}$$

where $\sigma^2_{\epsilon_1}$ and $\sigma^2_{\epsilon_2}$ are the measurement error variances associated with the observed ability

scores and reaction time.

To summarize, the BDCM-X used as the motivating model throughout this paper is a time-invariant group-based model. That is, most of the parameters in the model, except for the person-specific initial levels and constant slopes, are held constant across individuals and time. Making this assumption allows researchers to circumvent estimation issues in situations involving finite time length from each individual, and pre-estimate modeling parameters using a different sample prior to application of the controller to a new, validation sample. As shown in our illustrations, model-implied trajectories from the BDCM-X, which are time- and person-specific, can be combined with population norms to define the target functions for control purposes.

Estimation Algorithm

Latent Variable and Parameter Estimation via the Kalman Filter (KF)

Before any control theory algorithm can be applied, at least two elements have to first be estimated: the unknown modeling parameters (e.g., b_{11}—b_{22}), and the latent variable values, which are needed for computation of the control input. One well-known approach for accomplishing these purposes is to use the Kalman filter (KF). The KF estimates values of the current or future latent variables values (e.g., factor scores) given manifest data up to the current time point by minimizing prediction errors in the least squares sense (Zarchan & Musoff, 2000). By-products from performing the KF can be substituted into a log-likelihood function, which has known analytic form in the linear, normal special cases shown in Equations 1—2, and optimized, for example, via Newton-Raphson procedures to obtain estimates of the unknown modeling parameters (Chow et al., 2010; Shumway & Stoffer, 2006). This requires iterative calls of the KF for repeated evaluations of the log-likelihood function at different parameter values.

Some definitions of notation and key concepts are in order. Let $\mathbf{Y}_{i,1:k} \triangleq \{\boldsymbol{y}_{ij}, \boldsymbol{u}_{ij}; j = 1, \ldots, k\}$ denote the array of manifest observations, including exogenous input, available from time 1 to time k. Three types of state estimates and corresponding

covariance matrix (for quantifying uncertainty associated with the state estimates) are usually of interest:

1. The one-step-ahead *predicted* or *forecast* state values, $\boldsymbol{\eta}_{it|t-1} \triangleq E(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:t-1})$, and the associated covariance matrix, $\mathbf{P}_{it|t-1} \triangleq \text{Cov}(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:t-1})$, estimated using observations up to time $t-1$;

2. The *filtered* state values, $\boldsymbol{\eta}_{it|t} \triangleq E(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:t})$, and the associated covariance matrix, $\mathbf{P}_{it|t} \triangleq \text{Cov}(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:t})$, estimated using observations up to time t;

3. The smoothed state values, $\boldsymbol{\eta}_{it|T} \triangleq E(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:T})$, and the associated covariance matrix, $\mathbf{P}_{it|T} \triangleq \text{Cov}(\boldsymbol{\eta}_{it}|\mathbf{Y}_{i,1:T})$, estimated using observations up to time T, where T may correspond to the last time point of the data, or any later time point beyond t.

To compute optimal control inputs in the present study, we are interested in computing — or specifically, forecasting — estimates of $\boldsymbol{u}_{it}$ given $\mathbf{Y}_{i,1:t-1}$. Doing so requires use of all three sets of the state estimates summarized above. The procedures for doing so are outlined next.

With some initial guesses of the parameter values, and setting the initial conditions of the state estimates as: $\boldsymbol{\eta}_{i1|0} = \boldsymbol{a}_0$, and $\mathbf{P}_{i1|0} = \mathbf{P}_0$ (for alternative specifications of these initial conditions, see also Harvey (2001), and Zarchan and Musoff (2000)), the KF essentially involves sequentially going through a prediction and a filtering phase from time $t = 1, \ldots, T$ to obtain the predicted and filtered state estimates, respectively. During the prediction phase, we obtain:

$$\begin{aligned} \boldsymbol{\eta}_{it|t-1} &= \mathbf{B}\boldsymbol{\eta}_{i,t-1|t-1} + \boldsymbol{G}\boldsymbol{u}_{i,t-1}, \text{ and} \\ \mathbf{P}_{it|t-1} &= \mathbf{B}\mathbf{P}_{i,t-1|t-1}\mathbf{B}' + \boldsymbol{\Psi}_{\zeta}. \end{aligned} \tag{9}$$

This is followed by the filtering phase, from which we obtain:

$$\begin{aligned} \boldsymbol{\eta}_{it|t} &= \boldsymbol{\eta}_{it|t-1} + \mathbf{K}_{it}\left(\boldsymbol{y}_{it} - \boldsymbol{\Lambda}\boldsymbol{\eta}_{it|t-1}\right), \\ \mathbf{P}_{it|t} &= \left(\mathbf{I} - \mathbf{K}_{it}\boldsymbol{\Lambda}\right)\mathbf{P}_{it|t-1}^{-1}, \end{aligned} \tag{10}$$

where $\boldsymbol{K}_{it} = \mathbf{P}_{it|t-1}\boldsymbol{\Lambda}' \left(\boldsymbol{\Lambda}\mathbf{P}_{it|t-1}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}_{\epsilon}\right)^{-1}$, usually known as the Kalman gain matrix, determines how heavily the discrepancies between the predicted and actual measurements are weighted in updating filtered estimates. It may vary in dimension for each individual at each time point to accommodate partial missingness in some observed variables at particular time points.

Finally, when $\boldsymbol{\eta}_{it|t}$ and $\mathbf{P}_{it|t}$ are available from time 1 through T, the *smoothed* state estimates and the associated smoothed stated covariance matrix can be computed backward in time starting from time $t = T, \ldots, 1$ as:

$$\begin{aligned} \boldsymbol{\eta}_{it|T} &= \boldsymbol{\eta}_{it|t} + \tilde{\mathbf{P}}_{it}\left(\boldsymbol{\eta}_{i,t+1|T} - \boldsymbol{\eta}_{i,t+1|t}\right) \\ \mathbf{P}_{it|T} &= \mathbf{P}_{it|t} + \tilde{\mathbf{P}}_{it}\left(\mathbf{P}_{i,t+1|T} - \mathbf{P}_{i,t+1|t}\right)\tilde{\mathbf{P}}_{it}' \end{aligned} \tag{11}$$

where $\tilde{\mathbf{P}}_{it} = \mathbf{P}_{it|t}\mathbf{B}'\left(\mathbf{P}_{i,t+1|t}\right)^{-1}$. This process is one example of a Kalman smoother (KS) known as the fixed interval smoother (Shumway & Stoffer, 2006; Zarchan & Musoff, 2000). In most circumstances, particularly when process noises are present in the system, smoothed estimates provide more accurate estimates of the system's latent variable values than the filtered estimates because smoothing draws on information from more observations.

Parameter Estimation by Prediction Error Decomposition (PED)

The difference $\boldsymbol{e}_{it} \triangleq \left(\boldsymbol{y}_{it} - \boldsymbol{\Lambda}\boldsymbol{\eta}_{it|t-1}\right)$ shown in Equation 10 is often termed the vector of *innovations*, as it represents the difference between the predicted measurements, $E(\boldsymbol{y}_{it}|\mathbf{Y}_{i,1:t-1}) = \boldsymbol{\Lambda}\boldsymbol{\eta}_{it|t-1}$, and the actual measurements at time t, $\boldsymbol{y}_{it}$, or in other words, the new information brought in by the observations at time t. This term is also known as the one-step-ahead prediction errors (e.g., Chow et al., 2010; Durbin & Koopman, 2001). These prediction errors, the associated innovation covariance matrix, $\mathbf{F}_{it} \triangleq \mathrm{Cov}(\boldsymbol{e}_{it}) = \boldsymbol{\Lambda}\mathbf{P}_{it|t-1}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}_{\epsilon}$, together with other by-products from the KF procedures, can be substituted into a log likelihood function, then optimized via Newton-Raphson or other similar techniques, which would yield maximum likelihood

estimates of the unknown parameters in $\mathbf{B}$, $\mathbf{G}$, $\mathbf{\Lambda}$, $\mathbf{\Psi}_\zeta$ and $\mathbf{\Psi}_\epsilon$.

The log likelihood function can be written as:

$$LL_{KF}(\theta_k) = \frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(-p\log(2\pi) - log|\mathbf{F}_{it}| - \boldsymbol{e}_{it}'\mathbf{F}_{it}^{-1}\boldsymbol{e}_{it}\right), \quad (12)$$

where p is the number of manifest variables, which may be person-dependent in the presence of missing data. Equation 12 is known as the prediction error decomposition (PED) function, and maximizing this function with respect to the parameters in $\mathbf{B}$, $\mathbf{G}$, $\mathbf{\Lambda}$, $\mathbf{\Psi}_\zeta$ and $\mathbf{\Psi}_\epsilon$ results in maximum likelihood (ML) estimates of these parameters (Harvey, 1989; Ljung & Söderström, 1983; Schweppe, 1965; Shumway & Stoffer, 2006). In addition, when the parameters are constrained to be invariant across persons, the resultant model captures the pooled dynamics in the sample as a whole.

In sum, by first setting the parameters to some fixed initial values, each person's data are subjected to the KF algorithm and individual state estimates are thus obtained (from $i = 1, \ldots n$) by treating the parameters as fixed values. State estimates from these n individuals are then substituted into the PED function, the optimization (with Newton-Raphson procedures) of which generates updated parameter estimates for another iteration of the KF. This entire KF $\leftrightarrow$ PED cycle is repeated iteratively until some convergence criteria are met, at which point the final parameter estimates at convergence provide ML estimates of the parameters. Using the inverse of the negative numerical Hessian of the PED function at the point of convergence as an estimate of the asymptotic covariance matrix of the parameters, we compute standard error estimates as the square roots of the diagonal elements of this covariance matrix. As described elsewhere (Chow & Zhang, 2013; Harvey, 2001), information criterion measures such as the Akaike information criterion (AIC; Akaike, 1973) and Bayesian information criterion (BIC; Schwarz, 1978) can also be computed using the PED for model comparison purposes. These procedures have been implemented in the R package, *dynr* (Ou, Hunter, & Chow, 2019). We extend the functions available from *dynr* to use these KF-related by-products to implement the

constrained control input estimation procedures described next.

Constrained Control Theory Optimization

In the current context, we examine the extent to which students' deviations in ability from their desired levels can be reduced more efficaciously by optimizing, as opposed to dictating by design, the appropriate amount of training "dosage" to which each individual should be exposed at each particular time point. Training dosage, namely, u_{it} as shown in Equation (5), corresponds in the Math Garden example to a learner's weekly activity time using the app. Our goal is to use constrained control and estimation (Goodwin et al., 2005; Molenaar, 2010) to "forecast" the optimal amount of weekly training duration to spend on Math Garden, and contrast the simulated forecast results obtained using these "controlled" as compared to other alternative training schemes.

In practice, this constrained optimization of training duration is implemented as follows. We first fit the BDCM-X using an *estimation sample* to obtain estimated parameters, which are then used to set up a controller to be applied to the learning data of a *validation sample* of new Math Garden users. This controller provides recommendations on optimal training dosage (or duration) for each Math Garden user and each week to tailor to each user's learning efficacy.

Receding Horizon Linear Quadratic Controller (LQC). Working with the general state-space model in Equations (1)—(2), optimal values of $\boldsymbol{u_{it}}$ may be obtained by minimizing a quadratic cost function with respect to $\boldsymbol{u}_{it}$. In most engineering settings, the solution for $\boldsymbol{u}_{it}$ is typically derived recursively (i.e., one t at a time) over a control horizon between time t and $t+h$, where $h>1$ is called the control horizon. The quadratic cost function is defined as

$$\begin{aligned} J_{it}\left(\boldsymbol{\eta}_{i,t+\tau}^{\{\tau=0,\ldots,h\}}, \boldsymbol{\eta}^r, \boldsymbol{u}_{i,t+\tau}\right) &= \sum_{\tau=0}^{h-1}\left[(\boldsymbol{\eta}_{i,t+\tau}-\boldsymbol{\eta}^r)'\mathbf{Q}(\boldsymbol{\eta}_{i,t+\tau}-\boldsymbol{\eta}^r)+\boldsymbol{u}'_{i,t+\tau}\mathbf{R}\boldsymbol{u}_{i,t+\tau}\right]+ \\ &\quad (\boldsymbol{\eta}_{i,t+h}-\boldsymbol{\eta}^r)'\mathbf{Q}_h(\boldsymbol{\eta}_{i,t+h}-\boldsymbol{\eta}^r), \end{aligned} \tag{13}$$

where the desired reference levels of the latent processes are denoted as $\boldsymbol{\eta}^r$. The matrices $\mathbf{Q}$, $\mathbf{Q}_h$, and $\mathbf{R}$ are positive-definite design matrices chosen *a priori* to reflect, respectively, how heavily deviations of the latent processes from their desired levels should be weighted within the control horizon (for τ between 1 and $h-1$), at the end point of the control horizon (i.e., for $\tau = h$), and the costs associated with administration of higher training dosages (i.e., higher values) of $\boldsymbol{u}_{it}$.

