On the Origin of the Matthew Effect: Insights from a Quantitative Theoretical Model

Matthew T. McBee[1], Matthew C. Makel[2], & Natasha L. Godkin[1]

[1] East Tennessee State University

[2] Duke University Talent Identification Program

Author Note

Correspondence concerning this article should be addressed to Matthew T. McBee.

E-mail: mcbeem@etsu.edu

Abstract

A quantitative theoretical model of academic achievement, based on the concept of the Zone of Proximal Development (ZPD), has proven capable of reproducing the results of several empirical, longitudinal studies of achievement. This paper applies the model to understanding the origins of variance in achievement (at the individual and subgroup mean level of analysis), and the effect of gifted education initiatives on that variability.

Keywords: gifted, excellence gap, zone of proximal development, theoretical calculation

Word count: X

On the Origin of the Matthew Effect: Insights from a Quantitative Theoretical Model

Theoretical Framework

Current educational and psychological theories, being qualitative, are incapable of generating precise quantitative predictions – "point predictions" – that are the hallmark of the physical sciences (c.f. Cohen, 1994; Maxwell & Howard, 2009; Meehl, 1967). For theories to be falsifiable, they must make specific predictions. Only then can empirical research provide the type of severe tests (Mayo, 1996) that lead to meaningful scientific progress. We propose such a theory in this paper. Our theory of academic achievement growth can describe and predict the achievement trajectories of individual students, accounting for both individual and environmental factors.

Methods1

The Zone of Proximal Development (ZPD) is the basic psychological principle at the heart of this theoretical model. Although the ZPD concept is typically attributed to Lev Vygotsky, the phenomenon was actually first revealed by Dorthea McCarthy (1930). The ZPD describes how learning emerges from the interaction between a student's base of currently mastered skills and knowledge with a particular set of learning experiences, and refers specifically to the set of skills that a student cannot perform autonomously but can do with scaffolding from a teacher or more competent peer. In our theory, instruction that occurs within the ZPD leads to the most rapid, efficient learning.

Academic achievement

We define the term academic achievement to mean a student's location along a sequenced, developmental curriculum within a specific domain. Academic achievement therefore refers to the body of knowledge, understanding, and skills that a student has mastered at a particular moment in time. Academic growth or achievement growth is change (typically positive) in achievement over time. Our numeric scale for measuring achievement ranges from zero to positive infinity, where zero is a meaningful quantity representing no knowledge or skills in the domain. Specific values of achievement could refer to the location of specific skill milestones or content mastery. For example, one could (arbitrarily) set the mastery of mathematical operations expected of a typical third grader to a value of 0.40 on the achievement metric. ZPD. The ZPD is represented by a normal curve, normalized to a maximum value of one, whose peak is slightly ahead of the student's current achievement (Figure 1). In the figure, a vertical line at 0.25 depicts the student's current achievement; her ZPD is the shaded curve peaking at 0.30. The ZPD is not a single point but rather a "spread" of values varying in intensity; its peak describes the most efficacious level of instruction for stimulating academic growth.

Insert Figure 1 here

School curriculum.

School curriculum is co-located on the same scale with academic achievement in much the same way that IRT models place test scores and item difficulties on a common scale. For example, one could imagine that a typical Kindergarten reading curriculum ranges from 0.05 to 0.15, where 0.05 might include basic letter and phoneme awareness and 0.15 a limited ability to "sound out" simple words and knowledge of roughly 200 sight words. Within this academic year, the curriculum presented on a specific day would consist of a narrow slice or segment of this interval, and as the year progresses, the presented curriculum would steadily advance. Our model uses a normalized trapezoidal distribution to represent the school curriculum due to the many shapes it can assume with different choices of values for its parameters. Figure 2 illustrates how three students, with varying initial achievement, can experience the same school curriculum. Only the student in the center panel would derive strong benefit from exposure to this curriculum as it falls within his ZPD. The curriculum is too difficult for the first student and too easy for the third; neither of these students would experience much benefit. Our implementation of the model in our R package does allow for students to experience different curricula (e.g., remedial, typical, or advanced) on an individual basis.

