DOCUMENT RESUME

ED 464 110 TM 033 791

AUTHOR	Raiche, Gilles; Blais, Jean-Guy
TITLE	Practical Considerations about Expected A Posteriori Estimation in Adaptive Testing: Adaptive A Priori, Adaptive Correction for Bias, and Adaptive Integration Interval.
PUB DATE	2002-04-07
NOTE	26p.; Paper presented at the Biennial Meeting of the International Objective Measurement Workshop (11th, New Orleans, LA, April 6-7, 2002).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Adaptive Testing; Computer Assisted Testing; Estimation (Mathematics); Item Response Theory; Simulation; *Statistical Bias
IDENTIFIERS	*Expected A Posteriori Tests; A Priori Tests; Rasch Model

ABSTRACT

In a computerized adaptive test (CAT), it would be desirable to obtain an acceptable precision of the proficiency level estimate using an optimal number of items. Decreasing the number of items is accompanied, however, by a certain degree of bias when the true proficiency level differs significantly from the a priori estimate. G. Raiche (2000) has suggested that it is possible to reduce the bias, and even the standard error of the estimate, by applying to each provisional estimation one of a combination of these strategies: (1) the adaptive correction for bias proposed by R. Bock and R. Mislevy (1982); (2) adaptive a priori estimate; and (3) adaptive integration interval. A simulation study was conducted to explore these approaches. One thousand administrations of a CAT were simulated for varying proficiency levels. Expected a priori estimation of a unidimensional Rasch model with 40 quadrature points was used for provisional and final estimates. Simulation results demonstrate that bias can be reduced more effectively by using an adaptive adjustment of the estimation procedure than by considering only the final correction for bias proposed by Bock and Mislevey. Findings suggest the use of these adaptive estimate strategies in adaptive testing, especially the adaptive a priori combined with the adaptive integration interval. (Contains 2 figures, 12 tables, and 13 references.) (SLD)

Practical considerations about expected a posteriori estimation in adaptive testing: Adaptive a priori, adaptive correction for bias, and adaptive integration interval

Paper presented at the
Biennal International Objective Measurement Workshop

New Orleans

Gilles Raîche, University of Moncton
Jean-Guy Blais, University of Montréal
April 7, 2002

Abstract

In a computerized adaptive test, we would like to obtain an acceptable precision of the proficiency level estimate using an optimal number of items. Unfortunately, decreasing the number of items is accompanied by a certain degree of bias when the true proficiency level differs significantly from the a priori estimate. Raîche suggests that it is possible to reduced the bias, and even the standard error of the estimate, by applying to each provisional estimation one or a combination of the following strategies: adaptive correction for bias proposed by Bock and Mislevy (1982), adaptive a priori estimate, and adaptive integration interval.

Practical considerations about expected a posteriori estimation in adaptive testing: Adaptive a priori, adaptive correction for bias, and adaptive integration interval

Paper presented at the
Biennal International Objective Measurement Workshop, New Orleans

Gilles Raîche, University of Moncton
Jean-Guy Blais, University of Montréal
April 7, 2002

In a computerized adaptive test (CAT), we would like to obtain an acceptable degree of accuracy of the proficiency level estimate, $\hat{\theta}$, while using an optimal number of items. With the Rasch model and with the expected a posteriori estimation method (EAP), the number of items needed to obtain a standard error between 0.40 and 0.20 usually varies between 13 and 40 (Raîche et Blais, 2002). Unfortunately, a decrease in the number of items is accompanied by the presence of bias when the true proficiency level differs significantly with the a priori estimate, $\hat{\theta}_{ap}$. Bock and Mislevy (1982, p. 439-442) proposed a formula to reduce the bias based on the principle of inversion of the reciprocal of the variance of the estimate. The new proficiency level estimate, $\hat{\theta}_{bm}$, is:

$$\hat{\theta}_{bm} = \frac{\hat{\theta}}{1 - s_{\hat{\theta}}^{2}}. \qquad \textbf{Equation 1.}$$

When administering a CAT, they suggested applying this formula only to the final estimate of the proficiency level. Raîche (2000, p. 191-194) found that the Bock and Mislevy's correction was more useful when the administered number of items is over 13 and the standard error of the estimate is less than 0.40. Is it possible to find a way to diminish even more the degree of bias of the proficiency estimate, specifically when we have to use a small number of items or allow a larger standard error?