Kwon and Han (2005) showed that optimal values of $\boldsymbol{u}_{i,t+\tau}$ that would minimize the cost function shown in Equation (13), denoted herein as $\boldsymbol{u}^*_{i,t+\tau}$, can be computed as:

$$\boldsymbol{u}^*_{i,t+\tau} = -\boldsymbol{\Xi}\Big[\mathbf{L}_{i,t+\tau}\mathbf{B}\boldsymbol{\eta}_{i,t+\tau} + \boldsymbol{g}_{i,t+\tau+1}\Big] \tag{14}$$

where $\boldsymbol{\Xi} = \mathbf{R}^{-1}\mathbf{G}'\Big[\mathbf{I}_w + \mathbf{L}_{i,t+\tau+1}\mathbf{G}\mathbf{R}^{-1}\mathbf{G}'\Big]^{-1}$, where $\mathbf{I}_w$ denotes a w-dimensional identity matrix ; whereas $\mathbf{L}_{i,\cdot}$ and $\boldsymbol{g}_{i,\cdot}$ are obtained, starting from time $t+h$ (i.e., $\tau = h$) as:

$$\begin{aligned} \mathbf{L}_{i,t+h} &= \mathbf{Q}_h \\ \boldsymbol{g}_{i,t+h} &= -\mathbf{Q}_h\boldsymbol{\eta}^r, \end{aligned} \tag{15}$$

and then computed recursively backward in time for $\tau = h-1$ to 0 as:

$$\begin{aligned} \mathbf{L}_{i,t+\tau} &= \mathbf{B}'\mathbf{S}^{-1}\mathbf{L}_{i,t+\tau+1}\mathbf{B} + \mathbf{Q} \\ \boldsymbol{g}_{i,t+\tau} &= \mathbf{B}'\mathbf{S}^{-1}\boldsymbol{g}_{i,t+\tau+1} - \mathbf{Q}\boldsymbol{\eta}^r. \end{aligned} \tag{16}$$

To shed light on what these terms mean, it may be helpful to note that at $t = t+h$, $\boldsymbol{u}^*_{i,t+h} = \boldsymbol{\Xi}\mathbf{Q}_h\Big[\boldsymbol{\eta}^r - \mathbf{B}\boldsymbol{\eta}_{i,t+h}\Big]$. Thus, values of $\boldsymbol{u}^*_{i,t+h}$ are determined as proportionate to the amounts of deviations of the states' projected values, $\mathbf{B}\boldsymbol{\eta}_{i,t+h}$, from their target levels, $\boldsymbol{\eta}^r$. How much these deviations are weighted depends on the state deviation penalty matrix, $\mathbf{Q}_h$, and also the "input gain" matrix, $\boldsymbol{\Xi}$, which compares how important and costly it is to *not* reduce state deviations relative to the cost of administering the input, $\mathbf{R}$. These recursions allow backward propagation of the state values and target levels through $\mathbf{L}_{i,t+\tau}$ and $\boldsymbol{g}_{i,t+\tau}$, respectively. They eventually yield estimates for $\boldsymbol{u}^*_{i,t}$ that, when incurred on the states at time $t+1$, help minimize the states' deviations from their target levels over the

control horizon.

The control theory algorithm summarized in Equations (14)—(16) is one kind of controller known as the receding horizon Linear Quadratic Controller (LQC). It is linear in terms of the underlying state-space model linked to the controller, quadratic in the sense of the quadratic form of the cost function adopted in Equation (13), and the receding horizon refers to the property that the input is iteratively updated with a receding (declining or shrinking window) of future state values. This controller can be regarded as "deterministic" in the sense that it was originally designed for systems in which perfect knowledge of the latent states is available. In the present context, the controller is paired with different state estimators, thus resulting in different variations of the receding horizon LQC. These variations are described next.

Off-line and Finite Memory Linear Quadratic Controllers. We consider three approaches for obtaining state estimates for computation of the control input in Equation (14). The first is an off-line implementation of the fixed interval smoother (hence forth simplified as the Kalman smoother or KS) that provides estimates of the states conditional on the whole collection of time series (from $t = 1, \ldots$ T). This approach is said to be *off-line* because the state estimates are computed for the entire time series after all the data have already been collected, as opposed to *on-line* as the data arrive. Because the effects of the control input on the state processes are not taken into consideration in computing the state estimates, this first approach parallels an open-loop approach (Kuo, 1991). This specific LQC is referred to herein as the off-line LQC with KS state estimates.

The second and third approaches utilize a window of measurements to compute the state estimates in that particular window. Kwon and Han (2005) referred to these variations as Linear Quadratic Finite Memory Controllers. Such estimators are sometimes referred to as moving horizon estimators, and we refer to them herein as finite memory LQCs (FMLQCs). FMLQCs typically use a window of measurements of size n_h prior to the current time t, where n_h is known as the moving horizon window, to compute the state

estimates (Bavdekar, Bhushan Gopaluni, & Shah, 2013; J.B., Mayne, & Diehl, 2017). The second approach we considered combines the LQC algorithm with the KF for estimating the state values up to time t sequentially, based on observed data up to time t. We refer to this as the KF-based FMLQC. The KF-based FMLQC is a closed-loop estimation approach because effects of past control input values (up to time $t-1$) are taken into account in updating the state estimates at time t (Kuo, 1991). However, future effects of the control input on state values beyond time t are not incorporated into the state estimates.

The third approach is designed to apply the KS to a window of observations from time t to $t+h$ in estimating the state values at time t. As such, it uses the same future horizon of estimation window as the computation of the control input $\boldsymbol{u}_{it}^{*}$ in Equation (14). Because current and future measurements ($t+1$, …, $t+h$) are used in updating the state estimates at time t, these smoothed estimates take into consideration the effects of the control input on future state values in computing the control input values at time t. As such, this approach is also a closed-loop estimation approach. In summary, the three approaches we adopted for state estimation give rise to three variations of the LQC: (1) The off-line LQC based on off-line KS state estimates; (2) the FMLQC with KF state estimates administered for $t \in [t-n_h, \ldots, t]$; and (3) the FMLQC with KS state estimates administered for $t \in [t, \ldots, t+h]$.

The quadratic cost function in Equation(13) used in the present context may diverge from the needs of most educational applications in the sense that both positive and negative deviations from the target function are penalized equally. In practice, performing above the target function does not constitute a problem in instructional settings. In fact, over-performing may even be encouraged. We circumvent this limitation by adding constraints to the recommended input values to fall between the lower and upper limits of

$\boldsymbol{u}_{lower,i}$ and $\boldsymbol{u}_{upper,i}$, respectively. That is, we constrain that:

$$\boldsymbol{u}_{it}^{R*} = \begin{cases} \boldsymbol{u}_{it}^{*} & \text{if } \boldsymbol{u}_{lower,i} > \boldsymbol{u}_{it}^{*} > \boldsymbol{u}_{upper,i} \\ \boldsymbol{u}_{upper,i} & \text{if } \boldsymbol{u}_{it}^{*} > \boldsymbol{u}_{upper,i} \\ \boldsymbol{u}_{lower,i} & \text{if } \boldsymbol{u}_{i,t}^{*} < \boldsymbol{u}_{lower,i}. \end{cases} \tag{17}$$

For all our illustrations, we set $\boldsymbol{u}_{lower,i}$ to $\mathbf{0}$. Thus, even though a quadratic cost function is optimized, recommendations to reduce training durations are automatically ignored. Other approaches that utilize alternative (e.g., asymmetric) cost functions to more heavily penalize deviations in one direction than the other are highlighted in the Discussion section but are not considered here.

For design purposes, it is of interest to derive some indices for quantifying the costs and benefits associated with using a set of control inputs. We propose using the cost and benefit functions:

$$\begin{aligned} \text{Cost\_State} &= \sum_{i=1}^{n}\sum_{t=1}^{T-1}\left[(\boldsymbol{\eta}_{i,t} - \boldsymbol{\eta}_{i,t}^{r})'\mathbf{Q}(\boldsymbol{\eta}_{i,t} - \boldsymbol{\eta}_{i,t}^{r})\right] + (\boldsymbol{\eta}_{i,T} - \boldsymbol{\eta}^{r})'\mathbf{Q}_{h}(\boldsymbol{\eta}_{i,T} - \boldsymbol{\eta}^{r}) \\ \text{Cost\_Input} &= \sum_{i=1}^{n}\sum_{t=1}^{T}\boldsymbol{u}_{i,t}'\mathbf{R}\boldsymbol{u}_{i,t}, \\ \text{Relative Benefit} &= \frac{\text{BaseCost\_State} - \text{Cost\_State}}{\text{BaseCost\_State}}, \\ \text{Relative Cost} &= \frac{\text{Cost\_Input} - \text{BaseCost\_Input}}{\text{BaseCost\_Input}}, \end{aligned} \tag{18}$$

where Relative Benefit quantifies the change in the quadratic cost associated with state deviations under the current control input scheme relative to the quadratic state cost in a baseline condition, BaseCost_State, (e.g., when no control input is used). Positive (negative) values represent a reduction (increase) in state deviations from the target trajectory compared to baseline. In a similar vein, Relative Cost quantifies the change (specifically, increase) in the quadratic cost associated with the current control input scheme relative to the quadratic input cost in a baseline condition, BaseCost_Input, (e.g., when input values are not determined by the LQC). Positive (negative) values represent an

increase (decrease) in input costs compared to baseline.

Illustrative Simulations

To demonstrate the effects of the proposed LQCs , we simulated data using a univariate dual change score model, that is, a univariate special case of the BDCM-X in Equations 3—8 in which all terms associated with η_{2it} and y_{2it} were dropped.

We present five simulations designed to demonstrate the effects of the control theory input under: (I) a scenario in which some external shocks were applied to η_{1it} to induce transient effects in lowering individuals' latent ability levels; (II) a scenario in which the external shocks were applied both to η_{1it} and a_{1i}, leading to irreversible reductions in the individuals' asymptotic performance levels; (III) a scenario in which we contrasted the effects of the control input under conditions with static (i.e., person- and time-invariant) as compared to person- and time-specific target trajectory, $\boldsymbol{\eta}^r$; (IV) different choices of penalty weight for the input cost through variations in $\mathbf{R}$, and (V) use of the off-line LQC and other FMLQC variations. We also demonstrate within the context of these illustrations the advantages of the control theory approaches relative to a simpler, observed linear approach that recommends increments in practice duration as proportionate to individuals' *observed* negative deviations in performance scores in the absence of an operating state-space model.

In all five illustrations, we set $T = 20$ time points, $n = 3$ individuals, and adopted the following dynamic and measurement parameter values:

$$b_{11} = -0.20, g_1 = 0.70, \psi_{11} = 0.00, \sigma^2_{\epsilon_1} = 0.5,$$

and the following initial condition-related parameter values:

$$\mu_{\eta_1} = 1.00, \mu_{a_1} = 0.50, \sigma^2_{v_{\eta_1}} = 0.13, \sigma_{v_{\eta_1}, v_{a_1}} = 0.001, \sigma^2_{v_{a_1}} = 0.10.$$

These parameter values were chosen to mirror parameter estimates obtained from previous longitudinal modeling of student arithmetic learning using variations of the BDCM (Chow

et al., 2013). The process noise variance, ψ_{11}, was purposefully set to 0 to yield relatively smooth latent change trajectories that saliently reflect the effects of interest.

For all illustrations, we applied the following constrained control scheme: we set

$$\mathbf{Q} = \mathbf{Q}_h = \mathbf{\Lambda}'\mathbf{\Lambda} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} ,$$

whereas the input cost matrix, which consists of a single scalar, R, was varied by design of the illustrative simulations. The code for all the illustrations is included as supplementary material with this article.

Illustration I: Shocks to Latent Ability

A single exogenous variable, u_{it}, was generated with zero value for all individuals and all time points. Two randomly selected "shock points" were incurred on each individual's time series within the time window $2 <= t <= 10$, with shock magnitudes that were randomly sampled from a uniform distribution in the interval between -2.5 and -0.5. This demonstration is intended to mirror the arithmetic learning trajectories of students who display decrements in performance on two occasions due to unforeseen circumstances (e.g., due to illness or social distress), followed by gradual recovery from the shocks over time to return to their otherwise relatively smooth learning trajectories. Imagine now that it is possible to deliver u_{it} to boost the students' training via an app such as the Math Garden. As opposed to devising a "one-size-fits-all" booster training scheme to each individual, we illustrate the outcomes of using the constrained control input, u^*_{it}, in accelerating the students' return to their original learning trajectories.

We applied the off-line LQC in this illustration with two choices of control horizon window, $h = 4$ and 10. We set R to 10, namely, a relatively large value compared to the values in $\mathbf{Q}$, the cost matrix for deviations of state values. This choice was specifically made to emulate the scenario where administration of higher training dosage would be costly, as dictated by having an R value that was much higher than those in $\mathbf{Q}$ and $\mathbf{Q}_h$. Additionally, we set $\boldsymbol{\eta}^r$ to be time-varying and person-specific, with the value of $\boldsymbol{\eta}^r_{it}$ for

each person and time point set to be the person's predicted curve based on the univariate dual change score model. In other words, the reference levels of the latent processes are themselves dynamic, and are taken to be the individuals' predicted learning trajectories in the absence of any shocks or disturbances, namely,

$$\boldsymbol{\eta}_{it}^{r} = \mathbf{B}\boldsymbol{\eta}_{i,t-1}$$

We further assumed that all parameters were known and fixed at their true values. Furthermore, we imposed lower and upper limits on the permissible values of u_{it}^{R*} such that $0 \leq u_{it}^{*} \leq 2$. The upper limit was deliberately set to be low to constrain the effects of the control input to be small.

The true latent trajectories of one out of the three hypothetical individuals (ID 1) generated using u_{it}^{R*} and $u_{it} = 0$ are plotted in Figure 1, in the top two panels. It can be seen that in this illustration, immediately following the shock points, the individual's latent ability level, η_{1it} (top left panel), but not constant slope, a_{1i} (top right panel), showed abrupt reductions in level that eventually dissipated over time as T approached 20. With the use of the off-line LQC, higher dosages of u_{it}^{R*} were automatically determined and delivered immediately following these shock points, and the resultant, "controlled" latent trajectories (see the solid line marked with the symbol 'C') clearly show quicker return to the unperturbed trajectories compared with the original trajectories without the control input (see the solid line marked with the symbol 'N').

As expected, the dosage strength was proportionate to the extent of deviations from the target latent trajectories. For instance, surges in the dosage of u_{it}^{R*} were observed in ID 1 following the occurrence of two closely located shock points. The shape of the shaded region in the top two plots of Figure 1, which depicts values of u_{it}^{R*}, also helps provide a glimpse into the control input scheme devised for this individual. Given the relatively conservative control scheme used in this particular illustration, the administered u_{it}^{R*} never actually pushed the imposed upper limit of 2. Due to the imposition of a lower constraint

of 0 on u_{it}^{R*}, some "overcorrections" (i.e., ability exceeding and staying above the target level) were observed in this particularly individual.