Insert Figure 2 here

Home curriculum

Children learn a great deal at home, particularly in early childhood and in certain subjects such as reading. The rate of academic growth caused by learning experiences at home is a function of the overlap between the child's ZPD and what we call the "home curriculum", which we conceptualize as a downward-sloping function. Our R package implantation represents the home curriculum as a normalized beta distribution with the α parameter fixed to 1. Figure 3 displays the home curriculum function (with parameter β set to 5) with a superimposed ZPD for a child with achievement = 0.20. The implication of this downward slope is that academic growth due to home instruction is most rapid in early childhood and becomes increasingly reliant on school instruction in later grades.

Insert Figure 3 here

Individual difference parameters

Our model incorporates four child-level parameters: current achievement, learning rate, decay rate, and home environment; these create variability in growth trajectories. The learning rate represents individual differences in the pace of academic growth for students with identical levels of ZPD engagement, and can be considered to be a combination of general intelligence and motivation. The decay rate is the rate at which achievement will decay (due to forgetting) in the absence of instruction. The home environment parameter describes how effectively the home environment is at facilitating academic growth.

Model dynamics

The amount of learning during discrete time interval t for individual i is given by

$$
\begin{aligned}
Learning_{ti} = LearnRate_i * \\
\Big[\Big(dosage_{ti} * \int_0^\infty ZPD_{ti}(x) * S_t(x)\,dx\Big) + \\
(1 - dosage_{ti}) * HomeEnv_i * \int_0^\infty ZPD_{ti}(x) * H_t(x)\,dx\Big)\Big]
\end{aligned}
\tag{1}
$$