According to Raîche (2000; Raîche et Blais, 2001), the bias could further be reduced by applying this method of correction to each provisional estimate of the proficiency level (i.e. the adaptive correction for bias, ACB). Raîche also suggests that the bias could be also reduced by: 1) replacing the a priori estimate of each computation of the provisional estimate of the proficiency (i.e. the adaptive a priori estimate, AAP); 2) adjusting and adapting the integration interval of the a priori estimate (i.e. the adaptive integration interval, IN). Combining these three adaptive strategies could also be effective. While reducing significantly the bias of the estimate we will also obtain a more accurate standard error of the estimate. A simulation study is therefore proposed to verify the afore mentioned suggestions.

Method

One thousand administrations of a CAT were simulated for proficiency levels, θ, varying between -9 and 0. EAP estimation of an unidimensional Rasch model with 40 quadrature points, q, was used for provisional and final estimates. The formal provisional and final EAP proficiency level estimates involve numerical integration and are given as

$$\hat{\theta}_{eap} = \frac{\sum_{k=1}^{q} X_k L_j(X_k) A(X_k)}{\sum_{k=1}^{q} L_j(X_k) A(X_k)}. \qquad \textbf{Equation 2.}$$

Where X_k is one of the q quadrature points, varying usually between -4 and 4, $L_j(X_k)$ is the likelihood of the quadrature point, here the one parameter logistic model, and $A(X_k)$ is a weight associated with that point, A being N($\hat{\theta}_{ap}$,1).

The standard error of the estimated proficiency level is approximated by

$$S_{\hat{\theta}_{eap}} = \frac{\sum_{k=1}^{q}(X_k - \hat{\theta}_{eap})^2 L_j(X_k)A(X_k)}{\sum_{k=1}^{q} L_j(X_k)A(X_k)}. \qquad \textbf{Equation 3.}$$

The initial a priori estimate, $\hat{\theta}_{ap}$, was equal to zero. The next item, j+1, was chosen with a difficulty parameter, b_{j+1}, corresponding to the provisional estimate, $\hat{\theta}_j$, supposing that an ideal situation prevailed, that is one where there is always the availability of finding such an item in the pool. The stopping rule according to the number of items administered was fixed at different values between 1 and 60.

Each of the 1000 simulations considered the eight following strategies for the provisional proficiency estimation: (1) standard EAP ($\hat{\theta}_{eap}$), (2) adaptive Bock and Mislevy correction for bias ($\hat{\theta}_{acb}$), (3) adaptive a priori estimate ($\hat{\theta}_{aap}$), (4) adaptive integration interval ($\hat{\theta}_{in}$), (5) adaptive Bock and Mislevy correction for bias mixed with adaptive a priori estimate ($\hat{\theta}_5$), (6) adaptive Bock and Mislevy correction for bias mixed with adaptive integration interval ($\hat{\theta}_{icb}$), (7) adaptive a priori estimate mixed with adaptive integration interval ($\hat{\theta}_{iap}$), and (8) standard final Bock and Mislevy correction for bias ($\hat{\theta}_{fcb}$). Empirical bias and empirical standard error at each levels of the proficiency estimates were computed and compared.

When the provisional proficiency level is computed using the adaptive Bock and Mislevy correction for bias, the estimator is

$$\hat{\theta}_{acb} = \frac{\hat{\theta}_{eap}}{1 - s^2_{\hat{\theta}_{eap}}} . \qquad \textbf{Equation 4.}$$

If the provisional proficiency level is computed using the adaptive a priori strategy, the a priori estimate of the weighting function $A(X_k)$ is replaced with the j-1 provisional proficiency level estimate $\hat{\theta}_{eap}$, so A is now N($\hat{\theta}_{eap}$,1).

The adaptive integration interval estimates $\hat{\theta}_{in}$ forces the q quadrature points to vary between $-4 + \hat{\theta}_{eap}$ and $4 + \hat{\theta}_{eap}$. The interval is now adaptive and not only fixed between -4 and 4. In mixed strategies the integration interval become $[-4 + \hat{\theta}_{aap}, 4 + \hat{\theta}_{aap}]$ for $\hat{\theta}_{iap}$ or $[-4 + \hat{\theta}_{acb}, 4 + \hat{\theta}_{acb}]$ for $\hat{\theta}_{icb}$.

Following the simulations, only the estimates $\hat{\theta}_{eap}$, $\hat{\theta}_{in}$, $\hat{\theta}_{iap}$ and $\hat{\theta}_{icb}$ are of interest and consequently only their results will be presented. Particularly not interesting, the combination of adaptive a priori and adaptive Bock and Mislevy correction for bias produce an important inflation of the bias of the proficiency level estimate.

