charles river analytics

Charles River Analytics Inc.
625 Mount Auburn St.
Cambridge, MA 02138

www.cra.com

p: 617.491.3474
f: 617.868.0780

15 SEP 2009

Reference: Government Contract No. N00014-09-C-0050, "Enhancing Simulation-based Training Adversary Tactics via Evolution (ESTATE)"
Charles River Analytics Contract No. C08098

Subject: Contractor's Status Report: Quarterly Status Report #3
Reporting Dates: 6/15/2009 – 09/15/2009

Dear Dr. Hawkins,

The following is the Contractor's Quarterly Status Report for the subject contract for the indicated period. During this reporting period work has concentrated on Task 3: Enhance Adaptation Techniques and Task 4: Develop Trainee Model Processing.

1. Summary of Progress

During this reporting period, we extended our work in the conceptual formulation of ESTATE for Challenge / Response games. We explored and refined aspects of item response theory to support assessment and learning Challenge / Response games. We constructed multiple simulations to test our theoretical hypotheses.

1.1 Background - Item Response Theory

We employed conceptual framework in Figure 1 and applied Item Response Theory (IRT) (Baker, 2001) to help provide a computational foundation.

1.1.1 Item Characteristic Curve

In IRT, *ability* is used to represent and measure latent traits in individuals performing a function. We represent this term by θ. While θ can range from positive infinity to negative infinity, it is typically given a -3 to 3 range. For each item (or challenge), an individual has a probability of getting the item correct or incorrect. This probability is represented by P(θ). Since P(θ) is a function of θ, we can construct an *item characteristic curve* (ICC) that represents the probability of getting an item correct as a function of an individual's ability level. These ICCs are normally S-curves. The shape of these S-curves can be defined by several mathematical models. The *difficulty* of an item is a location index that describes where the item functions along the ability scale. For our purposes, this can be where is P(θ) = 50%. The *discrimination* of an item describes how well the item can differentiate between examinees having abilities below the item location and those having abilities above the item location (essentially the steepness of the ICC in the

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
15 SEP 2009		**00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Enhancing Simulation Based Training Adversary Tactics via Evolution (ESTATE)?	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Charles River Analytics Inc,625 Mount Auburn St,Cambridge,MA,02138**	8. PERFORMING ORGANIZATION REPORT NUMBER **Topic #N08-109**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

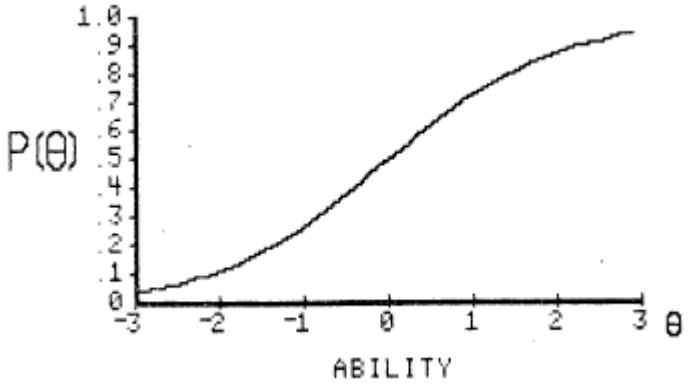
13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**7**	**14**	

middle, or the slope of the line where P(θ) = 50%). The *guessing* of an item describes how likely it is that an examinee will guess the answer correctly.



The equation for the three parameter ICC (Baker, 2001) is:

$$P(\theta) = c + (1-c)\frac{1}{1+e^{-a(\theta-b)}}$$

Where: b is the difficulty parameter
a is the discrimination parameter
c is the guessing parameter and
θ is the ability level

Note that in simulation, a response may be generated from this equation by setting a response value r such that: $r = P(\theta) < U(0,1)$; where $U(0,1)$ is a random number from the uniform distribution between 0 and 1, inclusive.

The single parameter model, or Rasch model, defined as the above ICC with a=1.0 and c=0.

1.1.2 Estimating an Examinees Ability

Given a set of ICCs and a history of results for an examinee, it is possible to estimate the examinees ability. The estimation equation for maximum likelihood is:

$$\hat{\theta}_{s+1} = \hat{\theta}_s + \frac{\sum_{i=1}^{N} a_i\left[u_i - P_i(\hat{\theta}_s)\right]}{\sum_{i=1}^{N} a_i^{\,2} P_i(\hat{\theta}_s) Q_i(\hat{\theta}_s)}$$

where: $\hat{\theta}_s$ is the estimated ability of the examinee at iteration *s*

a_i is the discrimination parameter of item i

u_i is the response made by the examinee to item i: 1 for correct, 0 for incorrect

$P_i(\hat{\theta}_s)$ is the probability of a correct response to item i under the given item characteristic curve.