Deriving control input dosages that are proportional to the amount of deviations from some target level is a key strength afforded by the control theory approach. But what exactly does such an approach add compared to a simpler, observed linear approach that recommends a fixed amount of increment in practice duration with every point that an individual performs below the target level? One key strength of the LQC and related approaches resides in its formulation within a state-space modeling framework and corresponding estimation algorithms that allow the computations of control input based on the latent variable estimates even on occasions where the observed data are missing. To demonstrate this point, we fitted a linear regression model predicting individuals' practice durations based on their observed negative deviations from the target performance levels (i.e., magnitudes of under-performance) at the previous time point, time $t-1$. The resultant intercept and regression coefficient estimate were used to compute predicted durations at time t based on individuals' negative deviations in performance level at time $t-1$. In addition, we randomly set 7 of the 20 occasions (35%) to be missing, and for these missing occasions for which no observed data were available to indicate the amounts of under-performance, we set the recommended training duration to be 0.

Results from this observed linear scheme are depicted in Figure 1 (marked with the symbol 'L'; see the first two rows of plots). It can be seen that the observed linear scheme did indeed produce recommended training durations that were close to those generated with the off-linear LQC, as proportionate to the amounts of under-performance shown by each individual. However, the observed linear scheme performed slightly worse than the off-line LQC, providing delayed duration forecasts to help close the performance gaps especially when the missing data were located close to the shock points. Thus, the LQC and related approaches provide a more "holistic" approach to optimizing the magnitudes of inputs needed to minimize the hypothesized (negative) deviations from the person-specific

target functions.

Finally, we repeated computation of the off-line LQC control input using $h = 10$. Relatively little differences were observed between the two control horizon windows. One relatively trivial difference was that the larger window size of $h = 10$, as compared to $h = 4$, tended to factor into consideration more of the incurred shocks, hence diagnosing higher dosages of input for the exact same amounts of shocks. This, coupled with $R = 10$, at times led to even greater overcorrections in the latent ability trajectories, yielding latent ability values that were greater above the specified target curve than when $h = 4$ was used.

Illustration II: Shocks to Latent Ability and Constant Slope

The shocks incurred in Illustration I were designed to impact individuals' ability levels in a transient way. That is, these shocks acted as unusual process noises, the effects of which persisted for some time but eventually dissipated over time. Thus, in the absence of further shocks, the individuals would, by nature of the hypothesized model, still return to their target trajectories even in the absence of any control input – albeit more slowly. However, in real-life educational settings, shocks to individuals' learning are unlikely to be transient, and might change individuals' learning dynamics or outcomes, leading to irreversible training disparities. To simulate this scenario, we incurred shocks to individuals' latent ability levels as in Illustration I, but additionally, also to their constant slopes, a_{1i}, by drawing from a uniform distribution that ranged from -0.1 to -0.01. Other settings were held identical to those in Illustration I. Such negative shocks to individuals' slopes are known to yield irreversible reductions in the asymptotes of individuals' trajectories (Chow, Hamaker, & Allaire, 2009). Thus, the shocks due to illness or social distress in the previous illustration are no longer a fleeting "nuissance" but rather, would prevent individuals from ever realizing their full potentials in the absence of any interventions.

The resultant trajectories of η_{1it} and a_{1i} for the same hypothetical participant (ID 1) are plotted in the middle row of Figure 1, and the trajectories of two additional participants

are shown in the bottom row. Plot of a_{1i} from ID 1 (middle right panel) highlights that the individual's constant slope was shocked at two distinct time points. Such shocks to the constant slope altered the asymptotic performance level of the individual. Specifically, in the absence of any control input, the individual's latent ability level (the line marked with the symbol 'N') approached a plateau at around 2, in contrast to around 3 as shown by the target trajectory. In this case, the off-line LQC essentially recommended continuous elevated input values, which had the effect of bringing the individual's controlled trajectory close to the target trajectory. The recommended control schemes for two other individuals (see bottom row of Figure 1) had distinct, individual-specific shapes, but shared the same characteristic of a heightened, sustained level of control input. In fact, the constant slope in this model is one example of a latent variable that is *uncontrollable*.[1] Thus, this particular illustration demonstrated a scenario where individuals were subjected to external influences that could lead to permanent, irreversible performance disparities relative to their original asymptotes (their maximum potentials). Even though the solution provided by control theory algorithm might not be ideal (i.e., continuous delivery of input is required), it is still of some utility in highligting the amounts of input needed by different individuals to close the disparities in ability levels.

Illustration III: Effects of Static, One-Size-Fits-All Target Level

This illustration was designed to demonstrate the consequences of using a person-invariant, or in other words, a "one-size-fits-all" target function for η_{1it} for all individuals in a study, defined as the fixed effects curve, $\eta^r_{1it} = \mu_{a1} + b_{11}\eta_{1it}$, starting identically at $\eta_{1i1} = \mu_{\eta_1}$ for all i. The resultant latent ability trajectories for all participants

---

[1] A dynamical system is said to be controllable if it is possible to drive the system into a particular state through the use of manipulable *control inputs* (e.g., interventions, treatment, training). Technically, a system is controllable when the $w \times (w + r)$ controllability matrix,

$$C = \begin{bmatrix} \mathbf{G} & \mathbf{BG} & \mathbf{B}^2\mathbf{G} & \dots & \mathbf{B}^{w-1}\mathbf{G} \end{bmatrix} \tag{19}$$

has rank w (Zarchan & Musoff, 2000). The BDCM-X model is uncontrollable when constant slopes are included as latent variables.

are shown in Figure 2 (see right panel), with the trajectories from Illustration I (see left panel) shown here for all participants on the same scales to facilitate comparisons. In this case, only the participants' latent ability levels, but not their constant slopes, were shocked at two randomly selected time points. For the same participants with otherwise identical trajectories in the absence of any control input (see the trajectories marked with "No Control"), the algorithm now hardly recommended any control input for IDs 1 and 2 because these individuals were consistently performing above the target function. An unconstrained control algorithm would have recommended negative input values, but the constraint that u_{it}^{*} be positive led to the control input profile see in the top right panel of Figure 2. In contrast, for ID 3, an individual who was performing distinctly below the fixed effects curve, much higher positive values of u_{it}^{*} were recommended throughout the study span (see bottom right panel) compared to the recommendations obtained under person-specific target functions (see bottom left panel). Forcing such universal standards on all students without regard to their own strengths and limitations might not serve any individual well in the end (Rose, 2016). This illustration thus pointed to the importance of selecting target functions that can infuse some population standards with each individual's unique learning characteristics in designing control theory interventions.

Illustration IV: Effects of Different Control Input Penalty Weights, R

This illustration serves to clarify the effects of changing the value of R, the cost for input administration. We varied the value of R to be .1, .5, 1, 2, 5, 10, and 100, while keeping the cost matrices for state deviations, $\boldsymbol{Q}$ and $\boldsymbol{Q}_h$, at the same values. The total quadratic costs associated with administering the input, Cost_Input, and the relative benefit gained in terms of reducing overall state deviations from the target level (see Equation 18) are plotted in Figure 3, under the scenarios considered in Illustrations 1 —3. Because no input was administered in the original scenario (i.e., BaseCost_Input = 0 in the baseline condition), RelativeCost as shown in Equation 18 is not defined. Thus, we only plotted RelativeBenefit against Cost_Input.

The plots, shown in Figure 3, indicate that under the scenario considered in Illustration I, the values of relative benefit were positive (indicating a reduction in total state deviations relative to target trajectories) only at larger values of R, suggesting that an overly liberal scheme to deploy control input (i.e., when the values of R were small, such as < 1) yielded control input costs that greatly outweighed the relative benefits. When only individuals' latent ability levels were shocked, the preferred value of R that maximized relative benefit occurred around $R = 10$ whereas when both latent ability and constant slope levels were shocked (as in Illustration II), the slightly lower control input cost of 5 (compared to 10) led to greater reduction in total state deviations under the scenario considered in Illustration II. In addition, when both of these latent components were shocked, the total costs associated with administering the input clearly increased.

Additionally, in the one-size-fits-all scenario with shocks to individuals' latent ability levels only, the input costs and relative benefits showed less variations as a choice of R. Slightly larger values of R (e.g., $2 \leq R \leq 10$), were still preferred, but the cost-and-benefit curve generally assumed a narrower range in relative benefits compared to the earlier scenarios. Overall, the results suggested that the optimal balance of relative benefits and costs of input from the control algorithm depended on the choice of R, and a thorough evaluation of such costs and benefits is imperative.

For comparison purposes, we also added the total input costs and relative benefits associated with the linear observed scheme to the plots, shown as light-shaded, vertical and horizontal lines, respectively, marked with the symbol 'L'. It can be seen that across all illustrations, the offline LQC provided higher relative benefits than the observed linear scheme at $R = 5$ and 10. It should be cautioned, however, that at other less optimized values of R, however, the observed linear scheme actually yielded greater relative benefits than the offline LQC, at comparable total input costs, and much reduced computational time. The reduced total input costs reflected in large part our simulation setting, which recommended no training whenever data were missing at time t-1. Regardless, this simpler

alternative scheme serves as a viable alternative especially in situations where there would be limited missingness from the participants, and underscored the importance of proper selection of the penalty weights, R and Q, in designing and use of the controllers.

Illustration V: Offline Compared to FMLQCs

In the illustrations presented thus far, the control inputs were computed after all the data have already been collected. In other words, the effects of "real-time" implementations and delivery of the control inputs were not reflected in the generated state trajectories, thus increasing the likelihood of "over-corrections" in some individuals' ability in the earlier illustrations. In this illustration, we used the KF- and KS-based FMLQCs in which state estimates were updated in a moving window by incorporating the control inputs computed for that moving window. As shown in Figure 4 (top panel), the real-time update of the control input via the KS-based FMLQC outperformed the other LQC variations by yielding more targeted and timely reductions of control input. The state trajectory controlled under these KS-FMLQC regulated input values (see trajectory labelled as "FMLQC w/ KS") now approached the target level more precisely and showed less over-corrections. Note that positive deviations in performance were generally disregarded based on our constraints on the control input, u_{it}^{R}. Thus some over-corrections were still present in some of the illustrative cases. The corresponding costs and benefit comparisons in the bottom panel suggested that the FMLQCs led to comparable cost and benefit curves compared to the off-line LQC. However, the KF-based FMLQC, in contrast to both the off-line and KS-based FMLQC, yielded less pronoounced increases in total reductions in state deviations compared to the baseline condition when no control input was used (see the dashed horizontal reference line).

Summary of Illustrations: We demonstrated the effects of three LQC through five illustrative simulations. These illustrations highlighted the effects of administering constrained control input under relatively simple scenarios with no process noises and a limited number of shocks to the system.

Empirically Motivated Simulations

To demonstrate the feasibility and utility of using control theory optimization in a real-world scenario, we used a subset of national Dutch elementary school students' data on the Math Garden in the current application. Math Garden is a computerized adaptive practice system that utilizes the Elo rating system developed for chess competitions to perform both person and item parameter estimations on the fly (Maris & van der Maas, 2012), thus allowing educators and researchers to bypass the need to implement expensive pre-testing of the item bank (Klinkenberg et al., 2011). Training dosage was operationalized in the context of the Math Garden data as each student's average weekly activity time on the website in hours, calculated using the timestamps associated with the users' responses.

Math Garden contains 15 games covering the math curriculum of elementary schools, including arithmetic operations such as addition, subtraction, multiplication, and division. For illustrative purposes, we used weekly ability estimates and reaction time data on the division task only as dependent variables for model fitting and control theory testing purposes. Previous analyses of the Math Garden data have focused primarily on students' performance on the addition and multiplication tasks (e.g., Jansen, Hofman, Savi, Visser, & van der Maas, 2016). Here, we chose to focus on the division task because student performance on this task showed clear improvements over time and across multiple grade levels, but also frequent shifts and deviations from an idealized population curve.

We demonstrate, by using the BDCM-X as our operating model, that a student's performance on the division task can be more efficaciously driven toward a pre-defined target level via the KS-based FMLQC. We focused on simultaneous modeling of individuals' latent ability and reaction time given the known reciprocal effects between reaction time and latent ability and their corresponding estimates. Inclusion of reaction time would allow us to address questions such as: among individuals with the same reaction time (or controlling for the effects of reaction time), whether longer training duration at the previous week helped promote more growth in division ability this week. Students' original

ability estimates from Math Garden lied on a Rasch-type scale (Brinkhuis et al., 2018). Prior to model fitting, the ability scores were recoded by adding a minimum constant to the scores so a score of zero corresponded to the minimum observed division ability in the sample. No recoding was performed on the covariate (training duration) or reaction time.

We sought to address the following questions:

1. In what ways, if at all, are the KS-based FMLQC-recommended training durations "better" compared to a fixed, one-size-fits-all training scheme in which all individuals adhered to a strict weekly practice duration of 14.36 minutes (the median practice duration of the whole population of Math Garden users; coinciding also with the approximate practice duration recommended by the app developers)?

2. In what ways, if at all, are the KS-based FMLQC-recommended training durations "better" compared to the original practice durations recorded for these students?

3. What are the effects of using a target function based strictly on population standards, such as the grade-normed median, as compared to one that integrates population standards and some person-specific information, such as each student's unique model-implied trajectory as in the illustrative simulations?

We answered these questions through a series of empirically-motivated simulations. Specifically, we first estimated the values of $\boldsymbol{\theta} = [b_{11}, b_{12}, b_{21}, b_{22}, g_1, g_2, \psi_{11}, \psi_{22}, \sigma^2_{\epsilon_1}, \sigma^2_{\epsilon_2}, \sigma_{v(.)}, \sigma^2_{v(.)}]'$, where $\sigma^2_{v(.)}$ and $\sigma_{v(.)}$ denote all the variance and covariance parameters for the random effects shown in Equation (7). Other parameters are as defined in Equations (5) and (8). This was done by fitting the BDCM-X model to data from approximately half of the sample ($n = 400$; referred to herein as the *estimation sample*). Then, using the data and change characteristics (e.g., observed training durations, initial level and constant slope estimates) from the remaining $n = 384$ *validation* sample, we conducted a series of empirically motivated simulations to address our questions of interest.