Results

A substantial reduction in the empirical bias was found especially with ACB and AAP conditions combined with the IN condition ($\hat{\theta}_{icb}$ and $\hat{\theta}_{iap}$), compared to standard EAP and IN strategies ($\hat{\theta}_{eap}$ and $\hat{\theta}_{in}$). Figure 1 depicts the bias of the proficiency level estimate according to the number of items administered and to the proficiency level. Tables 1 to 6 give details for each value of the proficiency level. Figure 2 depicts the empirical standard error of the proficiency level estimate according to the number of items administered and to the proficiency level. Tables 7 to 12 give details for each value of the proficiency level. Only the empirical standard error of the proficiency level estimate is showed in these tables. The usual theoretical standard error of the proficiency level estimate computed while administering a CAT (equation 3) is noted only when necessary.

Extreme values of the proficiency level (θ = -9 and θ = -6)

With our seventh strategy (AAP and IN), $\hat{\theta}_{iap}$, and with as few as sixteen items as well as a value of the proficiency level so extreme as -9 (table 1), the bias could significantly be reduced with scores ranging from a high of 5.331 ($\hat{\theta}_{eap}$) to a low of -0.099 ($\hat{\theta}_{iap}$), while the empirical standard error of $\hat{\theta}_{iap}$ is equal to 0.605 (table 6). The theoretical value of the standard error is 0.511. To obtain empirical standard errors of the proficiency level estimate of 0.40, 0.30, and 0.20, we have to administer 22, 29 and 45 items. At a value so extreme of the proficiency level, the standard error of $\hat{\theta}_{eap}$ is unstable and doesn't make any sense. With the sixth strategy (ACB and IN), $\hat{\theta}_{icb}$, it took 45 items to reduce the bias to -0.100, while the empirical standard error was kept at 0,248. It is to be noted that the value of the theoretical standard error of $\hat{\theta}_{icb}$, which was computed to reflect the proficiency estimate, was 0.233. So, this theoretical value is very close to the empirical value of $\hat{\theta}_{icb}$. When the proficiency level is as low as -9, $\hat{\theta}_{iap}$ is clearly the best estimate: less bias and less empirical standard error.

With a proficiency level at a value so extreme as -6 and with as few as nine items (table 2), the bias of $\hat{\theta}_{iap}$ could significantly be reduced to 0,061 while the bias of the usual $\hat{\theta}_{eap}$ would equal 3.236. The empirical standard error of $\hat{\theta}_{iap}$ is equal to 0.678 (table 7). The theoretical value of the standard error is 0.675. To obtain empirical standard errors of the proficiency level estimate of 0.40, 0.30, and 0.20 we have to administer 17, 25 and 43 items. At a value so extreme of the proficiency level, the standard error of $\hat{\theta}_{eap}$ is also unstable and doesn't make any sense. With the sixth strategy (ACB and IN), $\hat{\theta}_{icb}$, it took 22 items to reduce the bias to -0.102, while the empirical standard error was kept at 0,369. The value of the theoretical standard error of $\hat{\theta}_{icb}$, which was computed to reflect the proficiency estimate, was

0.328. So, this theoretical value is very close to the empirical value of $\hat{\theta}_{icb}$. Again, when the proficiency level is as low as -6, $\hat{\theta}_{iap}$ it is clearly the best estimate: less bias and less empirical standard error.

To be noted, the IN strategy is not very much more effective that the EAP strategy in all cases.

Usual values of the proficiency level (θ = -3 to θ = 0)

When the proficiency level is equal to -3, -2, -1 and 0, the results are similar but quite not so extreme. For example, when the proficiency level is equal to -3, the biais of $\hat{\theta}_{iap}$ is so low as -0.008 (table 3), compared to 0.822 with $\hat{\theta}_{eap}$, using only height items administered while the empirical standard error equals 0.602 (table 8) (theoretical standard error equal to 0.510). The $\hat{\theta}_{icb}$ estimate is quite closed with a bias equal to -0.088 and an empirical standard error equal to 0.578 (theoretical standard error equal to 0.490).

The $\hat{\theta}_{iap}$ estimate always shows less bias at any value of the proficiency level and the number of items administered is considerably reduced when $\theta \leq -2$. The $\hat{\theta}_{icb}$ estimate is not less of interest than the $\hat{\theta}_{iap}$ estimate, and their bias values are about the same. Consequently $\hat{\theta}_{iap}$ and $\hat{\theta}_{icb}$ are always less biased than $\hat{\theta}_{eap}$ or $\hat{\theta}_{in}$. Even if the empirical standard error of $\hat{\theta}_{iap}$ and $\hat{\theta}_{icb}$ are bigger than the one of $\hat{\theta}_{eap}$ or $\hat{\theta}_{in}$, the reduction in bias is to be seriously considered.