$Q_i(\hat{\theta}_s) = 1 - P_i(\hat{\theta}_s)$ is the probability of an incorrect response

Thus, a running estimate of an examinee's ability can be computed in simulation by computing the adjustment after each item result. Note that if the examinee answers either all or none of the items correctly then the estimation is either infinity or division by zero respectively.

The **precision** of this estimate is given by the calculation of the standard error:

$$SE(\hat{\theta}) = \frac{1}{\sqrt{\sum_{i=1}^{N} a_i^{\,2} P_i(\hat{\theta}_s) Q_i(\hat{\theta}_s)}}$$

The denominator is the square root of the estimate's denominator.

1.2 Applying Item Response Theory to ESTATE

Using Item Response Theory, we can think of the ESTATE conceptual formulation in another way. A trainee has an ability level at any given time, represented by θ. Since we can never know the true ability of the trainee, we can only estimate it. This estimation is assigned $\hat{\theta}_s$. Via simulation, we can bring the trainee ability against a challenge *c* and come out with a result *r*. We build up a repository of these interactions as a history of tuples $< c_i, \hat{\theta}_i, r_i >$. During diagnosis, we assess the current estimated ability level of the trainee based on the history of traces and determine $\hat{\theta}_s$. During adaptation, we attempt to find the optimal challenge c^* that will promote learning to serve the next round. c^* can be derived from finding the challenge such that the probability of getting that challenge correct given the currently estimated ability level of the trainee is greater than or equal to the probability of getting that challenge correct at the optimal ability level minus some delta. Formally, $P_{c*}(\hat{\theta}_i) + \Delta P >= P_{c*}(\theta^*)$. We can assume that $P_{c*}(\theta^*) = 0.5$, since at the target ability level, with the optimal challenge, the trainee has a 50% chance of responding to the challenge correctly. Furthermore, we can start with ΔP at 5% or 10% as an assumption of the zone of proximal development (ZPD). We can then adapt ΔP based on the current trend in answers being correct or incorrect in recent history. Based on this, $60\% >= P_{c*}(\hat{\theta}_i) >= 40\%$ with a $\Delta P = 10\%$.

Using our new formulation, we can adapt ESTATE diagram to represent this case. This diagram is shown in Figure 1.

1.2.1 Key Issues

1) **Bootstrapping**: Given the model above, $\hat{\theta}_s$ - the estimate of the trainee's ability - must be within a small error to derive a challenge problem that will fall in the ZPD and stimulate learning. ESTATE's estimated ability of the trainee must be close enough to the trainee's actual ability to be able to formulate a problem that is appropriately challenging. How many challenges must the trainee attempt before $\hat{\theta}_s$ falls within this error? This number must be small enough to reasonably require the trainees to attempt this many challenges before receiving learning gains from the system.

2) **Self-Sufficiency**: The input to the system should be as little as possible. Defining a curriculum of challenges, determining their difficulty, and ranking the abilities of training are extremely difficult and time consuming tasks for a training instructor and system developer. ESTATE should structure interactions to gather as much of this information as possible. Ideally, ESTATE should be given only a set of features used to create challenges and a scoring mechanism. The system should be able to assess trainees' ability and promote learning.

3) **Dynamics**: Traditional item response theory does not account for the possibility of learning as a result of attempting items. However, we expect the challenges in ESTATE to promote learning in the trainees. ESTATE must predict or assess learning gains to prevent its estimates of a trainee's ability from becoming inaccurate over time. ESTATE must balance choosing learning challenges with choosing assessment challenges.

1.3 Approach to Resolving Key Issues

1.3.1 Estimating number of Challenges

1.3.1.1 Standard Error Calculation

Given the model of item response theory in Section 1.1, how many items (of known challenge curves) must be attempted before the estimate of the learner's ability $\hat{\theta}_s$ is within a certain error, $\left|\hat{\theta}_s - \theta\right| < \varepsilon$? The standard error formula gives some indication of this number. For this example, we desire an error less than 5% (on a scale from [-3,3]). Thus letting $\varepsilon = 0.3$, we assume a ICC with difficulty = 0.0, discrimination = 1.0, and guessing = 0.0. From the standard error formula above:

$$0.3 = \frac{1}{\sqrt{(1^2 * 0.5 * 0.5) * i}}$$

$$i \approx 45$$

A trainee must attempt about 45 challenges before the system can (on average) select a challenge appropriate to stimulate learning. However, the 0.5 P and Q values are the best case scenario. The number of challenges needed may increase as the difficulty of the challenge strays from the trainee's ability. Also, this value quantifies the *average* error but does not provide a distribution or *maximum* error with which to estimate how effective such challenge selection will be across the population of trainees.