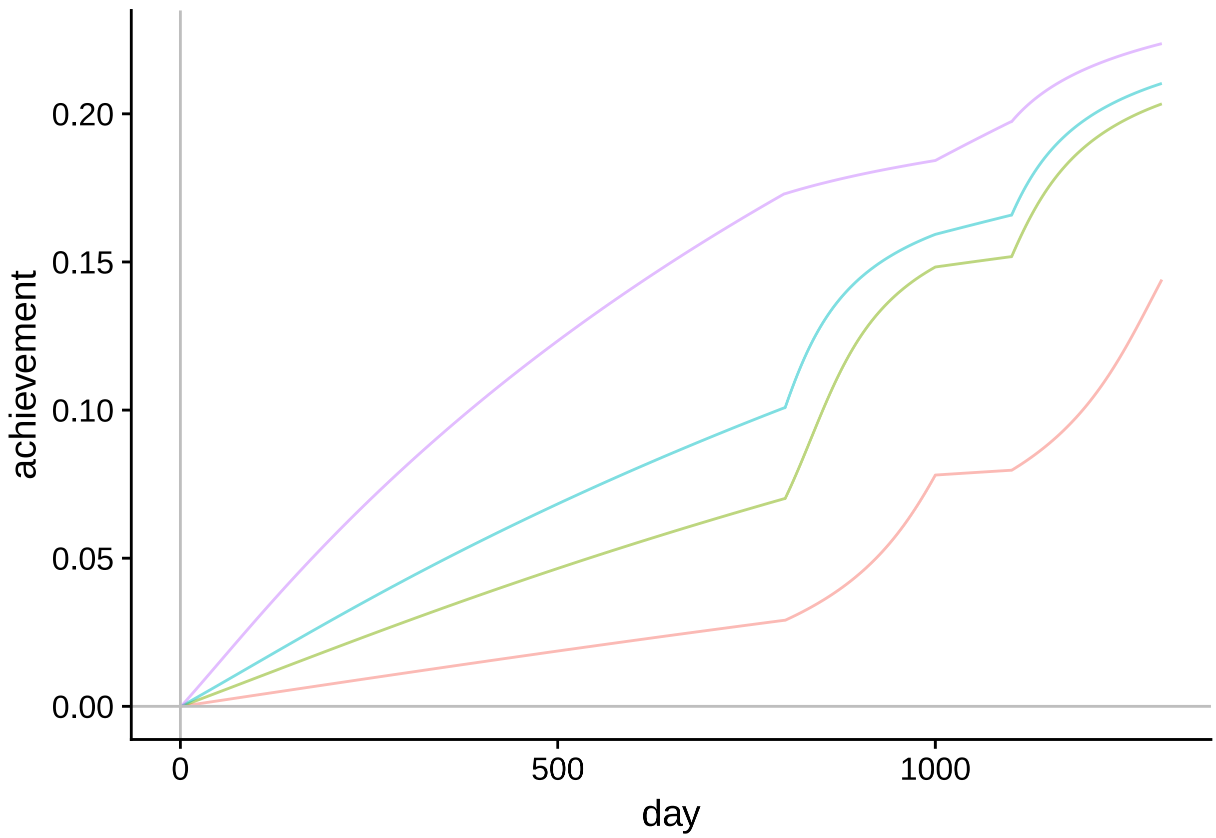

Figure 1

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

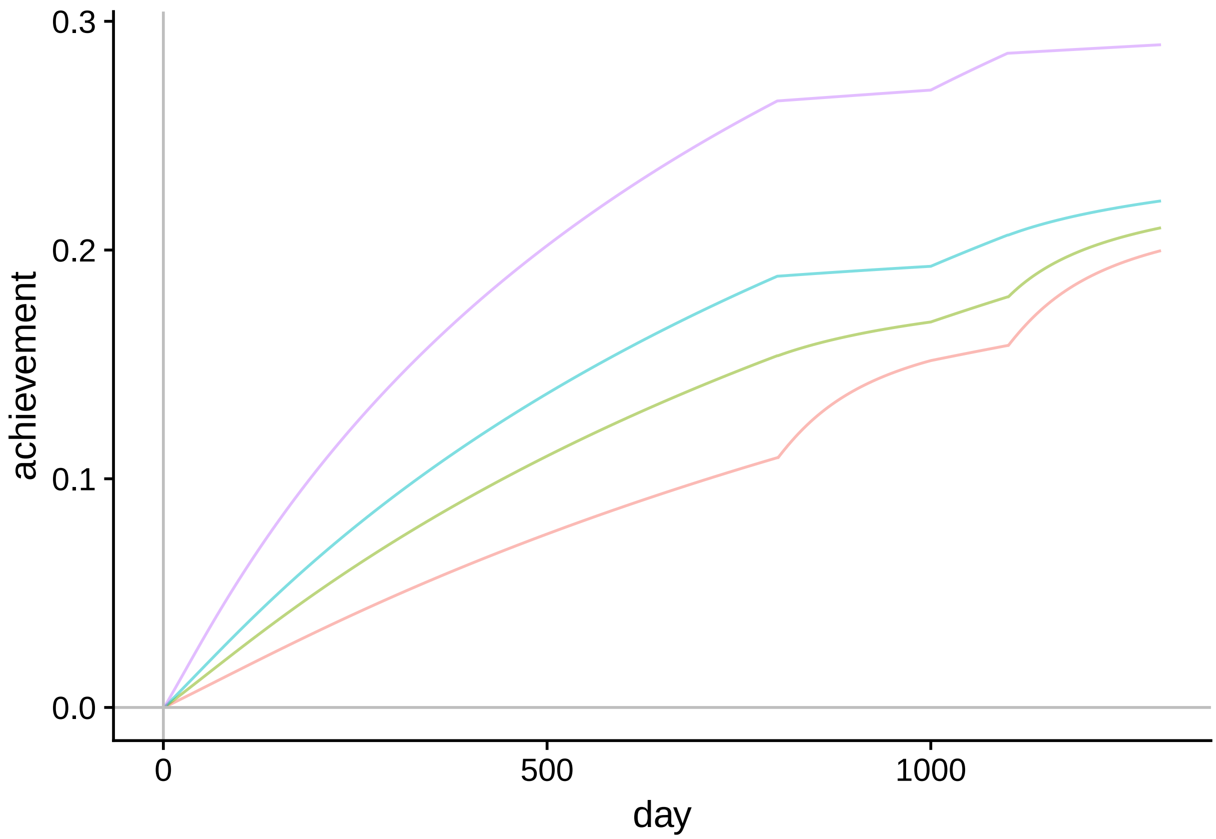

Figure 2

Participants

Material

Procedure

Data analysis

We used R (3.4.0, R Core Team, 2017) and the R-packages *cowplot* (0.9.2, Wilke, 2017), *dplyr* (0.7.4, Wickham, Francois, Henry, & Müller, 2017), *ggplot2* (2.2.1, Wickham, 2009), *papaja* (0.1.0.9492, Aust & Barth, 2017), and *ZPDGrowthTrajectories* (0.0.5, McBee, 2017) for all our analyses.