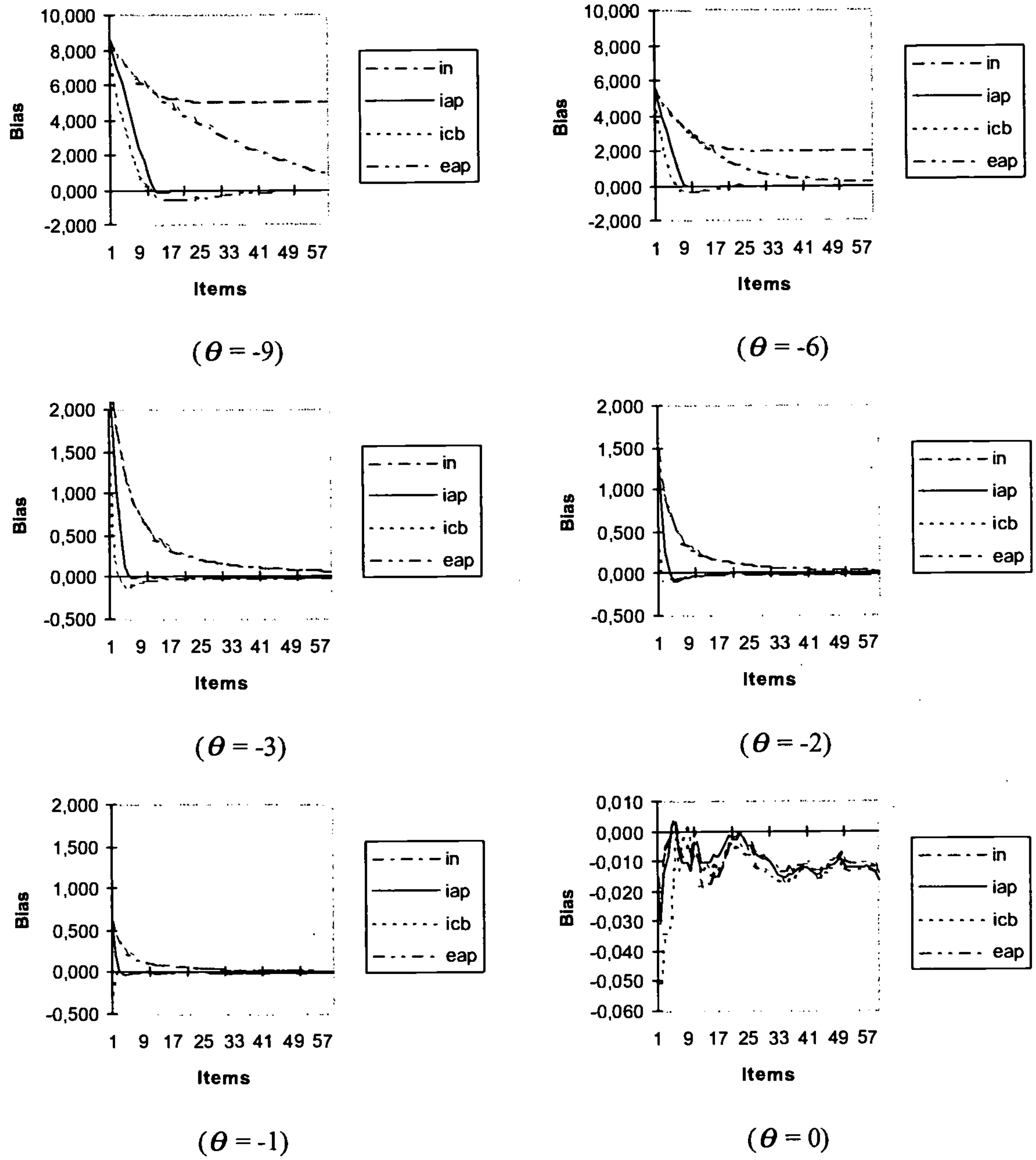

Figure 1. Bias of the proficiency estimate according to the number of items administered

