Theory vs Practice of Mastery Learning in the Cognitive Tutor: Principal Stratification on a Latent Variable

Adam C Sales John F Pane

1 Background / Context

The Cognitive Tutor (Anderson, Corbett, Koedinger, and Pelletier, 1995) is a piece of software designed to teach math, alongside traditional teachers. In the second year of a large-scale effectiveness study of the Cognitive Tutor Algebra I (CTAI) curriculum, the intervention had a moderate positive effect on high school post-test scores (Pane, Griffin, McCaffrey, and Karam, 2014). One of CTAI's mechanisms is "mastery learning" (Bloom, 1968): the software estimates students' skill mastery after each worked problem, and (ideally) only advances them to the next section after they have mastered all of the current section's skills. When each student advances at his or her own pace, academically diverse students can learn together in the same classroom. What role did mastery learning play in CTAI's successes? In practice, students will sometimes exhaust all of the problems in a section without mastering its skills, in which case they are "promoted" to the next section. Did students who were more likely to master worked sections also experience larger treatment effects?

Student mastery behavior is a variable defined subsequent to treatment assignment—students in the control condition do not use the software and therefore have no mastery data. A principal stratification analysis (Frangakis and Rubin, 2002; Page, 2012) could estimate the variance of treatment effects as a function of *potential* mastery behavior: how often students would master worked sections, were they assigned to treatment. Implicitly, this assumes that mastery behavior is measured without error, yet there are no error-free measurements of students' propensity to master sections. Further complicating matters, both the number of worked sections and which sections students worked varied widely between students in the treatment group. The typical principal stratification approach, assuming intermediate variables measured without error, may yield misleading or uninterpretable results when applied to mastery learning in CTAI.

2 Purpose / Objective / Research Question

This talk will address the substantive research question—what is the role of mastery learning in CTAI's success?—by also addressing the methodological question—principal stratification with a latent variable. We use item response theory (IRT; e.g. Embretson and Reise, 2013)

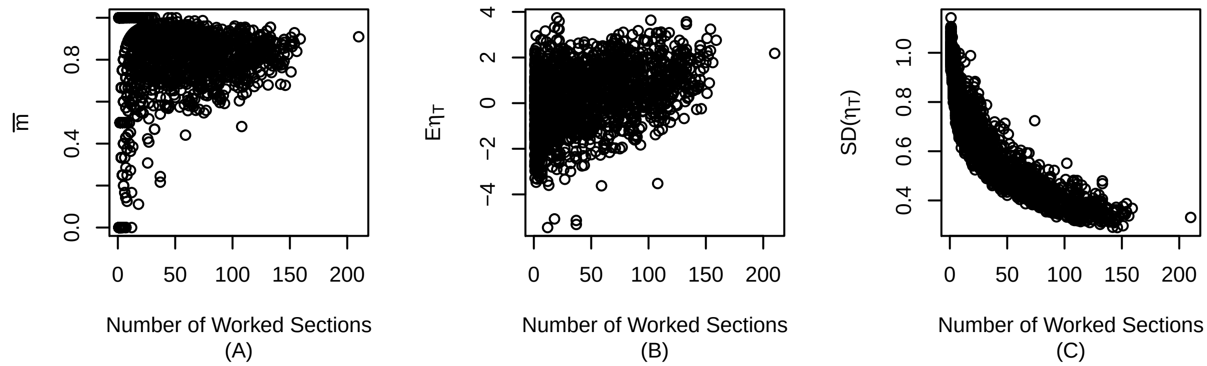

Figure 1: Measurement error in $\bar{m}$ and its solution: (A) plots $\bar{m}$ against the number of sections students worked, (B) plots means of a model-based measurement of students' propensity to master worked sections, and (C) plots the standard deviation of that measurement.

to measure student mastery potential as a latent variable within a principal stratification model. The result is a model that estimates the association between mastery and treatment effects, answering the question: do students who are more likely to master worked sections experience greater benefits?

3 Methodology: Principal Stratification and a Model for Mastery

Consider two measurements of students' propensity to master worked sections: $\bar{m}$ is the proportion of worked sections that students mastered, and η is the random "student ability parameter" from a Rasch model with section mastery as an outcome and section fixed effects. Model based measurement has at least three advantages over $\bar{m}$: first, $\bar{m}$ (Figure 1A), but not η (B), is strongly associated with the number of sections that students work . Second, the Rasch model automatically estimates the standard deviation of measurement for individual students' measurements η (C). Third, η accounts for the fact that different students work different sections.

To relate student mastery to treatment effects in a principal stratification context, write η, instead, as η_T—students' propensity to master worked sections *if assigned to the treatment condition*. Students in the control group, who had no access to CTAI, could never work or master sections, rendering the corresponding η_C irrelevant. To infer counterfactual η_T for students in the control group, we modeled η_T in the treatment group as linear in a set of pre-treatment covariates including pretest scores and demographics, with normal regression errors. Since assignment to CTAI was randomized, this model extrapolates to the control group, yielding a predictive distribution of η measurements for students in the control group. Finally, we modeled students' post-test scores as linear in covariates,treatment assignment, and latent mastery η_T, with normally-distributed errors. We fit the model in R and JAGS