To fit the BDCM-X model, an equally spaced model to the estimation sample,

missing data were inserted for the weeks during which no Math Garden activity was recorded. For these weeks with missing data, we imputed the value of 0 for training duration, and left the missing data for the ability and reaction time as they were to be handled via full-information maximum likelihood. After removing participants with excessive missingness (i.e., missing rate of $> 70\%$ or less than 6 non-missing observations), a sample of 784 students were retained. These students worked on the division task at their own schedules or as recommended by their schools, contributing data ranging from 6 to 282 weeks (median = 90 weeks) as they attended 3rd to 12th grade (median = 5th grade). After imputation of the missing training duration values with 0, corresponding to weeks on which the participants did not attempt any activities in Math Garden, the average amount of weekly Math Garden training duration recorded by the system was 0.12 hour (7.2 minutes), with a median of 0 and SD of 0.20. The median training duration prior to imputation was 0.24 hour (14.36 minutes).

Plots of the division ability scores from a sample of 100 randomly selected participants and their corresponding reaction time data are shown in Figure 5(A) and 5(B), respectively. As a comparison, we also plotted the median ability score of all 5th-grade (the median grade in the sample) students in the entire Math Garden database. The plot indicated that the current sample started out with an initial ability level that coincided closely with the 5th-grade median. Whereas some improvements were observed in many students over weeks, there was considerable heterogeneity in each individual's learning trajectory.

BDCM-X Modeling Results with the Estimation Sample

Results from fitting the BDCM-X model to the estimation sample suggested that with the exception of the measurement error variance for the division ability score, the process noise variance for reaction time, and some covariance terms among the random effects, all other parameters were significantly different from zero. Parameters that were not reliably different from zero, except for $\sigma^2_{\epsilon_1}$, were then fixed at zero, and empirical results from

fitting the refined model are shown in Table 1. The estimated means of the initial levels of ability and reaction were positive (μ_{η_1} and μ_{η_2}) and close to the empirical means of the observed ability scores and reaction time at time 1, with substantial interindividual differences. As noted, some of the covariances between random effects were not reliably different from zero, and were fixed at zero. Covariances that were retained and remained statistically significant included covariances between the random effects of initial division ability and initial reaction time ($\sigma_{v_{\eta_1},v_{\eta_2}}$), and between initial division ability and the constant change parameter for division ability ($\sigma_{v_{\eta_1},v_{a_1}}$). Estimates for these covariance terms as shown in Table 1 suggested that individuals who tended to have higher initial division ability also showed slightly longer reaction time, and those with higher initial ability were associated with higher constant slopes, a_{1i}.

The auto-proportion parameters for both latent ability and reaction time were both negative and significantly different from zero, suggesting that reduced latent growth in division ability tended to be observed for an individual when the individual's previous ability level at the previous week was high, especially as the individual approached his or her personal asymptote. Relatedly, there were small, reciprocal positive couplings between ability and reaction time, indicating that higher previous ability and higher reaction time at the previous week were associated, respectively, with greater latent changes in reaction time and latent ability this week. These findings were consistent with the design and adaptive nature of Math Garden – that is, a higher previous latent ability would prompt the system to present a student with more difficult items on the next trial. Taking the time to get these more challenging items correct (as opposed to resorting to hints or venturing guesses haphazardly), in turn, would yield a higher ability estimate for the student.

Previous week's training duration (activity time in Math Garden as measured in hours) was found to have significant positive effects on the current week's latent changes in division ability as well as reaction time. A larger amount of change was observed in latent ability level than in reaction time in seconds per hour of change in training duration,

possibly due to the limited changes individuals could display on reaction time under the system-imposed time limit. Although a direct comparison of the magnitudes of these control input-related coefficients would not be meaningful due to scaling differences between the constructs, the overall results from model fitting suggested that previous week's training duration could be a viable candidate as a control input to drive future changes in individuals' division ability.

Empirically Motivated Simulations Using the Validation Sample

Fixing the parameter values in $\boldsymbol{\theta}$ to the estimated values obtained from the estimation sample, we then performed a series of empirically motivated simulations to evaluate the effects of applying the KS-based FMLQC to the validation sample. Briefly, data were simulated in ways that mirror as closely as possible to the empirical characteristics of the validation sample. Specifically, we first applied the KS with $\boldsymbol{\theta}$ fixed at those obtained with the estimation sample to yield initial level and slope (i.e., η_{1i1}, η_{2i1}, a_{1i}, and a_{2i}) estimates for each individual in the validation sample. These initial level and slope estimates were used to generate simulated data sequentially (for $t = 2, \ldots, T_i$) based on Equations (5)-(8), and also to define person-specific target functions in some of the subsequent simulations. Process noises and measurement noises were added to the simulated data based on the normality assumptions outlined in Equations (5) and (8).

In short, our simulation specifications allowed us to make targeted manipulations of individuals' training durations according to different training schemes, while holding all other confounding factors constant – including initial conditions, parameter values, and sequences of process and measurement noises. We organized our simulation results based on the research questions outlined earlier. As in the illustrative simulation, we set $\mathbf{Q} = \mathbf{Q_h} = \boldsymbol{\Lambda}'\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ in all the empirically motivated simulations. Consonant

with the goal of the Math Garden app to improve students' arithmetic performance, we imposed a lower limit of $u_{lower,i} = 0$, and a constant, person-invariant upper limit as given by the 99th percentile of all students' weekly training duration, namely, $u_{upper} = 1.01$ hours (corresponding approximately to an average of 8.67 minutes per day). Finally, we set the finite memory window, n_h of the KS-based FMLQC to be 20.

Empirical Illustration I: KS-Based FMLQC-Recommendations Compared to Practicing at the Median Duration

Our first question of interest was whether and to what extent the KS-based FMLQC-recommended training durations led to greater training efficacy compared to a simpler, one-size-fits-all training scheme whereby all individuals adhered to a strict weekly practice duration. To test this question, we selected 14.36 minutes, the pre-imputation median practice duration of the whole population of Math Garden users as the fixed training duration against which the KS-based FMLQC-recommended training durations were compared. We set the target function to be each individual's grade-level median.

To compare the costs and benefits associated with the two training schemes, we computed and plotted the Relative Costs and Relative Benefits (see Equation 18) associated with the LQC training scheme, as compared to the fixed-duration scheme as a baseline across a range of values of input cost weight, R from 1 to 200. The corresponding relative cost and benefit values are shown in Figure 6. In the plot, positive (negative) values on the ordinate (vertical axis) indicate reductions (increases) in total state deviations under the LQC-recommended as compared to the fixed-duration training scheme. In contrast, the abscissa (horizontal axis) serves to highlight relative increases in total input cost, with positive (negative) values indicating increases (decreases) in total input cost under the LQC-recommended as compared to the fixed-duration scheme. Based on Figure 6, values of R that were equal to or higher than 2 were found to yield relative reductions in total state deviations compared to the target function (the grade-level median).

Of particular interest is the upper left region of that plot with an arrow. This region

captures instances where the KS-based FMLQC training scheme led to a reduction in total state deviations *as well as* a reduction in total input cost. That is, with values of R set at 100 or 200, it is possible for the students to practice less and still show lower total quadratic deviations in ability compared to the grade median. Note that even though the cost functions utilized are quadratic functions, instances where individuals performed above the grade median (i.e., positive deviations in ability levels) would automatically be ignored by the proposed FMLQC training scheme because of the constraint that $u_{it}^{*} \geq 0$. Thus, if the goal is for the students to perform at least as well as the grade median, the preferred FMLQC training scheme appeared to one that suggested relatively high penalty of administering training (R of 100 or 200), recommending training only for those instances where the students performed below the grade median, and in amounts that were proportionate to the deviations from the target function.

Empirical Illustration II: KS-Based FMLQC-Recommended Durations Compared to the Original Durations. Our second question of interest was whether and to what extent the LQC-recommended training durations would yield improved training efficacy compared to the original practice durations of the Math Garden users. As in illustration I, we computed and plotted the Relative Costs and Relative Benefits (see Equation 18) associated with the KS-based FMLQC training scheme, but now as compared to the original observed training durations, again across a range of values of input cost weight, R, from 1 to 200.

The relative benefits and costs, as plotted in Figure 7, were similar to those observed in Illustration I. That is, by setting the value of R to 100 or 200, it is possible for the students to practice less and still show lower total quadratic deviations in ability compared to the grade median. The only minor difference was the slight decreases in input costs compared to those observed under Empirical Illustration I.

Empirical Illustration III: Population Compared to Hybrid Target Functions

The first two illustrations were built on a nomothetic target function based on the population median. In practice, this target function might not serve the training goals of all individuals well. In this illustration, we explored the effects of using a hybrid target function that integrates population standards as well as some person-specific (idiographic) information, such as each student's own trajectory as implied by the BDCM-X model under no additional training.

We selected the alternative, hybrid target function by setting $\boldsymbol{\eta}^r$, the target ability to be $max\Big(\text{grade median}, E(\eta_{1it}|\eta_{1i1}, ..., \eta_{1i,t-1}, u_{it} = 0)\Big)$, namely, the higher value of the grade-normed median level, or individual i's BDCM-X model-implied latent ability trajectory. The latter was computed by setting the parameter values to those estimated using the estimated sample, and additionally, with each individual's initial level and constant slope set to the corresponding smoothed estimates for that individual at $t = 1$. This trajectory, appearing as a sigmoid-shaped curve, provided a set of alternative target functions toward which individuals' Elo scores could be driven if they happened to perform above their grade-level medians. The cost and benefit comparisons in Figure 8 revealed that when the hybrid target function was used, the KS-based FMLQC-recommended training scheme still yielded less total state deviations at reduced total input costs from this target function at $R = 100$ and 200 when compared to both a fixed-duration scheme (left panel), as well as the original observed training durations (right panel).

To inspect the ways in which the FMLQC-recommended training durations differed in magnitude and timing compared to individuals' original training durations, we plotted in Figure 9: individuals' observed ability (marked as "Observed") scores for four selected individuals (IDs 2, 3, 4 and 6); their predicted ability trajectories generated using the students' original durations, $E(\text{ability}_{it}|\text{original duration}_{i,t-1})$, denoted as "Original predicted ability" in the figure; their corresponding predicted ability generated using the FMLQC-recommended training durations, $E(\text{ability}_{it}|u^*_{it})$, marked as "Predicted ability with new u"; and the KS latent variable estimates obtained in finite memory (FM-KS)

windows (denoted as "FM KS estimates"). We also plotted the FMLQC-recommended and original training durations as shaded regions and unshaded regions marked with slanted lines, respectively. The absence of shading corresponded to periods during which the FMLQC recommended no training ($u^*_{it} = 0$).

The four illustrative students were selected from the larger validation sample of $n = 384$ because they underwent at least one transition to a higher grade during the observed span of the study, and are characterized by a range of ability levels. For instance, the target function for participant 6 was based largely on the grade-normed median curve. In contrast, participants 1, 3 and 4 consistently outperformed the grade median levels, and were thus assigned target trajectories based on their BDCS-X-implied growth trajectories. These students' observed Elo scores (marked as "observed in the plots) were interspersed with periods of positive as well as negative deviations from their target functions.

At R =100 (the top and middle rows of Figure 9), the LQC training scheme recommended more concentrated training durations on the occasions when individuals fell below their target levels (e.g., around $t \geq 45$ for ID 3), and at amounts that were proportionate to the magnitudes of negative deviations (i.e., how much the individual under-performed) from the individuals' target trajectories. Such heterogeneity in recommended training durations and timing further confirmed that using only one static population standard as the target level or a fixed-duration scheme might not be adequate to help each individual student realize his/her full learning potential.

To clarify the effects of using a smaller R, we plotted in the last row of Figure 9 the LQC-recommended training scheme and corresponding trajectories ("Controlled") for IDs 1 and 3 under $R = 10$. With the lower penalty value, greater magnitudes were generally recommended during the same periods of under-performance for the two individuals. However, because the FMLQC-recommended training "interventions" were never actually administered, the KS estimates of the latent variables used for computing u^*_{it}, which comprised weighted combinations of the observed data and model-implied trajectories

(similar in form to the sigmoid-shaped target trajectory) continued to suggest under-performance from the target trajectory as the state estimates were pulled down by these individuals' actual Elo scores. As a result, the last plot for ID 3 in Figure 9 underscored specifically a scenario where the KS-FMLQC algorithm did not work well. That is, this scenario may correspond to real-world situations where the training might not yield the intended outcomes for some individuals – for example, when the training delivered did not help improve learning for subgroups of students, or the recommendations were ignored altogether by the students. In this case, continuing to deliver the training recommendations at low R value would not help to reduce the total deviations, and additionally, could become very costly.

Note that the use of the FM-KS provided latent variable estimates that closely tracked the observed ability levels of the participants, and additionally provided imputed ability values for occasions with missing data. This is a useful property of the KF/KS procedures – that is, through a weighted average of model predictions and observed data, these procedures can yield latent variable estimates that track the observed data relatively closely even if the dynamic model used for forecasting purposes is imperfect. These latent variable estimates can, in turn, be used to compute optimal control input values. Despite the usefulness of these latent variable estimates, the discrepancies of the predicted ability scores (i.e., model-implied ability values without conditioning on observed data) relative to the FM-KS or observed ability scores still highlighted some inadequacies of the BDCM-X model in capturing the change characteristics of the students. We address some of these inadequacies in the Discussion section.

Overall, our three empirically motivated simulations served to demonstrate the utility and feasibility of using a constrained controller in conjunction with a group-based state-space model to improve the training efficacy of educational apps such as Math Garden. We found that with appropriate choice of input penalty, R, the LQC training schemes could yield increased benefits (in terms of minimizing deviations from target

performance levels) and reduced training durations compared to alternative training schemes such as the fixed-duration and the original observed training schemes.

