DOCUMENT RESUME

ED 120 261 TM 005 224

AUTHOR Waller, Michael I.
TITLE Estimating Parameters in the Rasch Model: Removing the Effects of Random Guessing.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RB-76-8
PUB DATE Mar 76
NOTE 20p.; Paper originally presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., April 1975)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS *Ability; *Guessing (Tests); *Item Analysis; *Mathematical Models; Probability; *Response Style (Tests); Simulation
IDENTIFIERS Abilities Removing Random Guessing Model; *Rasch Model

ABSTRACT

A method of estimating the parameters of the Rasch Model removing the effect of random guessing is presented. The procedure is an application of the ARRG (Abilities Removing Random Guessing) model recently developed for two parameter latent trait models. Under the Rasch model ARRG provides for estimation of abilities, removing the effects of random guessing, without requiring the use of a computer. Monte Carlo simulations are employed to examine the accuracy of the resulting ability estimates. (Author/RC)

ED120261

RB-76-8

RESEARCH BULLETIN

ESTIMATING PARAMETERS IN THE RASCH MODEL: REMOVING THE EFFECTS OF RANDOM GUESSING

Michael I. Waller

This Bulletin was originally presented at the Annual Convention of the American Educational Research Association, Washington, D. C., April 1975. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
March 1976

Estimating Parameters in the Rasch Model:

Removing the Effects of Random Gueasing

Abstract

A method of estimating the parameters of the Rasch Model removing the effect of random guessing is presented. The procedure is an application of the ARRG model recently developed for two parameter latent trait models. Under the Rasch model ARRG provides for estimation of abilities, removing the effects of random guessing, without requiring the use of a computer. Monte Carlo simulations are employed to examine the accuracy of the resulting ability estimates.

1. Introduction

This paper presents a method of estimating the parameters of the Rasch model removing the effects of random guessing. The method used is an application of the ARRG (Abilities Removing Random Guessing) model or procedure developed recently for two parameter latent trait models by Waller (1973, 1974a). Application to the Rasch model is more or less straightforward, and the *raison d'etre* of the present work is to present a modified version which under the Rasch model becomes computer free. This is possible because one characteristic which the one parameter Rasch model possesses and which the two and three parameter logistic latent trait models do not is that the raw score is a sufficient statistic for estimation of ability. Consequently, once the item difficulties in a test have been estimated, a user may simply add up an examinee's raw score and look up the examinee's ability in a table developed during estimation of item difficulties. The method presented here provides the same facility while concomitantly removing the effects of random guessing.

The underlying assumption of the ARRG procedure is that examinees who guess in an essentially random manner, do so on those items which are too difficult for them. Latent trait models such as Rasch possess two characteristics which enable us to apply this assumption when estimating an examinee's ability. First, these models may be stated in a form which enables us to obtain an estimate of the probability of a correct answer for each examinee's response to each item. This allows us to identify items which are too difficult for a given examinee, and to identify examinees for which a given item is too difficult.

Second, once a set of items is calibrsted, i.e., scaled difficulties have been determined, any subset of items may be used to estimate an examinee's ability. This allows us to remove any number of items in a given test and indeed different sets for different examinees while still allowing us to estimate every examinee on the same scale (Bock and Wood, 1971).

Given this assumption and these two characteristics of latent trait models one may obtain estimates of difficulties and of abilities removing random guessing by removing from the estimation procedure those item-person interactions in which a person responds to an item estimated to be very difficult for him. Sections 2 to 4 describe the models, calibration of the items, and ability estimation by two methods. One method yields maximum likelihood estimates and requires the continual use of a computer; the other method yields a table from which estimates of abilities may be taken directly. The last section examines the accuracy of these ability estimates through a Monte Carlo simulation.