Results

Discussion

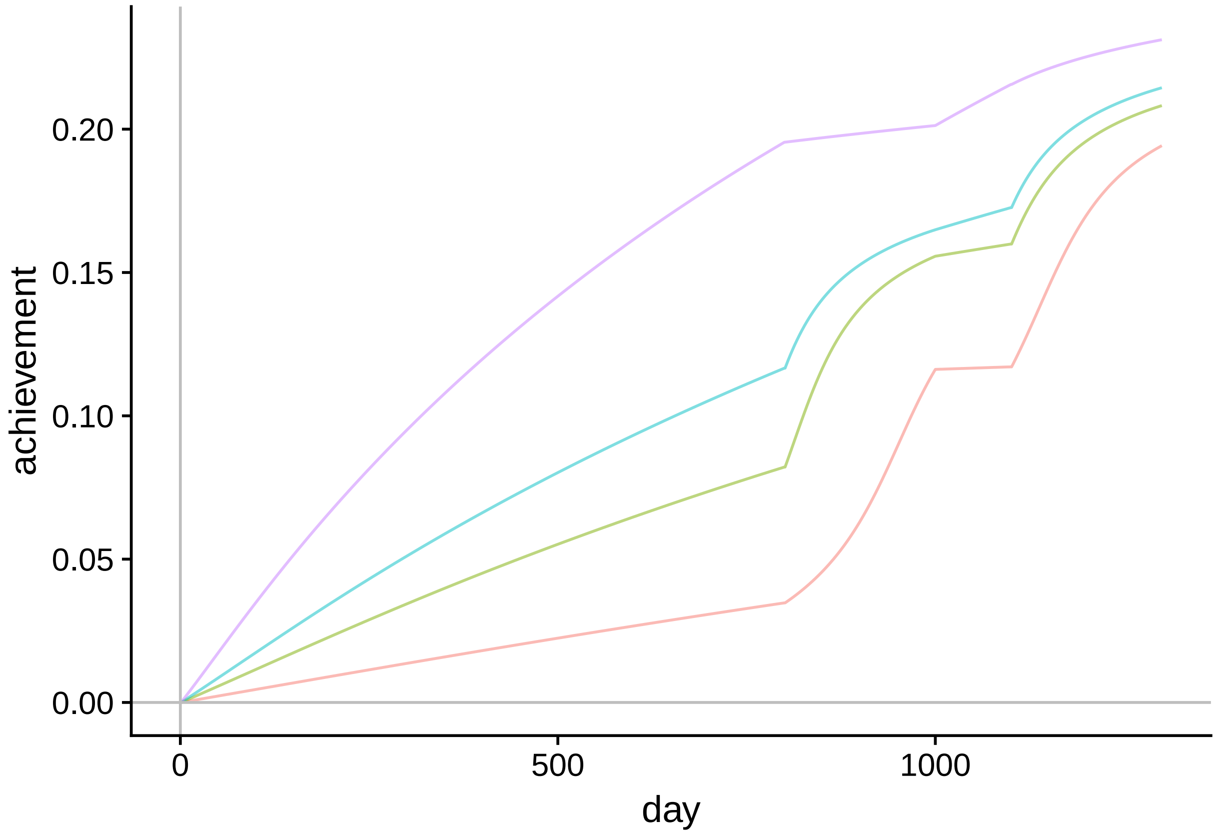

Figure 3

References

Aust, F., & Barth, M. (2017). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

McBee, M. (2017). *ZPDGrowthTrajectories: Simulates achievement growth trajectories using a theoretical model of academic achievement.*

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. Retrieved from http://ggplot2.org

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *Dplyr: A grammar of data*

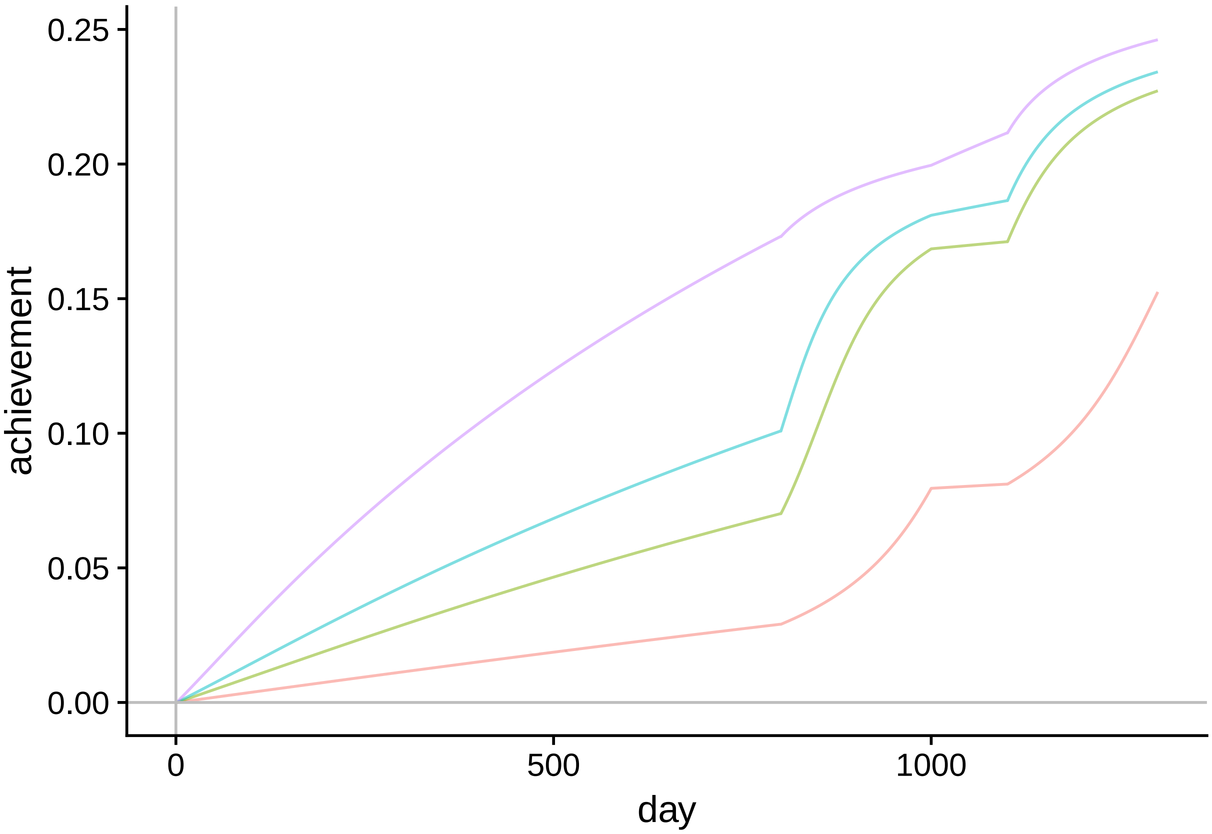

Figure 4

manipulation. Retrieved from https://CRAN.R-project.org/package=dplyr

Wilke, C. O. (2017). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=cowplot