Discussion

In this article, we proposed and evaluated three variations of the LQC with constraints to forecast the optimal weekly training durations for individual users of Math Garden, an educational app designed to enhance arithmetic learning for elementary school students. Population-level performance standards and individual learning information were used to construct person- and time-specific target performance trajectories, the deviations from which would trigger proportionally scaled training dosage to accelerate closing of such performance gaps. We demonstrated one possible way of integrating population standards with each student's own latent change information through a series of illustrative and empirically motivated illustrations, and showed that adoption of the control theory-guided, person- and time-specific training dosages could yield increased training benefits at reduced costs compared to students' actual observed training durations or a fixed-duration training scheme. In addition, actual user training data were used to guide the selection of the constraints on training. In the Math Garden application, these constraints included imposing an upper limit that was no more than one hour of training each week, and disregarding positive deviations, namely, instances where students over-performed compared to target levels.

We note here that the goal of our control theory application — namely, to control or manipulate some input (training duration) to minimize discrepancies from an objective function — has some conceptual similarity to the adaptive nature of the Elo system to tailor the difficulty levels of the assigned items to an individual's estimated ability level on the fly (Klinkenberg et al., 2011). However, the nature of the problems and estimation algorithms needed to fulfill these respective purposes are distinct because the matching of item and person characteristics is passive in adaptive systems such as the Elo system. That is, in the Elo system, the goal is to *assess* an individual's ability accurately (Park, Joo,

Cornillie, van der Maas, & Van den Noortgate, 2019), not to change, improve — or specifically, *control* — the ability of that person. In contrast, in a control theory application, the goal is to actively control the endogenous process (ability) by manipulating some exogenous variables in ways that would minimize discrepancies from an objective function of choice. To our knowledge, the current Math Garden-inspired application was the first application of such constrained control theory principles to large-sample real-world data in the social and behavioral sciences. The current work was also novel as a first attempt at combining a group-based state-space model that is non-stationary, namely, the BDCM-X, which postulates over-time changes in means as well as variance-covariance functions, with constrained LQCs. In the control theory literature, constrained LQCs are typically applied to stationary control problems, and at the individual level. Such direct application of the constrained LQCs was made possible through our use of person- and time-varying target functions. In addition, as distinct from previous applications that utilized true or strictly model-implied latent variable scores to compute control input values (Molenaar, 2010; Wang et al., 2014), we demonstrated the feasibility of using KS estimates, which combined information from model predictions and observed data beyond the current time point, for control input estimation at time t.

Promising applications of this technology in Math Garden and similar systems may focus on further refinements of the proposed algorithm to provide recommendations for durations spent on different types of exercises within games, as well as the selection of the games themselves. Examples of other applications that may benefit from use of control theory principles include apps that help individuals regulate their daily physical activity levels, educational apps targeting other learning domains such as reading, and mobile health devices that help individuals regulate their affect intensity and arousal levels.

Some software innovations are also available as part of this paper. We extended functions from the R package, *dynr* (Ou et al., 2019), to use by-products from the KF-related routines from this package to automate efficient computation of control input.

We provided the code for the illustrative simulations as supplementary material with this paper in hopes of facilitating further extensions and adoptions across a broad array of settings. In terms of computational time, there are notable discrepancies in the requisite computational time for applying the off-line LQC to all individuals' training durations at once after all the data have already arrived, as compared to applying FMLQCs as the data are collected. In the case of off-line LQC, it took only 32.66 seconds for us to forecast the input for all individuals and time points in our validation sample on a Mac computer with 2.3 GHz Intel Core i9 and 15 GB of 2400 MHz DDR4 memory. To perform KS-based FMLQC estimation of the training durations with a finite memory window of $n_h = 20$, repeated passing of information between R and the underlying C code in *dynr* is required. In this case, the computational time increased substantially to approximately 42.9 minutes. We note, however, that in most applications, FMLQCs only need to be applied to compute training durations one time step (or specifically, window) ahead as new data arrive. Each set of one-window-ahead forecasts for the entire validation sample requires approximately 2.612 seconds on the same computer.

The current study has several limitations. Currently, we used only a subset of the data from the Math Garden data base. These participants were specifically selected to have at least 5 practice sessions on the Math Garden. The extent to which our current results are generalizeable to all users of Math Garden is unclear and warrants more thorough investigation. In addition, the BDCM-X was specifically selected for this application because it captured the functional forms of the learning trajectories we observed in the empirical data. That is, we regarded this model more as a useful model rather than the true model of change. Knowledge concerning what omitted variables are responsible for driving real-world change processes is often limited in social and behavioral science applications. Our view was that the BDCM-X was useful as a building block to help individuals set learning target and obtain recommendations for training durations than other alternative (e.g., linear) training schemes. Caution would have to be exercised in drawing causal

inferences based on this model. As well, applicability of the proposed algorithms to other models, contexts, sample sizes, and design configurations will have to be examined more extensively. In addition, due to the widespread use of the Math Garden app in the Netherlands, population norms are available for designing the control theory algorithm in our empirical application. The plausibility of real-time adoption of the control theory algorithm recursively in other newer apps would have to be investigated with caution.

Several other design considerations have to be investigated more thoroughly in future studies to enable real-life adoption of the proposed approaches. First, our design may be regarded as a serendipitous design because computation of the control input takes no special considerations of the presence of the constraints in the first place (Goodwin et al., 2005). Such serendipitous designs may have reduced efficiency compared to designs that are truly optimized for constrained control purposes. A better alternative would be to perform direct constrained optimization by means of quadratic programming or other related estimation approaches (Bemporad, Morari, Dua, & Pistikopoulos, 2002; Goodwin et al., 2005; Seron, Goodwin, & De Doná, 2003). Second, even though a stochastic state-space model was used in the present article, the input values were computed via a deterministic LQC. That is, the state values were assumed "known and fixed" at the values of the conditional mean estimates from the KF or KS. This kind of control schemes works well when the separable principle holds, namely, when optimal control and state estimation can be decoupled under regularity conditions (Alspach, 1975). This was the case in the model considered in the present article, but this assumption may not hold in other empirical scenarios. In such cases, other stochastic control schemes may have to be utilized instead (Alspach, 1975; Bar-Shalom & Tse, 1976; Lu & Zhang, 2016). In addition, the values of the cost matrices, $\mathbf{Q}$, $\mathbf{Q}_h$, and $\mathbf{R}$, were selected in the current application heuristically through repeated trials. In the future, other more formal selection criteria or measures should be considered and further evaluated to help guide the selection of these cost matrices.

Finally, the quadratic cost function in Equation (13) penalizes positive deviations

from the target functions just as heavily as negative deviations. In practice, negative deviations (i.e., performing below the target functions) are of much greater concerns than positive deviations. We circumvented this limitation by using constraints to bypass recommendations to reduce training durations. An alternative would be to utilize nonlinear cost functions to explicitly target deviations in one direction (e.g., Taguchi, Chowdhury, & Wu, 2005; van den Berg, 2014; Zhang, Li, Wang, & Jin, 2014). Given the initial promise shown by our proof-of-concept simulations, further optimizations of the proposed estimation approaches are warranted.

Closing Remarks

The constant influx of new training options and educational apps in this digital age has provided students, educators, and training institutions with better and more inclusive ways of training students. Unfortunately, one-size-fits-all training is known to be inefficient. The appeal of personalized educational pathways is clear to many educators; however, the burden on the instructors to provide personalized training recommendations can be heavy. In this article, we presented and evaluated several variations of a constrained LQC that automate the delivery of optimal training dosage much in the way that the cruise control unit of a car regulates discrepancies between actual and target driving speed. While the overall designs and some of the results are still nascent, we hope that the proposed approach nevertheless provides a preliminary computational backbone to inspire more work to personalize the future of digital education.

References

Academy, K. (2017). *Khan academy.* Retrieved from https://www.khanacademy.org/ ([Online; accessed 20-October-2017])

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255-265.

Alspach, D. (1975). A stochastic control algorithm for systems with control dependent plant and measurement noise. *Computers & Electrical Engineering*, *2*(4), 297-306. Retrieved from https://www.sciencedirect.com/science/article/pii/0045790675900178 doi: https://doi.org/10.1016/0045-7906(75)90017-8

Åström, K. J., & Murray, R. M. (2008). *Feedback systems: An introduction for scientists and engineers*. New Kersey: Princeton University Press.

Bar-Shalom, Y., & Tse, E. (1976). Concepts and methods in stochastic control. In C. Leondes (Ed.), (Vol. 12, p. 99-172). Academic Press. doi: 10.1016/B978-0-12-012712-2.50009-3

Bavdekar, V. A., Bhushan Gopaluni, R., & Shah, S. L. (2013). A comparison of moving horizon and bayesian state estimators with an application to a ph process. *IFAC Proceedings Volumes*, *46*(32), 160-165. Retrieved from https://www.sciencedirect.com/science/article/pii/S1474667015382513 (10th IFAC International Symposium on Dynamics and Control of Process Systems) doi: https://doi.org/10.3182/20131218-3-IN-2045.00152

Bellman, R. (1964). Control theory. *Scientific American*, *211*, 186-200.

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica*, *38*(1), 3-20.

Biddlecomb, B. D. (2002). Numerical knowledge as enabling and constraining fraction knowledge: an example of the reorganization hypothesis. *The Journal of Mathematical Behavior*, *21*(2), 167 - 190. Retrieved from

http://www.sciencedirect.com/science/article/pii/S0732312302001177 doi: 10.1016/S0732-3123(02)00117-7

Blanton, M., Stephens, A., Knuth, E., Gardiner, A. M., Isler, I., & Kim, J. S. (2015). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. *Journal for Research in Mathematics Education*, *46*(1), 39–87. doi: 10.5951/jresematheduc.46.1.0039

Brinkhuis, M. J., Savi, A. O., Hofman, A. D., Coomans, F., van Der Maas, H. L., & Maris, G. (2018). Learning as it happens: A decade of analyzing and shaping a large-scale online learning system. *Journal of Learning Analytics*, *5*(2), 29-46. doi: 10.18608/jla.2018.52.3

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, *92*, 111-135.

Chow, S.-M., Grimm, K. J., Guillaume, F., Dolan, C. V., & McArdle, J. J. (2013). Regime-switching bivariate dual change score model. *Multivariate Behavioral Research*, *48*(4), 463-502.

Chow, S.-M., Hamaker, E. J., & Allaire, J. C. (2009). Using innovative outliers to detecting discrete shifts in dynamics in group-based state-space models. *Multivariate Behavioral Research*, *44*, 465-496. doi: 10.1080/00273170903103324

Chow, S.-M., Ho, M.-H. R., Hamaker, E. J., & Dolan, C. V. (2010). Equivalences and differences between structural equation and state-space modeling frameworks. *Structural Equation Modeling*, *17*(303-332). doi: 10.1080/10705511003661553

Chow, S.-M., & Zhang, G. (2013). Nonlinear regime-switching state-space (RSSS) models. *Psychometrika: Application Reviews and Case Studies*, *78*(4), 740-768.

Dowker, A. (2015). Individual differences in arithmetical abilities: The componential nature of arithmetic. In *The oxford handbook of numerical cognition.* (pp. 878–894). New York, NY, US: Oxford University Press.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199641178.001.0001

Goodwin, G., Seron, M. M., & de Don, J. A. (2005). *Constrained control and estimation: An optimisation approach* (1st ed.). London, United Kingdom: Springer-Verlag London Ltd.

Graeber, A. O., & Tirosh, D. (1990, December). Insights fourth and fifth graders bring to multiplication and division with decimals. *Educational Studies in Mathematics*, *21*(6), 565–588. doi: 10.1007/BF00315945

Hackenberg, A. J., & Lee, M. Y. (2015). Relationships between students' fractional knowledge and equation writing. *Journal for Research in Mathematics Education*, *46*(2), 196–243.

Hadass, R., & Bransky, J. (1991). Meanings of division and their implication for science teaching — a survey amongst elementary school teachers. *International Journal of Mathematical Education in Science and Technology*, *22*(2), 309-315. Retrieved from https://doi.org/10.1080/0020739910220216 doi: 10.1080/0020739910220216

Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Princeton, NJ: Princeton University Press.

Harvey, A. C. (2001). *Forecasting, structural time series models and the Kalman filter*. Cambridge, United Kingdom: Cambridge University Press.

Jansen, B. R., Hofman, A. D., Savi, A., Visser, I., & van der Maas, H. L. (2016). Self-adapting the success rate when practicing math. *Learning and Individual Differences*, *51*, 1 - 10. Retrieved from http://www.sciencedirect.com/science/article/pii/S1041608016301716 doi: https://doi.org/10.1016/j.lindif.2016.08.027

J.B., R., Mayne, D., & Diehl, M. (2017). *Model predictive control: theory, design, and*

computation (2nd ed.). Madison: Nob Hill Publishing.

Klinkenberg, S., Straatemeier, M., & van der Maas, H. L. J. v. d. (2011). Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education*, *57*(2), 1813 – 1824. Retrieved from http://www.sciencedirect.com/science/article/pii/S0360131511000418 doi: https://doi.org/10.1016/j.compedu.2011.02.003

Kuo, B. C. (1991). *Automatic control systems* (6th ed.). New Jersey: Prentice Hall.

Kwon, W. H., & Han, S. H. (2005). *Receding horizon control: model predictive control for state models*. London: Springer.

Lee, J.-E. (2007). Making sense of the traditional long division algorithm. *The Journal of Mathematical Behavior*, *26*(1), 48 - 59. Retrieved from http://www.sciencedirect.com/science/article/pii/S0732312307000065 doi: 10.1016/j.jmathb.2007.03.001

Li, Y., & Silver, E. A. (2000). Can younger students succeed where older students fail? An examination of third graders' solutions of a division-with-remainder (DWR) problem. *Journal of Mathematical Behavior*, *19*(2), 233–246. doi: 10.1016/s0732-3123(00)00046-8

Liu, J., Wang, W., Golnaraghi, F., & Kubica, E. (2010). A novel fuzzy framework for nonlinear system control. *Fuzzy sets and systems*, *161*, 186-200.