2. The Rasch and ARRG Models

Suppose that each of N subjects respond to M multiple choice items, each item containing A_i alternatives, $i=1,\ldots,M$. For the response of person n to item i, let:

$$a_{ni} = \begin{cases} 1 & \text{denote a correct response} \\ 0 & \text{denote an incorrect response.} \end{cases}$$

Then the probability that person n responds correctly to item i may be stated in Rasch model as:

$$(1) \qquad P_{ni} = Pr(a_{ni}=1) = \psi_{ni} = \frac{e^{b_n-d_i}}{1+e^{b_n-d_i}}$$

where b_n is the ability of person n and d_i is the difficulty of item i. This probability in the ARRG model becomes:

$$(2) \qquad P_{ni} = \begin{cases} \psi_{ni} & \psi_{ni} \geq P_c \\ 1/A_i & \psi_{ni} < P_c \end{cases} .$$

In (2) P_c is some small probability which is estimated during calibration of the test; the same value of P_c is used for all subjects and items.

Equation (2) reveals that the ARRG model divides the items into 2 sets:

$$S_n = [\text{Items such that } \psi_{ni} \geq P_c]$$

$$\overline{S}_n = [\text{Items such that } \psi_{ni} < P_c] .$$

Only items in S_n are used during ability estimation.

The ARRG procedure makes the assumption that P_c represents the probability below which some subjects answer randomly. While other subjects may choose to omit items rather than guess, the ARRG procedure allows us to ignore all responses with a low probability of a correct response and to do so without loss of precision in the resulting ability estimates (Waller, 1973 or 1974a, section 6). Omitted items in latent trait models are treated as incorrect responses for estimation of ability and difficulty. (Omitted items may be used as an adjunct to ARRG to obtain an independent estimate of the personality characteristic risk taking tendency as described in Waller (1974b). The estimate of risk taking tendency uses the items not used to estimate ability, i.e., items in $\overline{S}_n$.)

3. Test Calibration

Estimation of the item parameters, that is, calibration of the test, can be accomplished for the ARRG modification of the Rasch model by proceeding as in Waller (1974a). However, convergence during test calibration is found to be somewhat faster if an alternative procedure is employed. In the procedure used in this paper we adjust the responses prior to calibration to reflect the ARRG model and proceed with a free response analysis as in Wright and Panchapakesan (1969)[1].

The required adjustment is obtained as follows: A preliminary estimate of each subject's ability and each item's difficulty is obtained from a transformation of the raw proportion correct by the inverse of equation (1); i.e.,

$$b_n^o = \ln\left(\frac{P_n}{1-P_n}\right), \quad P_n = \text{raw proportion correct;}$$

(3)

$$d_i^o = \ln\left(\frac{P_i}{1-P_i}\right), \quad P_i = \text{raw difficulty.}$$

Then for each i and n, ψ_{ni} is estimated using (1). The response value to be used during ARRG calibration, say a^*_{ni}, is obtained from

(4) $$a^*_{ni} = \begin{cases} a_{ni} & \psi_{ni} \geq P_c \\ 0 & \psi_{ni} < P_c \end{cases} .$$

[1]The author wishes to thank Benjamin Wright for suggesting this method of calibration.

As described in Wright and Panchapakesan (1969) an estimate of ability is produced for each raw score as a part of the free response calibration process. In the ARRG procedure these estimates, based on the adjusted responses a^*_{ni}, are used only as starting points for the estimation of each subject's ability as described in the next section.

The proper value for P_c is obtained as follows; A cutoff probability, P_c, is chosen at the beginning of a calibration run and remains constant during the run. A χ^2 goodness of fit statistic (see e.g. Wright and Panchapakesan, 1969) is calculated for the test during each run. The value of P_c is then incremented and a second calibration is performed. After each calibration run, the fit is examined. The χ^2 statistic is found to decrease, as P_c is increased, to some minimum level and then increase (Waller, 1974a). The calibration run with the minimum χ^2 identifies the value to be used for P_c during estimation of ability, and the item difficulties obtained during this run become the estimated item difficulties of the test. Note that if the minimum χ^2 calibration occurs with P_c equal to zero this is an indication that the data are essentially free from random guessing, and in this case the resulting parameter estimates are identical to those produced by a standard Rasch calibration.