Table 1.
Bias of the proficiency estimate according to the number of items administered (θ = -9)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	8.434	8.434	7.218	8.434	31	3.199	-0.014	-0.309	5.017
2	8.012	7.699	5.991	8.012	32	3.095	-0.016	-0.286	5.014
3	7.666	6.942	4.968	7.666	33	2.994	-0.019	-0.265	5.012
4	7.365	6.186	4.074	7.366	34	2.893	-0.020	-0.244	5.010
5	7.097	5.430	3.264	7.099	35	2.793	-0.020	-0.224	5.008
6	6.852	4.675	2.516	6.856	36	2.696	-0.019	-0.206	5.007
7	6.625	3.919	1.822	6.632	37	2.599	-0.018	-0.189	5.005
8	6.414	3.165	1.197	6.424	38	2.509	-0.016	-0.172	5.005
9	6.214	2.416	0.670	6.231	39	2.417	-0.016	-0.159	5.004
10	6.024	1.673	0.265	6.054	40	2.327	-0.015	-0.145	5.003
11	5.844	0.969	0.025	5.892	41	2.237	-0.015	-0.135	5.003
12	5.671	0.372	-0.193	5.745	42	2.151	-0.016	-0.126	5.002
13	5.505	-0.008	-0.349	5.615	43	2.067	-0.016	-0.117	5.002
14	5.345	-0.127	-0.438	5.503	44	1.983	-0.016	-0.109	5.001
15	5.191	-0.132	-0.509	5.409	45	1.902	-0.014	-0.100	5.001
16	5.042	-0.099	-0.539	5.331	46	1.825	-0.013	-0.093	5.001
17	4.897	-0.083	-0.571	5.267	47	1.751	-0.011	-0.085	5.001
18	4.757	-0.062	-0.583	5.216	48	1.678	-0.010	-0.079	5.001
19	4.620	-0.043	-0.582	5.175	49	1.605	-0.010	-0.073	5.001
20	4.486	-0.031	-0.574	5.142	50	1.535	-0.009	-0.068	5.000
21	4.356	-0.019	-0.563	5.116	51	1.468	-0.012	-0.065	5.000
22	4.229	-0.016	-0.543	5.094	52	1.401	-0.012	-0.061	5.000
23	4.105	-0.009	-0.519	5.077	53	1.339	-0.012	-0.058	5.000
24	3.984	-0.007	-0.496	5.064	54	1.279	-0.012	-0.055	5.000
25	3.865	-0.012	-0.469	5.052	55	1.221	-0.012	-0.052	5.000
26	3.749	-0.011	-0.442	5.043	56	1.166	-0.013	-0.050	5.000
27	3.636	-0.014	-0.415	5.036	57	1.114	-0.013	-0.046	5.000
28	3.523	-0.015	-0.386	5.029	58	1.064	-0.013	-0.045	5.000
29	3.413	-0.012	-0.359	5.024	59	1.018	-0.014	-0.043	5.000
30	3.306	-0.013	-0.334	5.020	60	0.972	-0.017	-0.045	5.000

Table 2.

Bias of the proficiency estimate according to the number of items administered (θ = -6)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	5.434	5.434	4.218	5.434	31	0.716	-0.008	-0.036	2.021
2	5.012	4.699	2.991	5.012	32	0.674	-0.011	-0.035	2.018
3	4.666	3.944	1.983	4.666	33	0.630	-0.012	-0.034	2.015
4	4.365	3.188	1.138	4.366	34	0.594	-0.012	-0.032	2.013
5	4.097	2.435	0.478	4.099	35	0.561	-0.013	-0.030	2.010
6	3.852	1.698	0.120	3.856	36	0.530	-0.013	-0.028	2.009
7	3.626	0.994	-0.099	3.633	37	0.501	-0.013	-0.026	2.007
8	3.417	0.424	-0.236	3.428	38	0.480	-0.010	-0.022	2.006
9	3.218	<u>0.061</u>	-0.298	3.236	39	0.454	-0.010	-0.020	2.005
10	3.030	-0.085	-0.322	3.060	40	0.433	-0.011	-0.019	2.005
11	2.851	-0.093	-0.320	2.898	41	0.412	-0.011	-0.018	2.004
12	2.681	-0.085	-0.317	2.754	42	0.394	-0.014	-0.019	2.003
13	2.520	-0.076	-0.309	2.626	43	0.375	-0.015	-0.019	2.003
14	2.365	-0.063	-0.288	2.517	44	0.359	-0.015	-0.018	2.002
15	2.216	-0.050	-0.262	2.424	45	0.346	-0.013	-0.016	2.002
16	2.075	-0.029	-0.233	2.346	46	0.334	-0.013	-0.015	2.002
17	1.937	-0.024	-0.208	2.281	47	0.323	-0.012	-0.013	2.002
18	1.807	-0.017	-0.182	2.230	48	0.314	-0.012	-0.012	2.001
19	1.684	-0.007	-0.155	2.188	49	0.302	-0.012	-0.012	2.001
20	1.569	-0.002	-0.135	2.155	50	0.293	-0.011	-0.010	2.001
21	1.459	-0.003	-0.117	2.127	51	0.280	-0.013	-0.012	2.001
22	1.352	-0.001	<u>-0.102</u>	2.104	52	0.270	-0.013	-0.012	2.001
23	1.261	0.005	-0.087	2.087	53	0.261	-0.013	-0.011	2.001
24	1.171	0.001	-0.077	2.072	54	0.254	-0.013	-0.011	2.001
25	1.086	-0.002	-0.069	2.060	55	0.246	-0.013	-0.010	2.001
26	1.013	-0.001	-0.059	2.050	56	0.237	-0.013	-0.011	2.001
27	0.941	-0.005	-0.055	2.042	57	0.231	-0.013	-0.011	2.001
28	0.873	-0.007	-0.050	2.035	58	0.225	-0.014	-0.012	2.001
29	0.819	-0.008	-0.045	2.030	59	0.218	-0.015	-0.012	2.000
30	0.763	-0.008	-0.040	2.025	60	0.209	-0.017	-0.015	2.000