Ljung, L., & Söderström. (1983). *Theory and practice of recursive identification*. MIT Press.

Lo, J.-J., & Watanabe, T. (1997). Developing ratio and proportion schemes: A story of a fifth grader. *Journal for Research in Mathematics Education*, *28*(2), 216–236. Retrieved from http://www.jstor.org/stable/749762

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Lu, Q., & Zhang, X. (2016). A mini-course on stochastic control. In *Control and inverse problems for partial differential equations* (p. 171-254). Retrieved from https://www.worldscientific.com/doi/abs/10.1142/9789813276154_0004 doi: 10.1142/9789813276154_0004

Lütkepohl, H. (2005). *Introduction to multiple time series analysis* (2nd ed.). New York, NY: Springer-Verlag.

Maris, G., & van der Maas, H. (2012). Navigating massive open online courses. *Psychometrika*, *77*(4), 615-633. doi: 10.1007/s11336-012-9288-y

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analysis with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (p. 139-175). Washington, DC: American Psychological Association.

Molenaar, P. C. (2010). Note on optimization of individual psychotherapeutic processes. *Journal of Mathematical Psychology*, *54*(1), 208–213.

Mulligan, J. T., & Mitchelmore, M. C. (1997). Young children's intuitive models of multiplication and division. *Journal for Research in Mathematics Education*, *28*(3), 309–330. Retrieved from http://www.jstor.org/stable/749783

Ou, L., Hunter, M. D., & Chow, S. (2019). What's for dynr: A Package for Linear and Nonlinear Dynamic Modeling in R. *The R Journal*. Retrieved from https://journal.r-project.org/archive/2019/RJ-2019-012/index.html doi: 10.32614/RJ-2019-012

Park, J. Y., Joo, S. H., Cornillie, F., van der Maas, H. L., & Van den Noortgate, W. (2019). An explanatory item response theory method for alleviating the cold-start problem in adaptive learning environments. *Behavior research methods*, *51*(2), 895-909. doi: 10.3758/s13428-018-1166-9

Rivera, D. E., Pew, M. D., & Collins, L. M. (2007). Using engineering control principles to inform the design of adaptive interventions A conceptual introduction. *Drug and*

Alcohol Dependence, *88S*(S31-S40).

Rose, T. (2016). *The end of average: How we succeed in a world that values sameness*. San Francisco, CA: HarperOne. Retrieved from http://www.harpercollins.com/9780062358363/the-end-of-average

Savi, A. O., van der Maas, H. L. J., & Maris, G. K. J. (2015). Navigating massive open online courses. *Science*, *347*(6225), 958–958. Retrieved from https://science.sciencemag.org/content/347/6225/958 doi: 10.1126/science.347.6225.958

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464. doi: 10.1214/aos/1176344136

Schweppe, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*, *11*, 61-70.

Seron, M. M., Goodwin, G. C., & De Doná, J. A. (2003). Characterisation of receding horizon control for constrained linear systems. *Asian Journal of Control*, *5*(2), 271-286.

Shumway, R. H. (2000). Dynamic mixed models for irregularly observed time series. *Resenhas- Reviews of the Institute of Mathematics and Statistics*, *4*, 433-456.

Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications*. New York, NY: Springer Texts in Statistics.

Taguchi, G., Chowdhury, S., & Wu, Y. (2005). *Tagushi's quality engineering handbook*. Hoboken, NJ: Wiley & Sons. doi: 10.1002/9780470258354

van den Berg, J. (2014). Iterated lqr smoothing for locally-optimal feedback control of systems with non-linear dynamics and non-quadratic cost. In *2014 american control conference* (p. 1912-1918). doi: 10.1109/ACC.2014.6859404

Wang, Q., Molenaar, P., Harsh, S., Freeman, K., Xie, J., Gold, C., ... Ulbrecht, J. (2014). Personalized state-space modeling of glucose dynamics for type 1 diabetes using continuously monitored glucose, insulin dose, and meal intake: An extended kalman

filter approach. *Journal of Diabetes Science and Technology*, *8*(2), 331-345. Retrieved from http://dst.sagepub.com/content/8/2/331.abstract doi: 10.1177/1932296814524080

Zarchan, P., & Musoff, H. (2000). *Fundamentals of Kalman filtering: A practical approach. Progress in Astronautics and Aeronautics*. Virginia: American Institute of Aeronautics and Astronautics, Inc.

Zhang, J., Li, W., Wang, K., & Jin, R. (2014). Process adjustment with an asymmetric quality loss function. *Journal of Manufacturing Systems*, *33*(1), 159-165. Retrieved from https://www.sciencedirect.com/science/article/pii/S0278612513001015 doi: https://doi.org/10.1016/j.jmsy.2013.10.001

Table 1
Results from Fitting the BDCM-X to the Empirical Math Garden Division Scores and Corresponding Reaction Time.

	Estimate	Std. Error	t value	ci.lower	ci.upper	Pr(>|t|)
b_{11}	-0.034	0.002	-20.674	-0.037	-0.031	0.000
b_{22}	-0.013	0.002	-7.830	-0.016	-0.009	0.000
b_{12}	0.036	0.008	4.276	0.019	0.052	0.000
b_{21}	0.001	0.000	3.009	0.000	0.001	0.001
g_1	0.634	0.026	24.775	0.584	0.684	0.000
g_2	0.025	0.010	2.520	0.006	0.045	0.006
ψ_{11}	0.720	0.008	88.255	0.704	0.736	0.000
$\sigma^2_{\epsilon_1}$	0.001	0.001	1.400	-0.000	0.002	0.081
$\sigma^2_{\epsilon_2}$	4.615	0.052	88.101	4.512	4.717	0.000
μ_{η_1}	10.986	0.217	50.734	10.562	11.411	0.000
μ_{a_1}	0.451	0.046	9.700	0.360	0.542	0.000
μ_{η_2}	6.584	0.087	75.668	6.414	6.755	0.000
μ_{a_2}	0.099	0.010	9.393	0.078	0.119	0.000
$\sigma^2_{v_{\eta_1}}$	18.747	1.323	14.165	16.153	21.341	0.000
$\sigma_{v_{\eta_1},v_{a_1}}$	0.525	0.067	7.799	0.393	0.657	0.000
$\sigma_{v_{\eta_1},v_{\eta_2}}$	1.128	0.365	3.087	0.412	1.844	0.001
$\sigma^2_{v_{a_1}}$	0.058	0.007	8.646	0.045	0.071	0.000
$\sigma^2_{v_{\eta_2}}$	2.195	0.196	11.222	1.811	2.578	0.000
$\sigma^2_{v_{a_2}}$	0.001	0.000	5.706	0.001	0.001	0.000

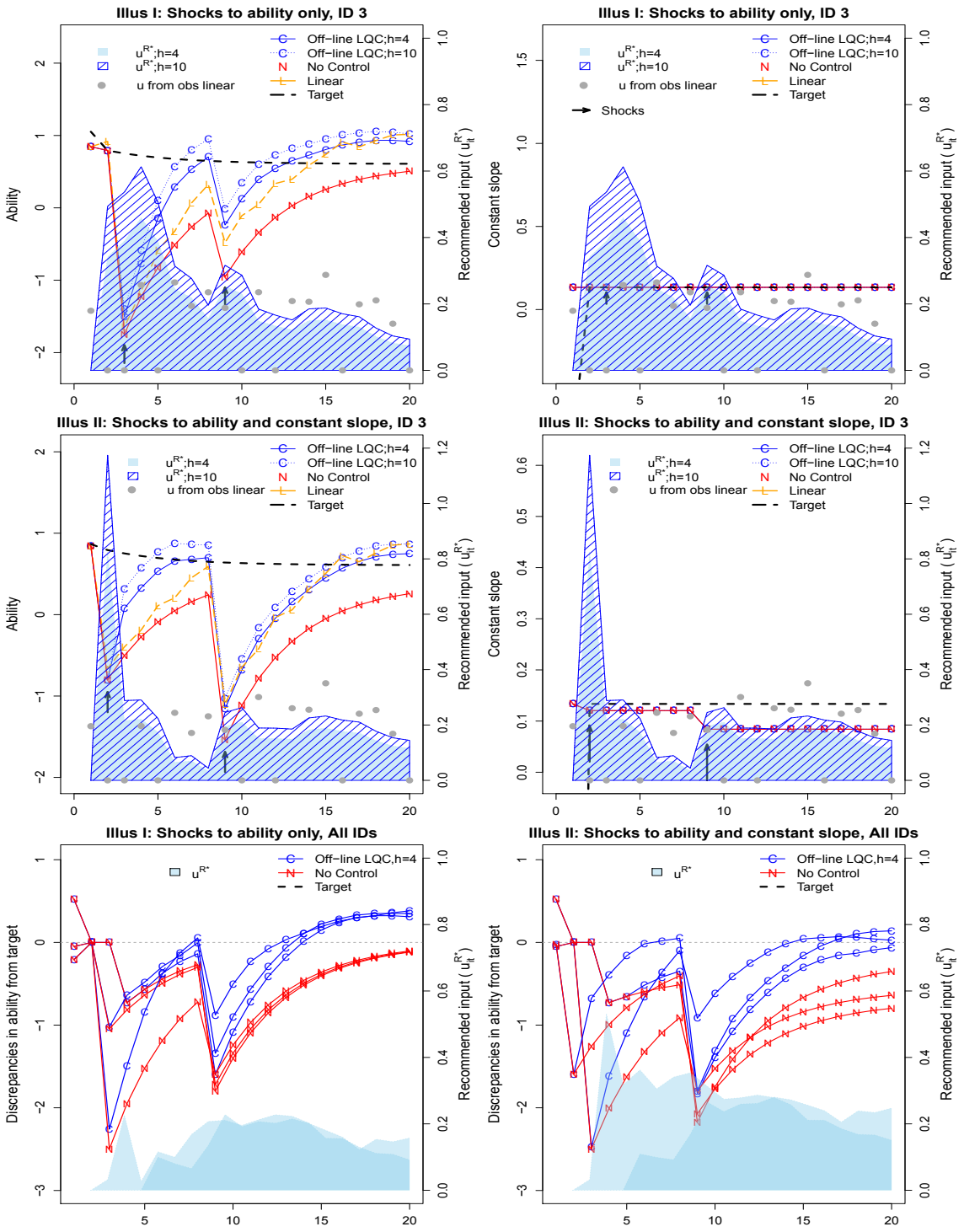

Figure 1. Top row: Simulated trajectories from Illustration I, with shocks to the latent ability levels, η_{1it} (left plot), but not the constant slope, a_{1i} (right plot). Bottom row: Simulated trajectories from Illustration II, with shocks to the latent ability levels, η_{1it} (left plot), but also the constant slope, a_{1i} (right plot). The plots depict the participants' latent ability levels, target trajectories, shock points, and trajectories with and without use of the control input, u_{it}^{R*}. In the plots of the model-implied ability scores for ID 3 (first two rows, column 1), we also added the trajectories and recommended durations based on the observed linear scheme ('Linear'). Note that two ordinates are used on the left and right sides of the plot to better reflect the scales of the latent variable and control input, respectively.

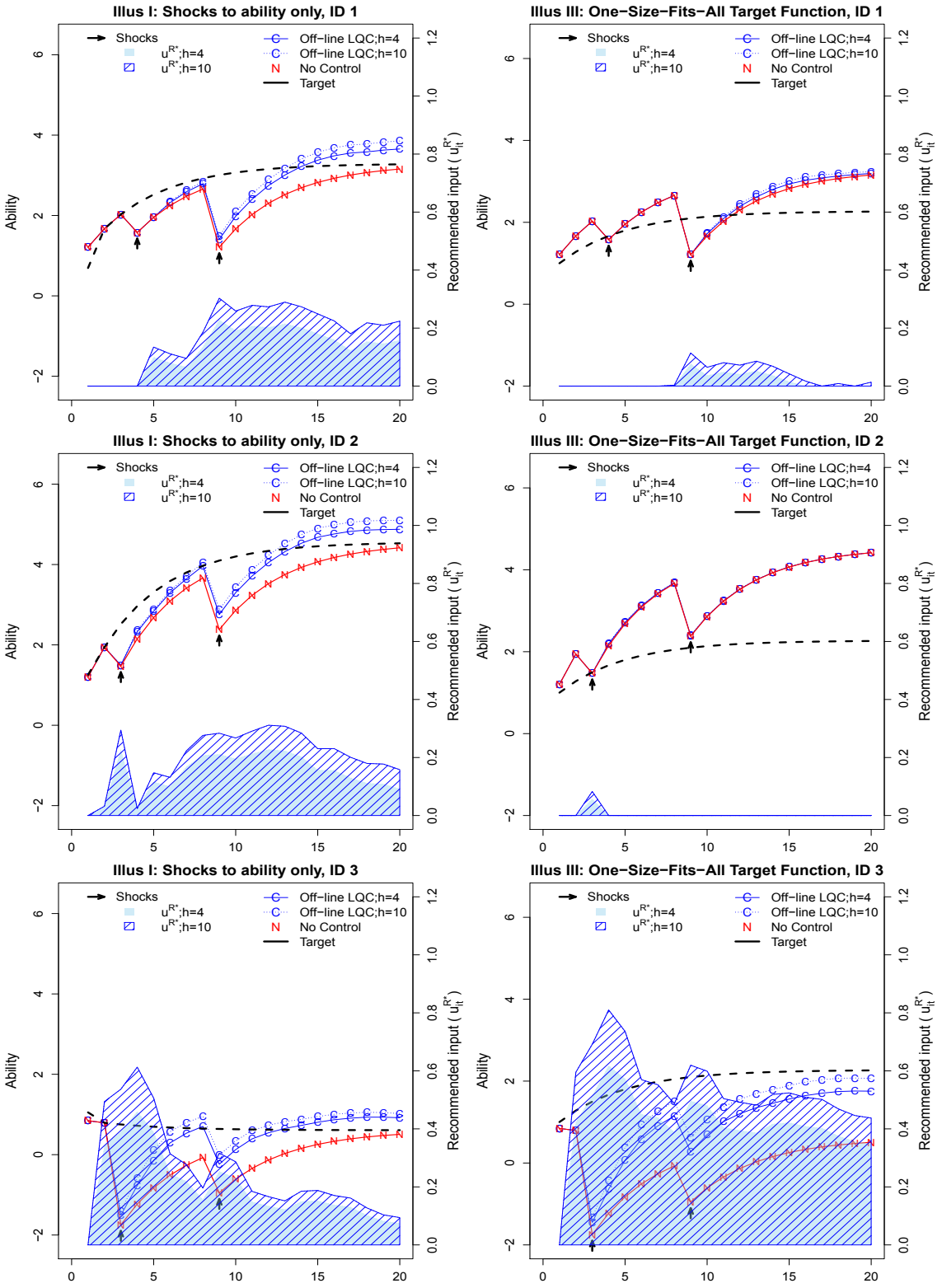

Figure 2. Simulated trajectories from Illustration I with person-specific target trajectories (left panel), and Illustration III with a person-invariant, one-size-fits-all target trajectory (right panel) plotted on the same scales.