4. Ability Estimation

After the items are calibrated, estimation of ability under the ARRG procedure can be performed in either of two ways. The first method, here labelled MAX, produces maximum likelihood estimates (MLEs) and requires the use of a computer. The second method, here labeled TABLE, produces approximate MLEs but after calibration does not require a computer.

The first method yields maximum likelihood estimates by means of Newton Rapheson iteration. This method is identical to that used in free response Rasch estimation (Wright and Panchapakesan, 1969) with the exception that a subset of items reflecting the ARRG model is used. This is accomplished as follows:

Given the initial estimate of ability from (3), the set of item difficulties, and the value of P_c determined during calibration, the estimate of ability for person n at the k^{th} iteration is given in the Newton Rapheson approach by:

$$(5) \qquad b_n^{(k)} = b_n^{(k-1)} - \ell_b/\ell_{bb}.$$

Here ℓ_b and ℓ_{bb} are the first and second partial derivatives with respect to b of the log likelihood function, i.e.,

$$(6) \qquad b_n^{(k)} = b_n^{(k-1)} - \sum_{S_n} [a_{ni} - \psi_{ni}] / \sum_{S_n} [\psi_{ni}(1-\psi_{ni})] \quad .$$

In (6) ψ_{ni} is evaluated using the $(k-1)^{st}$ estimate of b_n; the ARRG model is reflected in the fact that the set S_n consists of only those items for which $\psi_{ni} \geq P_c$.

The second method, TABLE, consists of developing a table utilizing the estimated abilities, difficulties and P_c from the minimum χ^2 calibration run. For each raw score the table contains the estimated ability for that score together with a list of the items for which ψ_{ni} (as calculated using the ability associated with that score) is greater than the cutoff probability, P_c. An examiner obtains the estimated ability by adding up the subject's raw score, finding the subset of items in the table associated with that raw score, and then calculating an adjusted raw score based only on the items in this subset. The subject's estimated ability is the ability associated with this <u>adjusted</u> raw score.

For example, consider the two simulated response vectors, v_1 and v_2, taken from the Monte Carlo simulation described later. Items in these vectors range from easy items on the left to difficult items on the right. Both vectors yield a raw score of 14 correct out of 45.

$$v_1=\{\underline{1011000011001000000000000}10101010011010010000\}$$

$$v_2=\{\underline{1111111101111100000000100}00000000000000000000\} .$$

Table 1 presents the estimated Rasch abilities from the minimum

----
Insert Table 1 about here

χ^2 calibration run on this set of simulated data, and the subset of items for which the estimated probability of a correct response is greater than the cutoff point associated with this run, $P_c = 0.12$. We see that for a raw score of 14 this subset contains items one to twenty-five (underlined above); and consequently, the subset containing these items will be used to obtain these simulated examinees' estimated ARRG ability. The "examinee" corresponding to the first response vector achieved six correct responses in this subset and his estimated ARRG ability is the estimated Rasch ability corresponding to a raw score of six, $\hat{\theta}_1 = -3.577$. The second "examinee" achieved all fourteen correct responses in this subset and his estimated ability is the same as the Rasch ability corresponding to his adjusted raw score of fourteen, $\hat{\theta}_2 = -1.677$.

While the disparity between the estimated abilities might seem unwarranted (given that both "examinees" achieved the same raw score), the difference reflects the fact that the first response vector simulates someone of low ability who guesses in a more or less random manner at items which he doesn't know, while the second vector simulates someone of a higher ability but someone who omits items he doesn't know, omitted items being scored as incorrect. The estimated standard error of these two ability estimates are $\sigma_1 = 0.538$ and $\sigma_2 = 0.459$. Since the true ability, i.e., those used to generate these data, are $\theta_1 = -3.41$ and $\theta_2 = -1.58$, it's clear that 95% confidence intervals placed around the estimated abilities cover the true values.