1.3.1.2 Estimating Trainee Ability using Simulation – Random Challenges

To explore these theoretical results further, we simulate the estimation of a trainee's ability based on attempting challenges of uniform random difficult across [-3, 3]. We vary the size of the history of the estimate (window size) basing the estimate on the most recent 20, 40, 60, etc. runs. To obtain reliable estimates, the history is filled and the estimates are simulated for the next 1000 challenges. This is one simulation run, and 10 runs are performed recording the average error percentage and the maximum error percentage.

The results are of this simulation are presented in Figure 2 and Figure 3. These results confirm the average error percent crossing the 0.05 mark at about 45 challenges, but the maximum error is still around 20% at this point! The standard error only accounts for the middle two thirds (66%) of the population, and *the estimates for the other third will fall outside of the acceptable range*. To achieve a standard error percent of 0.03 or less, approximately 130 challenges are needed, 0.02 is reached at approximately 250-300 challenges. To achieve a *maximum* error percent of 0.05 or less, 360-400 challenges are required.

The required number of challenges to achieve a precise ability estimate is too large for ESTATE's purposes. If each challenge takes a trainee 10 minutes (as is estimated), *this requires a time commitment from the trainee from 26 hours to about 66 hours* (depending on desired precision) to ensure an estimate with enough precision to enable learning. Also, *each time the trainee learns significantly (beyond the 5% error) from interacting with the system this calibration of the estimate must be repeated.* Clearly, an accurate and precise estimate must be obtained from less trainee interaction.

1.3.1.3 Estimating Trainee Ability using Simulation – Tailored Challenges

Perhaps the selection of challenge difficulty as uniformly distributed is increasing the estimate dramatically. If the challenge difficulty is varied to 'hone in' on the trainee's ability, the precision may be improved.

To test this hypothesis, we compare three methods of challenge difficulty selection in simulation. The first is as before, uniformly distributed, the second is randomly distributed about the estimate (a Gaussian distribution with mean as the theta estimate and standard deviation as 0.5), and the third is the current theta estimate itself.

The results are presented in Figure 4. The use of the theta estimate as the basis for choosing challenges during estimation seems to decrease the number of challenges

needed by 25% to 50%, with little discernable difference between using the theta estimate itself and introducing some randomness about the estimate. This is the same improvement achieved by the computerized GREs and other IRT test systems. However, since our challenges are expected to be much longer than an individual test question, by our approximation above this improved estimation *may still require 13 to 33 hours of trainee time before the required precision is achieved*. The individual challenges are not returning enough information from each challenge attempt, the single bit of correct/incorrect is insufficient to rapidly converge on a precise estimate.

1.3.1.4 Estimating Trainee Ability using Simulation – Continuous Response

ESTATE's challenges are intended to require much longer time commitment than the traditional item response theory item, which is often a single question on a test. Longer challenges incur longer assessment and calibration times. However, ESTATE may also be able to gather more information from each challenge. Instead of receiving a single correct/incorrect bit of information from the challenge, a normalized score may be assigned to each challenge performance. Under this scheme the score is a number in the range [0,1] instead of either a 0 or a 1. This increased information decreases the number of challenges required to compute a precise estimate.

To test the effectiveness of this approach, we recreate the simulation to incorporate continuous response scores. We assume a trainee will not always score the same on a challenge of a particular difficulty. Depending on a large number of random factors (e.g. trainee alertness, rest, or distraction) the trainee may display a particular level of ability that is normally distributed about his average ability; his performance will vary slightly from challenge to challenge. Because the item response curve is no longer a probability curve, but an indication of expected score, a value for the displayed ability is calculated before the item response function is applied, calculating a score for the challenge.

$$r_i = f_i(N(\theta, \sigma^2), a, b, c)$$

where r_i is the score returned for the *i*th challenge

$f_i(z,a,b,c)$ is the challenge curve for the *i*th challenge, with parameters *a, b, c*

$N(x,t)$ a normally distributed random variable with mean *x* and standard dev *t*

θ is the trainee's actual ability.