Table 3.
Bias of the proficiency estimate according to the number of items administered (θ = -3)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	2.439	2.439	1.232	2.439	31	0.168	-0.005	-0.009	0.168
2	2.040	1.731	0.229	2.040	32	0.161	-0.006	-0.011	0.161
3	1.718	1.040	-0.013	1.718	33	0.152	-0.008	-0.013	0.152
4	1.457	0.463	-0.099	1.457	34	0.147	-0.009	-0.014	0.147
5	1.237	0.129	-0.129	1.237	35	0.142	-0.009	-0.013	0.142
6	1.066	0.019	-0.109	1.066	36	0.137	-0.008	-0.013	0.137
7	0.930	-0.011	-0.106	0.930	37	0.133	-0.008	-0.013	0.133
8	0.822	-0.008	-0.088	0.822	38	0.131	-0.006	-0.010	0.131
9	0.728	0.000	-0.069	0.728	39	0.127	-0.006	-0.010	0.127
10	0.641	-0.002	-0.057	0.641	40	0.123	-0.005	-0.010	0.123
11	0.578	0.008	-0.042	0.578	41	0.121	-0.005	-0.010	0.121
12	0.518	0.003	-0.039	0.518	42	0.116	-0.006	-0.012	0.116
13	0.465	-0.008	-0.039	0.465	43	0.110	-0.008	-0.013	0.110
14	0.430	-0.002	-0.031	0.430	44	0.108	-0.008	-0.013	0.108
15	0.398	-0.002	-0.028	0.398	45	0.106	-0.006	-0.011	0.106
16	0.372	0.002	-0.022	0.372	46	0.103	-0.006	-0.011	0.103
17	0.341	-0.001	-0.021	0.341	47	0.102	-0.004	-0.010	0.102
18	0.321	-0.001	-0.016	0.321	48	0.100	-0.004	-0.009	0.100
19	0.302	0.001	-0.010	0.302	49	0.097	-0.004	-0.009	0.097
20	0.286	0.003	-0.006	0.286	50	0.095	-0.004	-0.008	0.095
21	0.270	0.005	-0.005	0.270	51	0.092	-0.006	-0.010	0.092
22	0.257	0.004	-0.004	0.257	52	0.088	-0.007	-0.010	0.088
23	0.246	0.007	-0.002	0.246	53	0.086	-0.007	-0.010	0.086
24	0.233	0.005	-0.003	0.233	54	0.085	-0.007	-0.010	0.085
25	0.219	0.002	-0.005	0.219	55	0.083	-0.007	-0.010	0.083
26	0.211	0.001	-0.005	0.211	56	0.081	-0.007	-0.010	0.081
27	0.198	-0.003	-0.008	0.198	57	0.080	-0.007	-0.009	0.080
28	0.189	-0.004	-0.007	0.189	58	0.077	-0.008	-0.010	0.077
29	0.183	-0.005	-0.008	0.183	59	0.075	-0.009	-0.011	0.075
30	0.175	-0.005	-0.008	0.175	60	0.071	-0.011	-0.014	0.071