During ability estimation with either method, the original responses of each examinee, a_{ni} not a^*_{ni}, are used. An examinee's response is never changed from right to wrong when estimating his ability; only the number of items used to estimate an examinee's ability may be changed. As indicated in the introduction, latent trait models allow estimation of abilities on the same scale using any subset of calibrated items. This characteristic may be seen in equation

(6) of the MAX method: responses to items not in the set used to estimate ability, S_n, do not affect estimation in any way.

5. A Monte Carlo Simulation

Evaluation of ARRG ability estimation may be made through a Monte Carlo simulation. A sample of abilities of size 480 from a logistic distribution was generated simulating both guessing and non-guessing subjects. The simulated test was constructed to have 45 items ranging in difficulty from -4.4 to +4.4 in steps of .2. A non-guessing response vector was generated by setting the value of his ability parameter, calculating the true probability of a correct response, ψ_{ni}, for each item and then comparing this value to a random number, r_{ni}, between zero and one: i.e.,

$$a_{ni} = \begin{cases} 1 & \text{if } \psi_{ni} \geq r_{ni} \\ 0 & \text{otherwise} \end{cases} .$$

A guessing subject's response vector was generated in the same manner with the exception that whenever the calculated probability, ψ_{ni}, was less than P_c, the random number was used to yield a simulated correct response on one-fifth of such items (thus simulating a five choice test). The sample contained approximately 33% guessers and the remainder non-guessers.

The simulated test thus generated was calibrated using the ARRG model and the free response model. As indicated earlier, if the best fitting calibration run occurs with $P_c = 0$, ARRG estimates are identical to free response estimates; that is, the free response analysis is subsumed by the ARRG analysis.

The comparison of the ARRG estimates to free response estimates is included to highlight the effect of ignoring random guessing when it is present in the form modeled by the ARRG procedure. When calibrating a real test, the χ^2 statistic will indicate whether anything will be gained by employing a cutoff point greater than zero.

The ability for each subject was estimated by the two methods presented for the ARRG model and by the free response model. The samples were then split into two subsamples, one composed of the simulated guessers, the other, simulated non-guessers. Table 2

Insert Table 2 about here

presents functions of the first four moments of the samples of the estimated abilities. These are compared to the corresponding moments of the sample of true abilities (also shown); all significant differences are indicated.

For each subject an estimate of the asymptotic variance of the subject's estimated ability may be calculated as

$$(7) \qquad \hat{\sigma}^2_{\hat{\theta}} = -1/\ \ell_{\hat{\theta}\hat{\theta}} = 1/\sum_{S_n} [\psi_{ni}(1-\psi_{ni})],$$

where the summation is again taken over only those items for which $\psi_{ni} \geq P_c$. Using (7) 95% confidence intervals may be placed around each subject's estimated ability. Table 3 presents the

Insert Table 3 about here

number of times, in each group of simulated subjects, such confidence intervals fail to cover the true ability.

Results from Tables 2 and 3 indicate that only the free response analysis of data contaminated by guessing fails to adequately estimate the true abilities. As would be expected from a free response analysis of data contaminated by guessing, the mean of the simulated guessers is significantly overestimated relative to the true mean. Also, this subsample is seen to contain a positive, though not satistically significant, skew relative to the true sample of abilities, as well as a significantly smaller variance. Both ARRG methods produce samples of estimated abilities which are statistically equivalent to the sample of true abilities.

While the TABLE method does not produce exact maximum likelihood estimates of ability (except for those subjects whose ability estimate is based on the entire set of items), the ability estimates produced by this method in conjunction with their estimated errors are found to produce an accurate estimate of an examinee's ability. The TABLE method also possesses the characteristic sought in this work that after calibration of an instrument by a test publisher, a user is not required to possess high speed electronic computing facilities to obtain accurate estimates of ability essentially uncontaminated by random guessing.

REFERENCES

Bock, R.D. and Wood, R. (1971) "Test Theory." Annual Review of Psychology, Vol. 22, pp. 193-224.

Waller, Michael I. (1973) "Removing the effects of random guessing from latent trait ability estimates." Unpublished Ph.D. Dissertation, University of Chicago.