To calculate an estimation of the trainees ability from a score, we can derive the displayed ability for that score by using the inverse challenge curve function (from the three parameter model above):

$$\theta_i = f^{-1}(r_i) = b + \frac{\ln((1 - r_i)/(r_i - c))}{-a}$$

where θ_i is the displayed ability for the *i*th challenge

r_i is the score returned for the *i*th challenge

a, *b*, *c* are the item response curve parameters

The normal distribution of displayed ability allows the displayed ability to be calculated simply as the means of all of the samples, $\hat{\theta} = mean(\theta_i)$. The simulation results are presented in Figure 5, again varying size of history as windowSize.

These results show a sizeable improvement over earlier estimates with correct/incorrect information. Even accounting for a large standard deviation of displayed ability (1.0), the estimate converges on the actual ability of the trainee rapidly. The mean error crosses the 5% mark at about 7 challenges, down from 45, and the maximum error reaches 7% at about 45 challenges, down from 190. Also, the error scales linearly with the standard deviation of the displayed ability. If the standard deviation can be assumed to be 0.5, then the estimate will converge twice as fast. This method of scoring challenges on a continuous scale *reduces the trainee's time commitment before learning from days to minutes or hours.*

1.3.2 Estimating both the Challenge Curve and Trainee Ability

Since ESTATE may be generating the challenges that trainees attempt, we cannot assume that we will have a well-defined challenge curve for each challenge. ESTATE must estimate both the trainee's ability and the challenge curve simultaneously. Since the ICC depends on the estimate of ability and the estimate of ability depends on the performance from an ICC, ESTATE must make an assumption about either the abilities of the trainees or the shape of the challenge curve.

In the case where the challenge curve cannot be assumed, assumptions about the trainees' abilities may be made. Because the trainees' abilities are due to a large number of possible factors, the central limit theorem indicates that the abilities may be assumed to be normally distributed – such an assumption is often used initially for data concerning human performance. The shape of the challenge curve may now be estimated from the set of scores.

First the estimated points on the ability/score graph will be computed, then a spline curve will be used to interpolate the function representing these points. We make the additional assumption that the challenge curve is monotonically increasing: higher displayed ability will result in an equal or higher score. The scores are ordered by increasing value, and the abilities are calculated as if constructing a normal probability plot:

$$\hat{\theta} = f(x) = G(U(x))$$

where *U(x)* are the uniform order statistic medians

G(x) is the percent point function (inverse of the cumulative distribution function) of the normal distribution

A cubic spline may be interpolated from these points to create an estimate of the challenge curve. The details of this interpolation are beyond the scope of this document.

Figure 6 presents the results of one such estimation. 20 trainees with abilities sampled from a normal distribution, $N(\mu = 0, \sigma^2 = 1.0)$, each attempt a challenge, displaying ability with a small variance from their actual ability ($\sigma^2 = 0.1$). As is evident from the figure, the challenge curve is estimated with a high degree of accuracy (average error = 0.018%).

The estimate curve above will only match with the actual curve if the assumed mean and standard deviation of the trainees' ability is correct. If, for instance, a class of new trainees enters in to the system, we may still assume that their abilities are normally distributed, but they may yet be at the low end of the ability range. Figure 7 presents the results of performing the same challenge curve estimation when the abilities of the trainees are normally distributed about the first half of the ability scale, $N(\mu = -1.5, \sigma^2 = 0.5)$. Here, the lower half of the challenge curve is stretched across the entire range of abilities, resulting in a distorted view. It is important to note that the challenge curve is still a good estimate (i.e. low error interpolation) of the difficulty of the challenge for this set of trainees. Such challenge curves will still allow selection of challenges to fall within the ZPD as long as the abilities of the trainees fall within the initial range. The estimated ability of the trainees will be normalized across the entire range of ability.

1.3.3 Continuous Estimation and Learning

The challenge curve estimation above does not yet consider learning over time. How does learning affect the accuracy of the estimate? Can ESTATE promote continuous learning using the above approach? We measure the effectiveness of this approach in simulation.

Given a set of low ability trainees and a set of challenges with a full range of difficulty, can ESTATE reliably target trainees' ZPD and promote learning over time? Our simulation is initialized with a group of trainees with abilities averaging -2.5 on a [-3,3] scale (std dev is 0.15) and a set of 100 challenges (using the Rasch model) with difficulties spaced equally along the same range. A trainee's skill will improve by a small increment, 0.05, if his expected score is between 60% and 70%, the ZPD for this simulation. Given the parameters above, there is always at least one 'correct' challenge to present to a trainee. The simulation estimates the challenge curve based upon the history of scores. The next challenge for a trainee is chosen by finding the challenge with the expected score, based on the estimated curve, that falls within the range above.