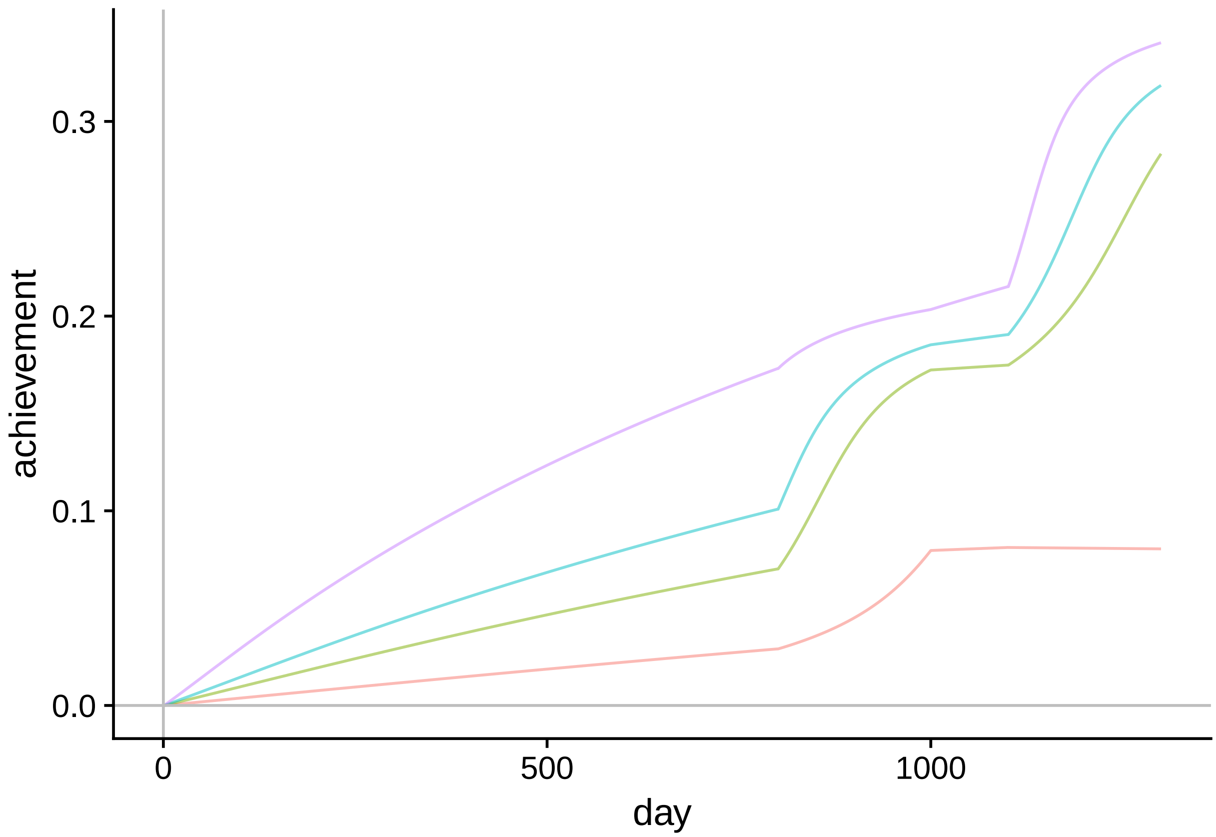

Figure 5