Waller, Michael I. (1974a) "Removing the effects of random guessing from latent trait ability estimates." RB-74-32, Educational Testing Service, Princeton, New Jersey.

Waller, Michael I. (1974b) "Estimating Guessing Tendency." RB-74-33, Educational Testing Service, Princeton, New Jersey.

Wright, B. D. and Panchapakesan, N. (1969). "A procedure for sample free item analysis." Educational and Psychological Measurement. Vol. 27, No. 1.

TABLE 1

n	$\hat{b}_n$	Items for Which $\psi_{ni} \geq P_c = 0.12$	NO. OF ITEMS FOR WHICH $\psi \geq 0.12$
1	-5.95481	1 2 4 3	4
2	-5.14989	1 2 4 3 6 5 7	7
3	-4.62980	1 2 4 3 6 5 7 10 8 9	10
4	-4.22489	1 2 4 3 6 5 7 10 8 9 11 13	12
5	-3.88209	1 2 4 3 6 5 7 10 8 9 11 13 12	13
6	-3.57777	1 2 4 3 6 5 7 10 8 9 11 13 12 14	14
7	-3.29914	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16	16
8	-3.03932	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17	18
9	-2.79283	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19	19
10	-2.55460	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21	21
11	-2.32836	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22	22
12	-2.10639	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22	22
13	-1.88950	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22	22
14	-1.67672	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25	25
15	-1.46730	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27	27
16	-1.26063	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27	27
17	-1.05617	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28	28
18	-0.85354	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29	29
19	-0.65240	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30	30
20	-0.45243	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31	31
21	-0.25338	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33	32
22	-0.05507	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32	33
23	0.14269	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35	35
24	0.34006	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35	35
25	0.53721	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36	36
26	0.73430	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38	37
27	0.93154	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40	38
28	1.12913	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39	41
29	1.32735	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43	43
30	1.52657	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43	43
31	1.72723	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43	43
32	1.92992	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
33	2.13541	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
34	2.34473	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
35	2.55921	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
36	2.78066	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
37	3.01158	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
38	3.25542	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
39	3.51178	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
40	3.80696	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
41	4.13142	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
42	4.52071	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
43	5.02652	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45
44	5.81919	1 2 4 3 6 5 7 10 8 9 11 13 12 14 15 16 18 17 19 20 21 22 23 24 25 26 27 28 29 30 31 33 32 34 35 36 38 40 37 42 39 41 43 45 44	45

Table 2

Moments of the Recovered Samples of Abilities

Analysis	Mean	Variance	Skewness	Kurtosis	Fit χ^2 D.F.=135
Free Response	0.1211[a]	2.142[b]	0.1030	0.6794	267.81
Guessers	0.3095[c]	1.191[c]	0.6545	1.2963	
Non Guessers	0.0261	2.238[b]	-0.0849	0.3196	
ARRG-MAX	-0.0171	2.301	0.1721	0.4482	164.06
Guessers	-0.0247	3.032	0.4938	0.6772	
Non Guessers	-0.0133	2.693	0.0184	0.3242	
ARRG-TABLE	-0.0887	2.916	0.1758	0.4472	164.08
Guessers	-0.0716	3.123	0.4903	0.6386	
Non Guessers	-0.0974	2.824	0.0075	0.3369	
TRUE	-0.0770	2.769	0.1498	0.4674	
Guessers	-0.0407	2.957	0.3001	0.7382	
Non Guessers	-0.0604	2.692	0.2414	0.1369	

Sample Size

Full Sample	477
Guessers	160
Non Guessers	377

a $P < .009$
b $P << .001$
c $P < .001$

Table 3

Number of Estimated Confidence Intervals
Failing to Cover the True Ability

Analysis	Sample N	Number Failing To Cover Ability	P Less Than
Free Response	477	59	.0001
Guessers	160	37	.0001
Non Guessers	377	22	.0776
ARRG MAX	477	21	.9542
Guessers	160	5	.9161
Non Guessers	377	16	.5205
ARRG TABLE	477	24	.5540
Guessers	160	8	.5506
Non Guessers	377	16	.5205