Figure 8 presents the results of this simulation. After the initial estimation of the curve, the choice of challenge briefly matches the theoretical best choice. At about the 7th round of challenges, the estimate begins to depart sharply from the best choice. At about the 14th round of challenges, the estimate is no longer able to choose a challenge in the ZPD, and the learning of the trainees is halted.

These results occur because the trainee's abilities climb out of the range of the estimated challenge curve. The challenge curve is attempting to estimate a score for an ability for which it has not yet seen. In order to provide an accurate estimate, the curve needs to be calibrated not only once, but after learning occurs. Figure 9 presents the same simulation if recalibration is introduced after every 7th round. Here, the estimated result keep pace with the learning of the trainees, and the choices based on the estimate follow closely with the theoretically best choices.

As Figure 9 indicates, *ESTATE can use its estimation of trainee's abilities to promote continuous situational learning.* If the abilities of the trainees are normally distributed, *ESTATE can automatically discover the challenge appropriate for a particular trainee at a particular time.*

2. Scheduled Items

During the next reporting period, we plan to focus on the following tasks:

- Research automated challenge generation using challenge problem elements
  - Investigate machine learning techniques to discover how to estimate trainee response to generated challenges based on past challenges
  - Formulate feature vector representation for challenges
  - System experimentation using simulation to validate automated challenge generation approach
- Examine plausibility of running human subjects research with academic partner to gather empirical data based on framework
- Review exploring alternative trainee model representations (e.g., strategies, not abilities) for a stricter model-based approach as opposed to the data-driven approach of ability estimation and challenge generation