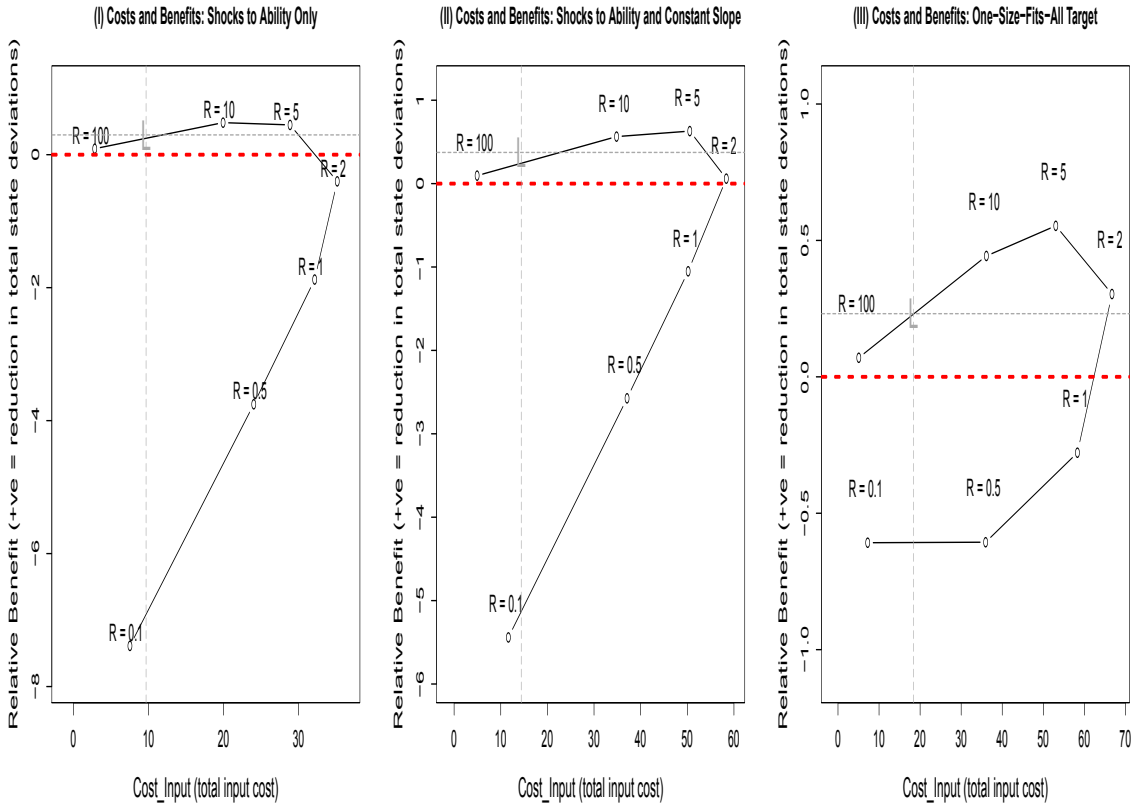

Figure 3. The total quadratic costs associated with administration of the off-line LQC input, Cost_Input, plotted against the relative reduction in total state deviations, Relative_Benefit, under Illustration I, with shocks to individuals' latent ability levels only (left); under Illustration II, with shocks to both their latent ability as well as constant slope levels (middle); and under Illustration III, with shocks to indviduals' latent ability levels only, but with a one-size-fits-all target function as defined by the fixed effects curve. In all plots, only the off-line LQC costs and benefits under a control horizon window of $h = 4$ are shown. The total input costs and relative benefits associated with the linear observed scheme are shown as light-shaded, vertical dashed and horizontal dotted lines, respectively, marked with the symbol 'L').

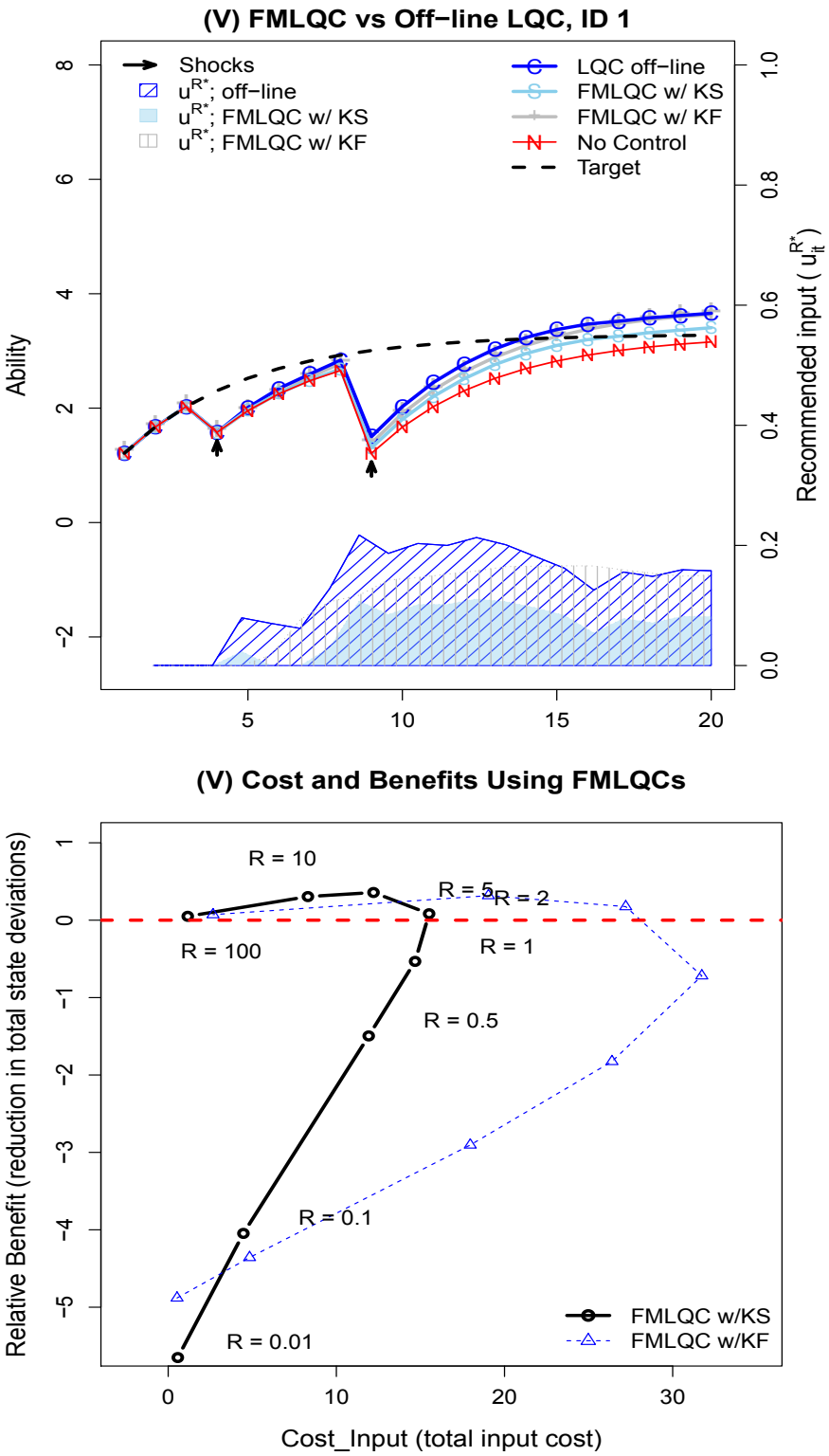

Figure 4. (Top) State trajectories and control inputs computed off-line in comparison to using KF- and KS-based FMLQCs in Illustration V. (Bottom) The corresponding total quadratic costs associated with input administration, Cost_Input plotted against the relative reduction in total state deviations, Relative_Benefit.

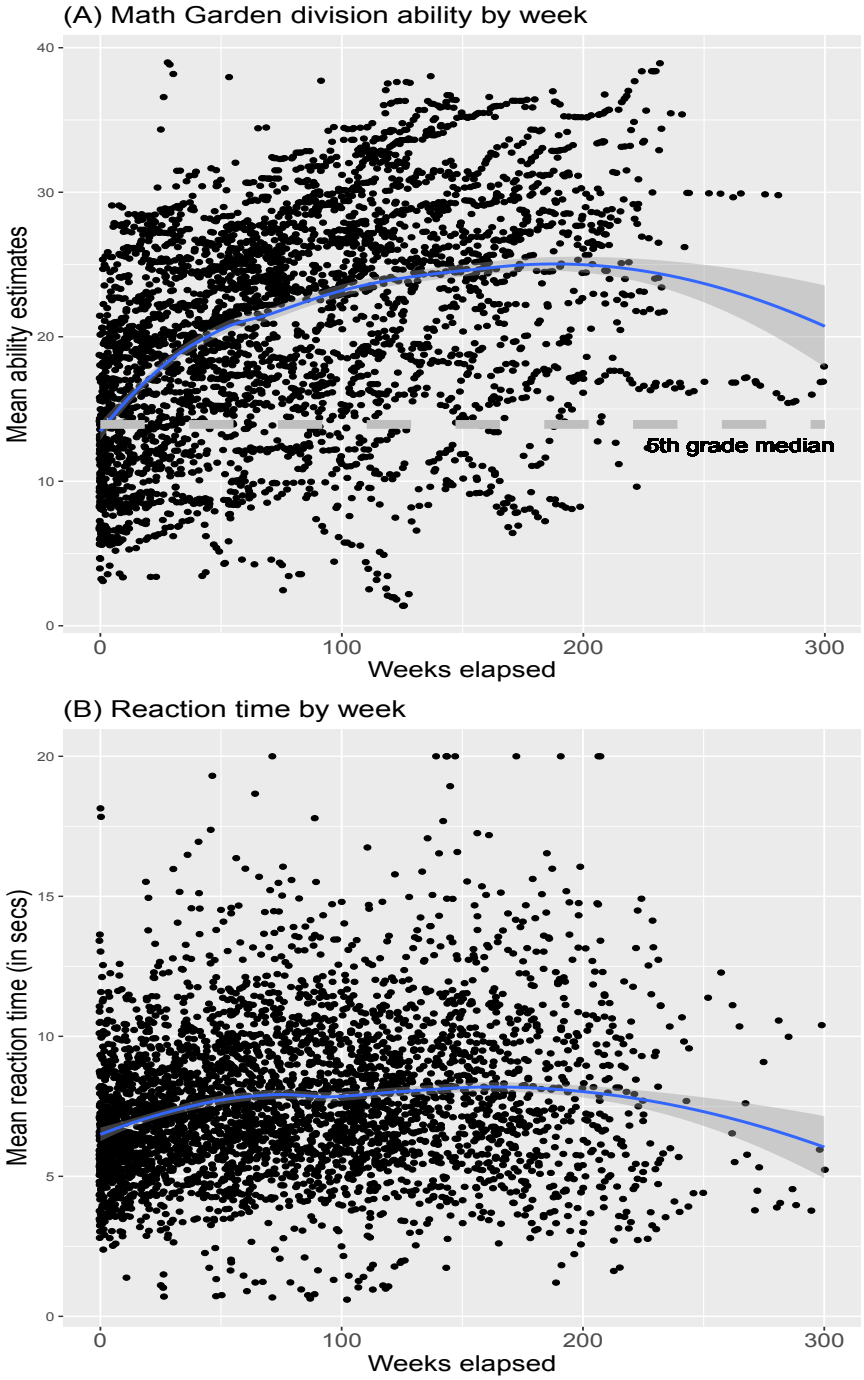

Figure 5. Plots of (A) Average weekly division ability estimates from 100 randomly selected Math Garden users; (B) Average weekly per-item reaction time on the division task for these randomly selected users. In each plot, the thick solid line with shaded region is the smoothed loess curve and its corresponding 95% confidence intervals. The Math Garden time limit for the division items was 20 seconds.