Table 4.
Bias of the proficiency estimate according to the number of items administered (θ = -2)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	1.454	1.454	0.282	1.454	31	0.083	-0.007	-0.008	0.084
2	1.105	0.806	-0.063	1.105	32	0.079	-0.008	-0.010	0.079
3	0.866	0.281	-0.071	0.866	33	0.073	-0.010	-0.011	0.074
4	0.693	0.007	-0.053	0.693	34	0.071	-0.011	-0.011	0.071
5	0.576	-0.084	-0.052	0.577	35	0.068	-0.011	-0.012	0.069
6	0.496	-0.085	-0.046	0.497	36	0.066	-0.010	-0.011	0.066
7	0.430	-0.075	-0.049	0.431	37	0.064	-0.009	-0.011	0.065
8	0.380	-0.057	-0.038	0.381	38	0.065	-0.007	-0.009	0.065
9	0.339	-0.047	-0.030	0.339	39	0.063	-0.007	-0.009	0.063
10	0.301	-0.041	-0.027	0.302	40	0.061	-0.007	-0.009	0.061
11	0.272	-0.027	-0.013	0.273	41	0.060	-0.007	-0.009	0.060
12	0.242	-0.027	-0.016	0.242	42	0.057	-0.008	-0.010	0.057
13	0.214	-0.027	-0.018	0.214	43	0.054	-0.009	-0.011	0.055
14	0.202	-0.022	-0.012	0.202	44	0.053	-0.009	-0.011	0.053
15	0.187	-0.019	-0.015	0.187	45	0.053	-0.008	-0.009	0.053
16	0.178	-0.015	-0.009	0.179	46	0.052	-0.008	-0.009	0.052
17	0.161	-0.016	-0.011	0.162	47	0.053	-0.006	-0.008	0.053
18	0.154	-0.011	-0.008	0.154	48	0.051	-0.006	-0.008	0.052
19	0.148	-0.004	-0.004	0.148	49	0.050	-0.006	-0.008	0.050
20	0.142	0.000	-0.001	0.143	50	0.049	-0.005	-0.007	0.050
21	0.137	0.001	0.000	0.138	51	0.046	-0.007	-0.009	0.047
22	0.130	-0.001	-0.001	0.130	52	0.045	-0.007	-0.009	0.045
23	0.125	0.001	0.001	0.126	53	0.044	-0.008	-0.010	0.044
24	0.119	-0.001	-0.002	0.119	54	0.043	-0.008	-0.009	0.043
25	0.111	-0.002	-0.005	0.111	55	0.042	-0.007	-0.009	0.043
26	0.107	-0.003	-0.005	0.107	56	0.041	-0.008	-0.009	0.042
27	0.099	-0.005	-0.008	0.100	57	0.040	-0.008	-0.009	0.041
28	0.094	-0.006	-0.009	0.095	58	0.039	-0.008	-0.010	0.039
29	0.091	-0.006	-0.008	0.092	59	0.037	-0.009	-0.011	0.037
30	0.087	-0.006	-0.009	0.087	60	0.034	-0.012	-0.013	0.034

Table 5.

Bias of the proficiency estimate according to the number of items administered (θ = -1)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	0.604	0.604	-0.248	0.604	31	0.038	-0.007	-0.001	0.038
2	0.421	0.208	0.030	0.421	32	0.035	-0.008	-0.003	0.036
3	0.346	0.023	0.020	0.346	33	0.033	-0.009	-0.006	0.033
4	0.279	-0.027	0.008	0.279	34	0.032	-0.010	-0.006	0.032
5	0.232	-0.031	0.002	0.232	35	0.030	-0.011	-0.006	0.031
6	0.207	-0.017	0.013	0.207	36	0.030	-0.010	-0.006	0.030
7	0.172	-0.016	0.007	0.172	37	0.029	-0.010	-0.006	0.029
8	0.158	-0.007	0.006	0.158	38	0.031	-0.008	-0.004	0.031
9	0.144	0.001	0.007	0.144	39	0.029	-0.008	-0.004	0.029
10	0.124	0.000	0.005	0.124	40	0.028	-0.008	-0.004	0.028
11	0.118	0.005	0.006	0.118	41	0.028	-0.008	-0.004	0.028
12	0.107	0.002	0.000	0.107	42	0.026	-0.009	-0.005	0.026
13	0.089	-0.008	-0.011	0.089	43	0.024	-0.010	-0.007	0.024
14	0.088	-0.005	-0.006	0.088	44	0.023	-0.009	-0.007	0.023
15	0.081	-0.005	-0.006	0.081	45	0.024	-0.008	-0.005	0.024
16	0.080	-0.001	0.001	0.080	46	0.023	-0.008	-0.005	0.024
17	0.072	-0.003	0.000	0.072	47	0.024	-0.007	-0.004	0.024
18	0.072	0.001	0.004	0.072	48	0.023	-0.006	-0.003	0.023
19	0.071	0.004	0.007	0.071	49	0.023	-0.006	-0.003	0.023
20	0.069	0.007	0.007	0.069	50	0.023	-0.005	-0.002	0.023
21	0.067	0.006	0.008	0.067	51	0.020	-0.006	-0.004	0.020
22	0.063	0.004	0.006	0.063	52	0.019	-0.007	-0.005	0.019
23	0.061	0.005	0.007	0.061	53	0.019	-0.007	-0.005	0.019
24	0.056	0.003	0.005	0.056	54	0.018	-0.007	-0.005	0.019
25	0.052	-0.001	0.004	0.052	55	0.018	-0.007	-0.005	0.018
26	0.051	-0.001	0.005	0.051	56	0.018	-0.007	-0.005	0.018
27	0.047	-0.004	0.001	0.047	57	0.017	-0.006	-0.004	0.017
28	0.043	-0.005	0.001	0.043	58	0.015	-0.007	-0.005	0.016
29	0.041	-0.006	0.000	0.041	59	0.014	-0.008	-0.006	0.014
30	0.039	-0.007	0.000	0.039	60	0.011	-0.010	-0.009	0.011