Sincerely,

Brad Rosenberg
Principal Investigator

3. Figures

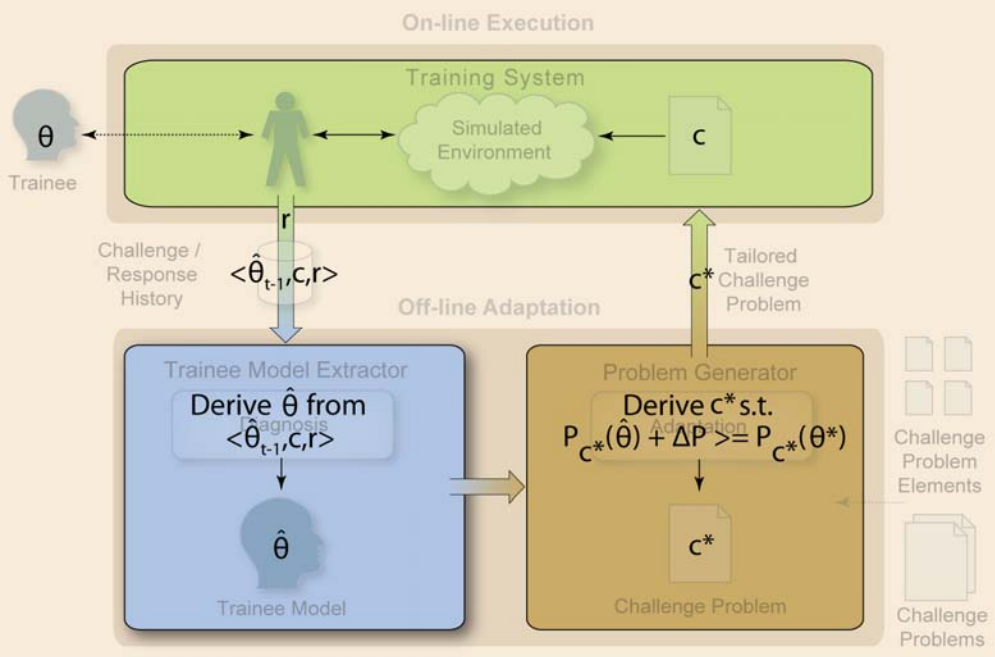


Figure 1: Apply Item Response Theory to ESTATE

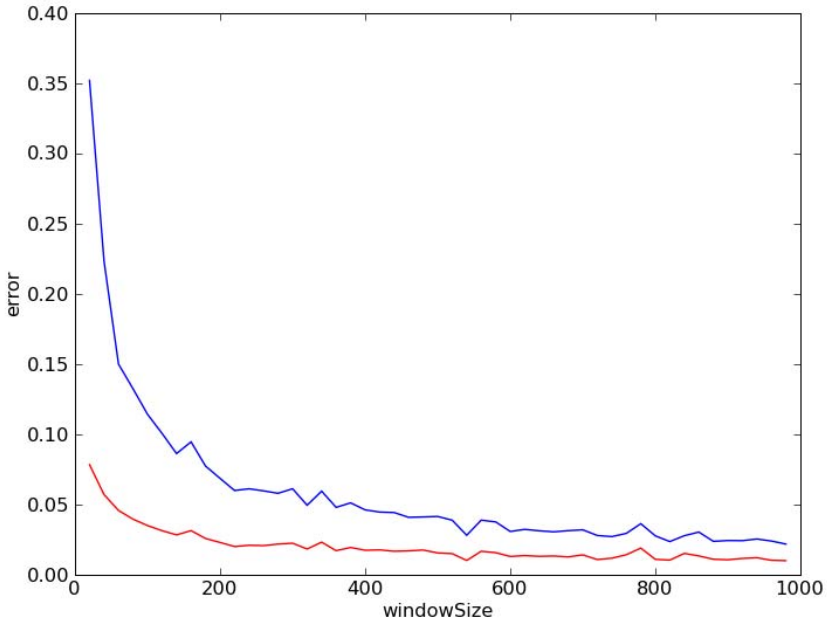


Figure 2: Average Error by window size

10 runs, 1000 challenges each, varying the window size of the estimate function: 20 – 1000 stepping by 20. Window was filled in simulation before recording data. Actual theta = -1, no learning. Average error in estimate % (red) and the Maximum error in estimate % (blue).

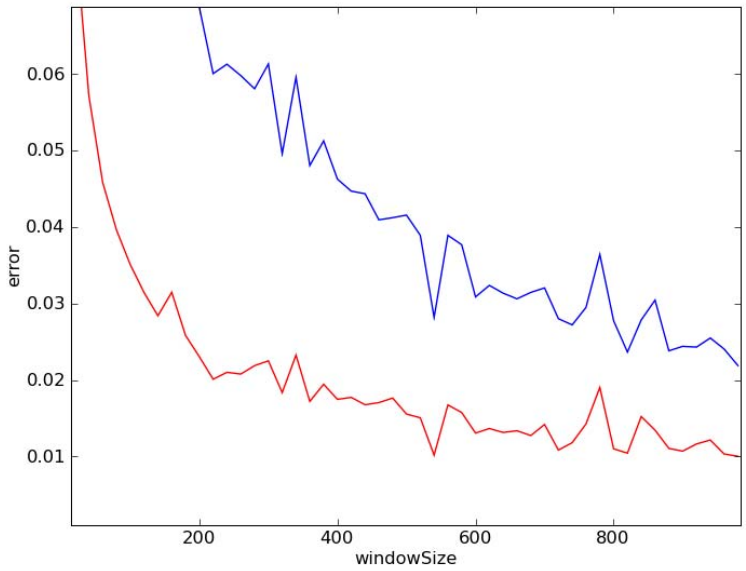


Figure 3: Average Error by window size, scaled

Scaled to lower 7%

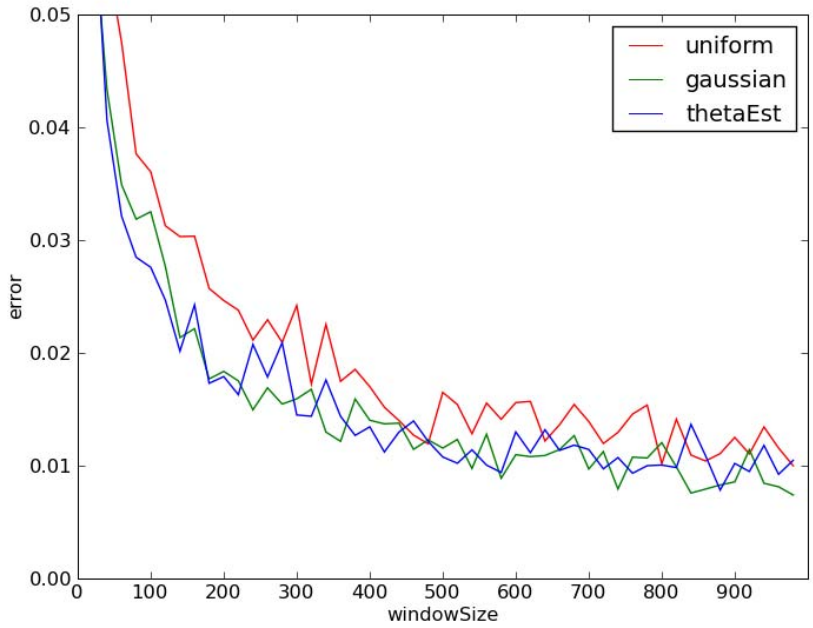


Figure 4: Average Error by window size, comparing item selection methods

```
uniform: random difficulty -3 to 3
gaussian: gaussian sample with mean as theta estimate and std dev as 0.5
thetaEst: the difficulty is set at the current theta estimate
```