Empirical Illustration I: LQC vs. Fixed-Duration Training Scheme

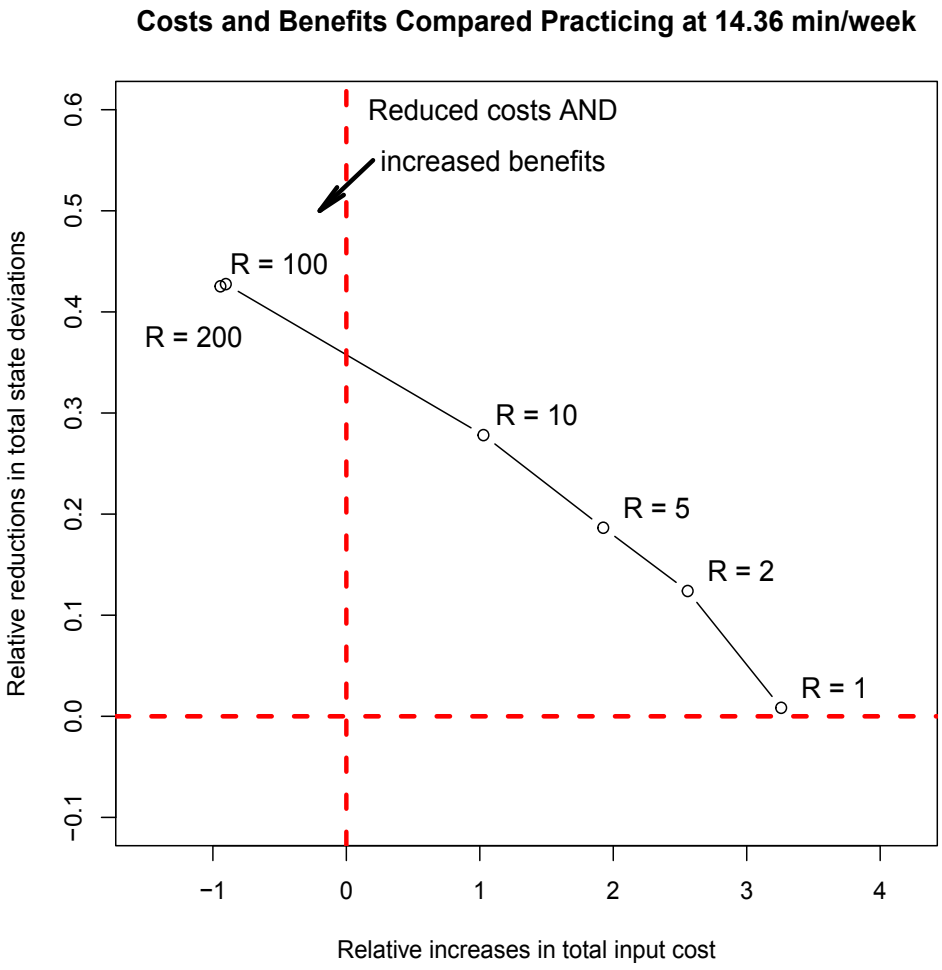

Figure 6. A summary of the cost and benefit values of empirical illustration I. The abscissa (horizontal axis) depicts the relative increases (+ indicates increased costs; - indicates reduced costs) in total input costs of the LQC-recommended training durations compared to the total input costs associated with practicing at a constant scheme of 14.36 minutes per week across different values of R. The ordinate (vertical axis) shows the corresponding relative reductions (+ indicates reduction; - indicates increase) in total state deviations under the LQC-recommended as compared to the fixed-duration training scheme.

Empirical Illustration II: LQC vs. Original Training Durations

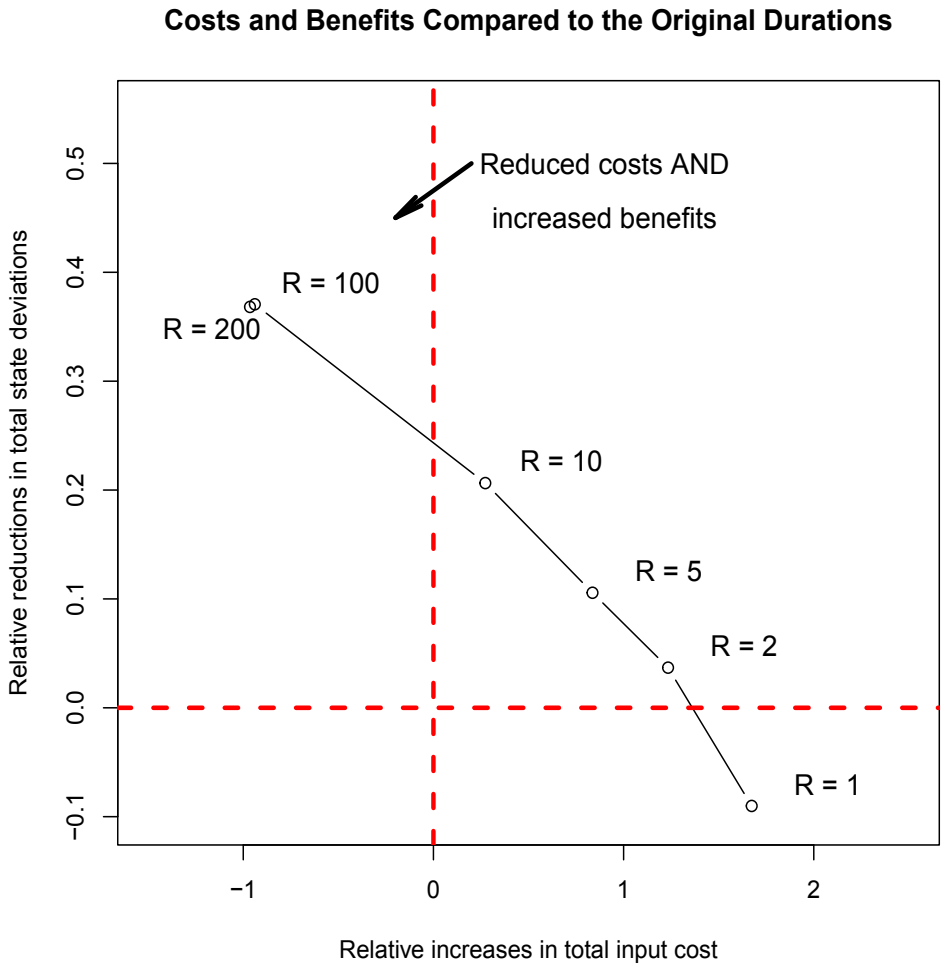

Figure 7. A summary of the cost and benefit values of empirical illustration II. The abscissa (horizontal) axis depicts the relative increases (+ indicates increased costs; - indicates reduced costs) in total input costs of the MHE-LQC-recommended training durations compared to the total input costs associated with the participants' original observed training durations across different values of R. The ordinate (vertical) axis shows the corresponding relative reductions (+ indicates reduction; - indicates increase) in total state deviations under the LQC-recommended as compared to the constant training schemes.

Empirical Illustration III: Hybrid Target Function

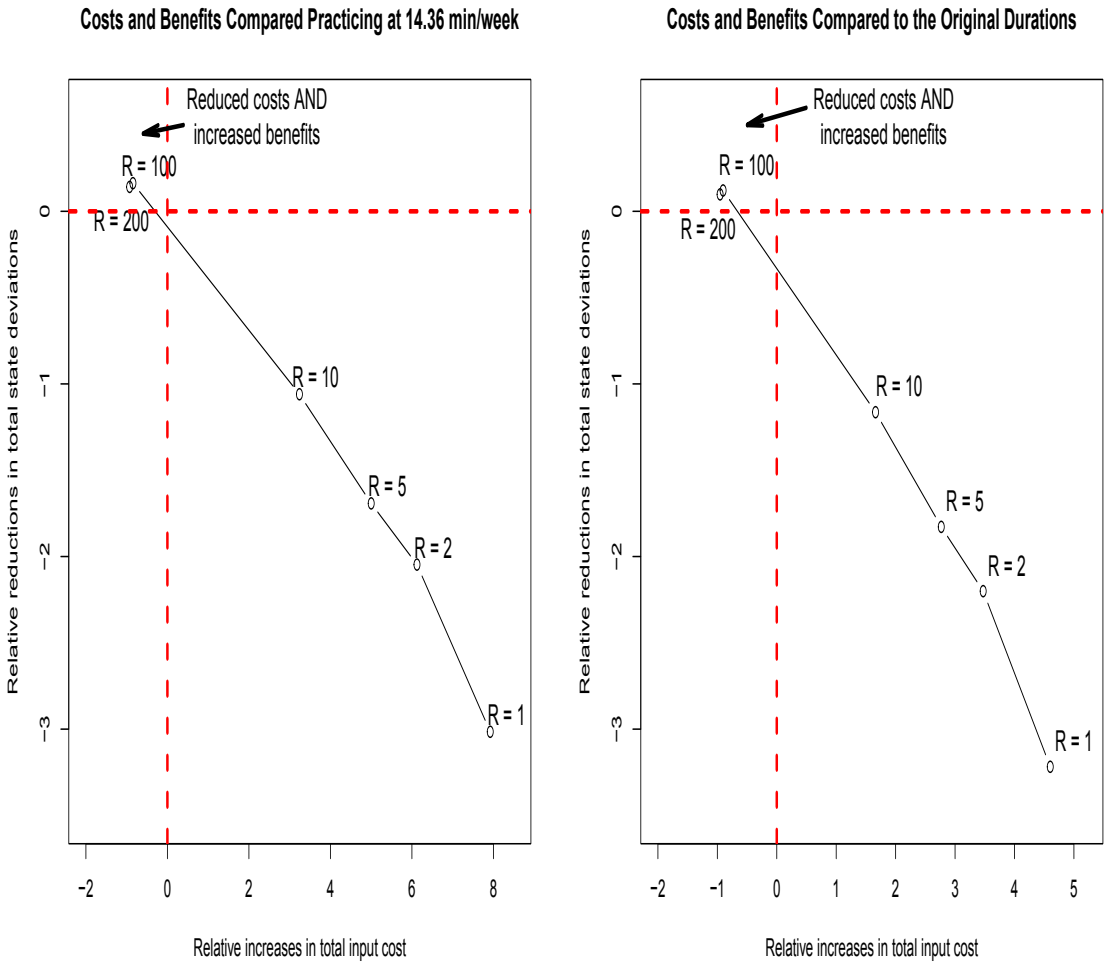

Figure 8. A summary of the cost and benefit values of empirical illustration III, with hybrid target function that integrates grade-norm median and person-specific change information. The left panel summarizes the costs and benefits under the KS-based FMLQC training scheme compared to the constant-duration training scheme; the right panel plots the cost and benefit comparisons under the KS-based FMLQC training scheme as compared to the observed training scheme.

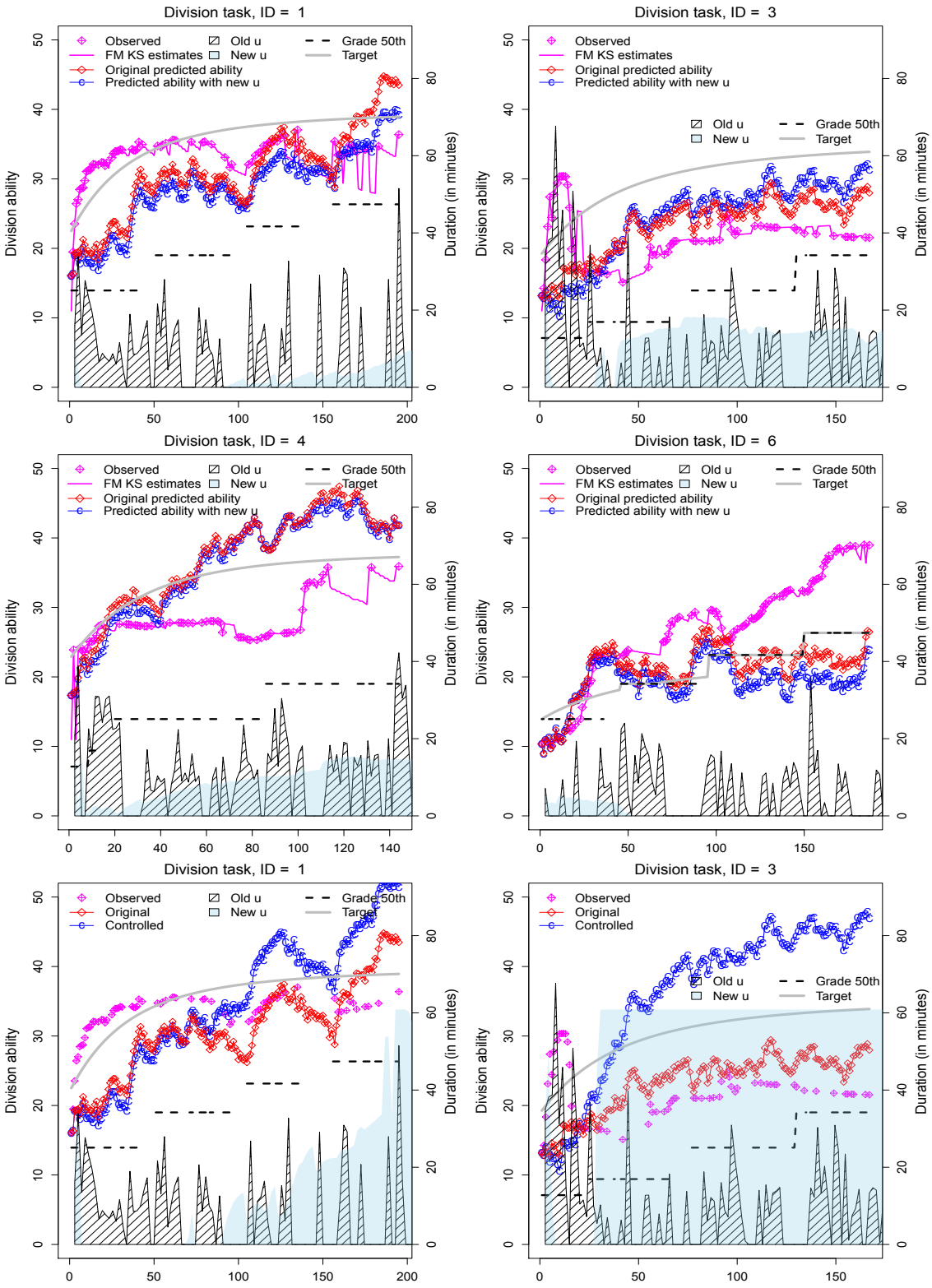

Figure 9. Top and middle rows: Results from applying the FMLQC with KS state estimates at $R = 100$ to the Math Garden data. Bottom row: Corresponding results at $R = 10$ to the Math Garden data. Predicted ability refers to model-implied ability trajectories generated in the absence of process noises. Observed = observed ability scores; FM KS = latent ability estimates from the KS-based FMLQC; Old u = original training durations; New u = KS-based FMLQC-recommended training durations, u^{R*}; Grade 50th = grade median scores; Target = Reference target, η^r.