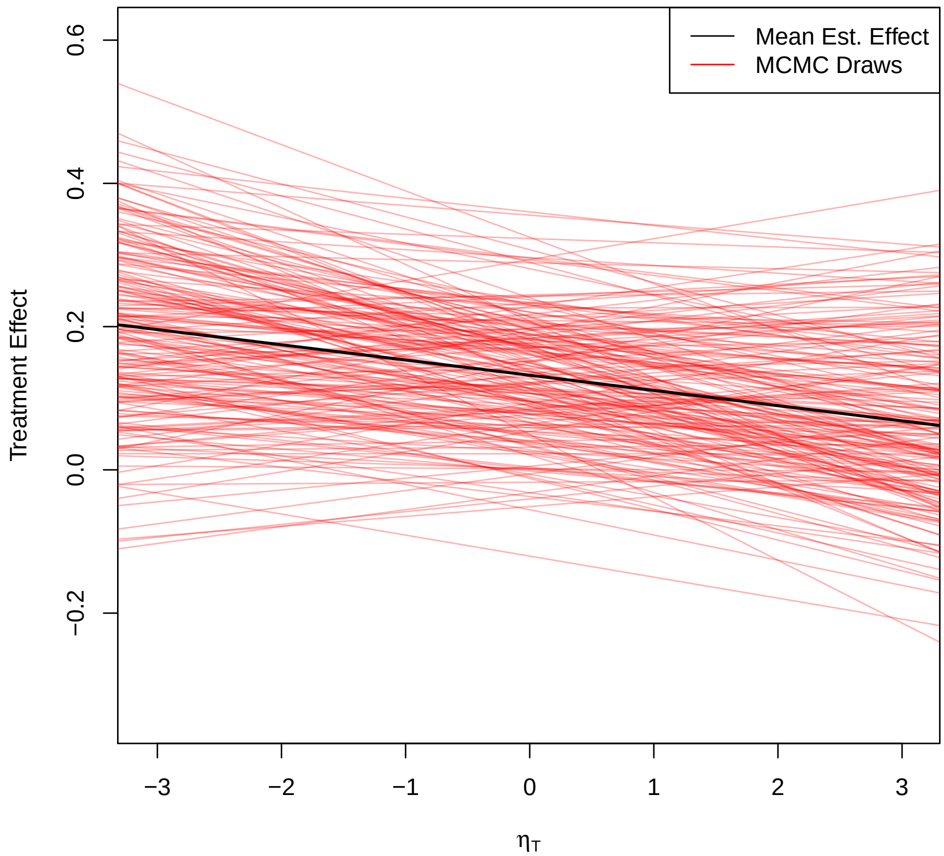

Figure 2: The estimated treatment effect as a function of students' propensity to master a section, $\mathbb{E}[Y_T - Y_C|\eta_T]$. Red lines are draws from the posterior distribution of the treatment effect function, and the black line is the mean of the posterior.

(Plummer, 2016; R Core Team, 2016) with weakly-informative priors.

4 Findings / Results

Figure 2 displays the posterior mean and posterior draws for the estimated function $\tau(\eta_T)$. In approximately 81% of of the MCMC runs, the slope of $\tau(\eta_T)$ was negative, implying that students with a higher potential to master sections that they work tended to experience lower treatment effects. These model results do not allow us to rule out a positive slope for $\tau(\eta)$, but they do allow us to rule out particularly large positive slopes. It is evident that mastery learning, as measured by η_{Ti}, is not the primary driver of the CTAI effect.

5 Conclusions

Theses results seem to challenge the theory underlying the Cognitive Tutor. If mastery learning were a central driver of the CTAI effect, we would expect the treatment effect to be larger for students who are more likely to achieve mastery. In fact, the opposite seems to be the case—either average effects decrease with students' mastery propensity, or they are nearly unrelated. On the other hand, the latent parameter η_T is a baseline student characteristic, that correlates with other baseline variables. If the CTAI effect were larger for struggling or lower performing students than for stronger students, we would expect to see a negative correlation between η_T and treatment effects. That said, the results here downplay the contribution of mastery learning, relative to other factors, in driving CTAI's effectiveness. However, these theoretical failures do not imply a failure in practice. In particular, struggling students, who are less likely to achieve mastery, are also most in need of help. The results here suggest that the Cognitive Tutor is not failing them.

Operationalizing students' potential mastery via the Rasch parameter η_T has clear advantages over the simpler approaches previously available. That said, it poses some philosophical challenges, which we plan to discuss. Latent variable principal stratification has the potential to merge the fields of psychometrics and causal inference, leading to more and more nuanced scientific discoveries.

References

John R Anderson, Albert T Corbett, Kenneth R Koedinger, and Ray Pelletier. Cognitive tutors: Lessons learned. *The journal of the learning sciences*, 4(2):167–207, 1995.

Benjamin S Bloom. Learning for mastery. instruction and curriculum. regional education laboratory for the carolinas and virginia, topical papers and reprints, number 1. *Evaluation comment*, 1(2):n2, 1968.

Susan E Embretson and Steven P Reise. *Item response theory for psychologists*. Psychology Press, 2013.

Constantine E Frangakis and Donald B Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.

Lindsay C Page. Principal stratification as a framework for investigating mediational processes in experimental settings. *Journal of Research on Educational Effectiveness*, 5 (3):215–244, 2012.

John F Pane, Beth Ann Griffin, Daniel F McCaffrey, and Rita Karam. Effectiveness of cognitive tutor algebra i at scale. *Educational Evaluation and Policy Analysis*, 36(2): 127–144, 2014.

Martyn Plummer. *rjags: Bayesian Graphical Models using MCMC*, 2016. URL `https://CRAN.R-project.org/package=rjags`. R package version 4-6.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org/`.