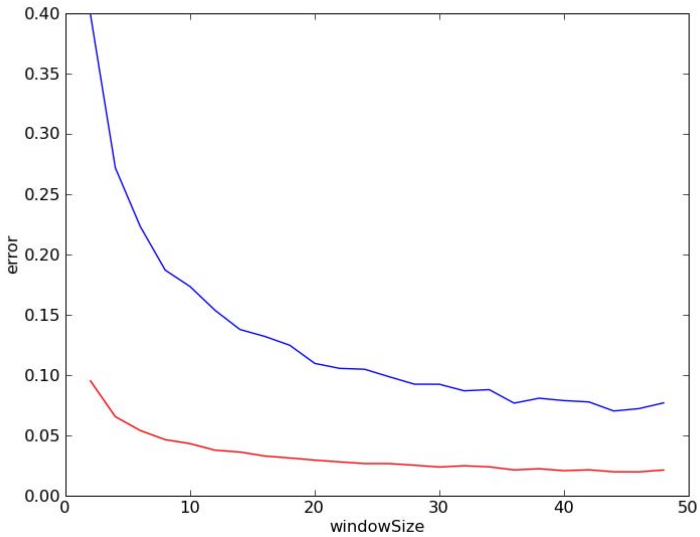


Figure 5: Average Error by window size, continuous response

% error by change in window size, continuous answer. Red is average, blue is maximum. Assumes Gaussian distribution of displayed ability, translated to score based on challenge response curve. Std dev set at 1.0 for all responses. Error scales linearly with standard deviation. (e.g. std dev 0.5 results in half the error).

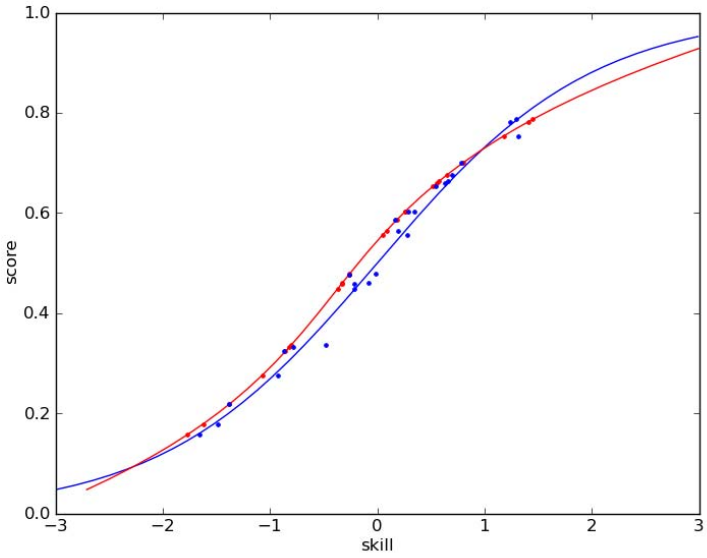


Figure 6: Curve estimation with full range of abilities

Blue line is actual challenge curve. Blue points are simulated scores with display ability normally distributed about actual ability (std dev = 0.1). Red line is estimated challenge curve. Red points are estimated skill for each score. 20 trainees, 1 attempt each. Error is 0.018%

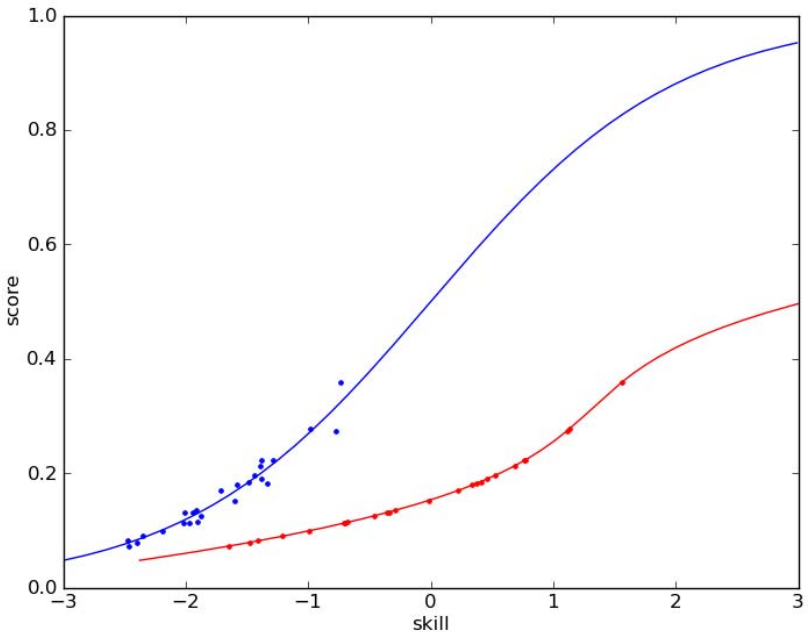


Figure 7: Curve estimation with half range of abilities

Blue line is actual challenge curve. Blue points are simulated scores with display ability normally distributed about actual ability (std dev = 0.1). Red line is estimated challenge curve. Red points are estimated skill for each score.

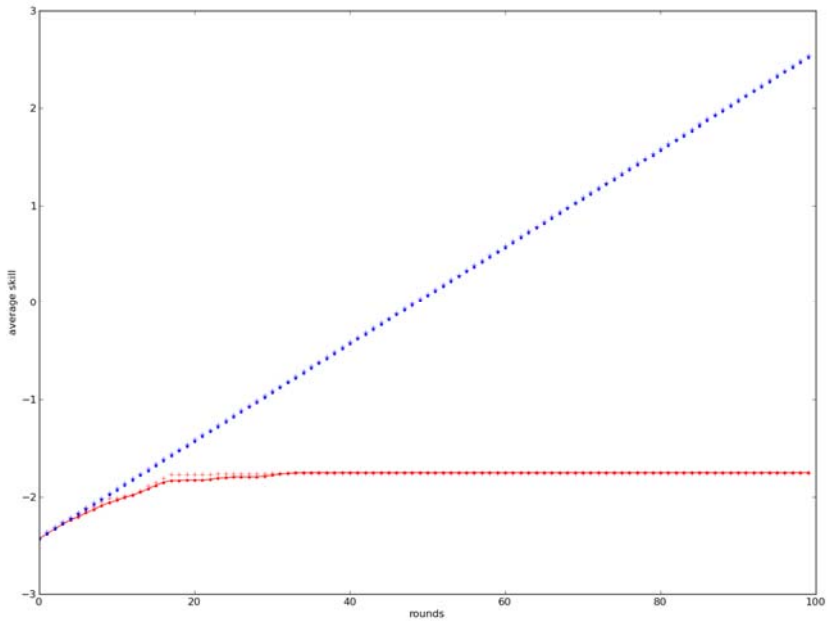


Figure 8: Learning induced without recalibration

Filled points are the mean ability, '+' points are the median ability. Blue points are theoretical best choice. Red points are chosen using estimated values.

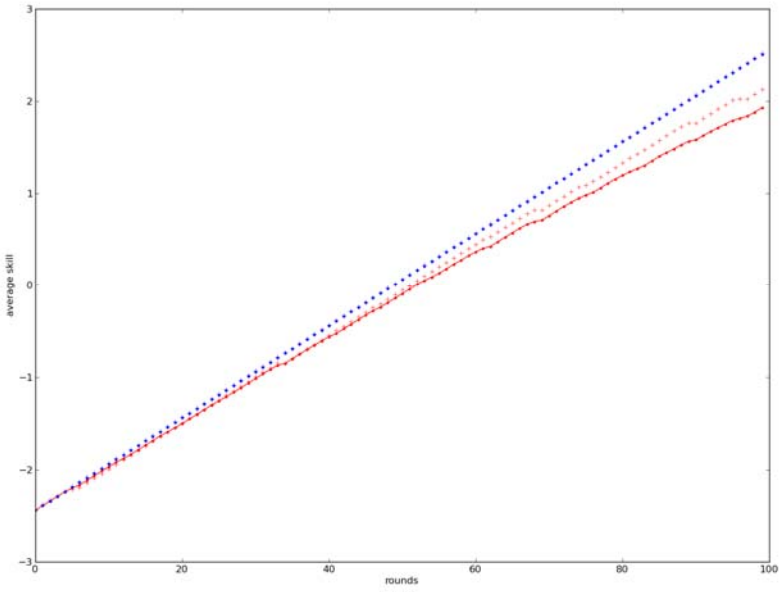


Figure 9: Learning induced with recalibration

Filled points are the mean ability, '+' points are the median ability. Blue points are theoretical best choice. Red points are chosen using estimated values.