Table 6.
Bias of the proficiency estimate according to the number of items administered ($\theta = 0$)

ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$	ITEM	$bias_{in}$	$bias_{iap}$	$bias_{icb}$	$bias_{eap}$
1	-0.016	-0.016	-0.050	-0.016	31	-0.010	-0.010	-0.014	-0.010
2	-0.023	-0.031	-0.051	-0.023	32	-0.011	-0.012	-0.015	-0.011
3	-0.006	-0.014	-0.034	-0.006	33	-0.013	-0.014	-0.016	-0.013
4	-0.003	-0.008	-0.034	-0.003	34	-0.013	-0.015	-0.017	-0.013
5	0.004	0.002	-0.016	0.004	35	-0.013	-0.016	-0.018	-0.013
6	0.003	0.003	0.001	0.003	36	-0.012	-0.015	-0.016	-0.012
7	-0.006	-0.007	-0.014	-0.006	37	-0.013	-0.014	-0.016	-0.013
8	-0.008	-0.011	-0.003	-0.008	38	-0.011	-0.013	-0.014	-0.011
9	-0.005	-0.010	0.001	-0.005	39	-0.011	-0.012	-0.013	-0.011
10	-0.010	-0.013	-0.006	-0.010	40	-0.011	-0.012	-0.013	-0.011
11	-0.007	-0.004	0.000	-0.007	41	-0.011	-0.012	-0.013	-0.011
12	-0.010	-0.004	-0.006	-0.010	42	-0.012	-0.013	-0.014	-0.012
13	-0.019	-0.010	-0.014	-0.019	43	-0.013	-0.014	-0.016	-0.013
14	-0.017	-0.010	-0.011	-0.017	44	-0.013	-0.014	-0.015	-0.013
15	-0.017	-0.011	-0.014	-0.017	45	-0.012	-0.013	-0.014	-0.012
16	-0.014	-0.008	-0.011	-0.014	46	-0.011	-0.013	-0.014	-0.011
17	-0.015	-0.009	-0.013	-0.015	47	-0.010	-0.012	-0.012	-0.010
18	-0.013	-0.007	-0.012	-0.013	48	-0.009	-0.011	-0.012	-0.009
19	-0.008	-0.004	-0.006	-0.008	49	-0.009	-0.010	-0.012	-0.009
20	-0.004	-0.001	-0.004	-0.004	50	-0.008	-0.009	-0.011	-0.008
21	-0.003	-0.002	-0.005	-0.003	51	-0.010	-0.011	-0.012	-0.010
22	-0.003	-0.002	-0.006	-0.003	52	-0.010	-0.012	-0.013	-0.010
23	-0.001	-0.001	-0.004	-0.001	53	-0.011	-0.012	-0.013	-0.011
24	-0.004	-0.003	-0.007	-0.004	54	-0.011	-0.012	-0.013	-0.010
25	-0.006	-0.005	-0.008	-0.006	55	-0.010	-0.012	-0.013	-0.010
26	-0.004	-0.005	-0.008	-0.004	56	-0.010	-0.012	-0.013	-0.010
27	-0.006	-0.007	-0.011	-0.006	57	-0.010	-0.012	-0.013	-0.010
28	-0.007	-0.008	-0.012	-0.007	58	-0.011	-0.013	-0.013	-0.011
29	-0.008	-0.009	-0.013	-0.008	59	-0.012	-0.013	-0.013	-0.012
30	-0.009	-0.010	-0.013	-0.009	60	-0.014	-0.016	-0.016	-0.014

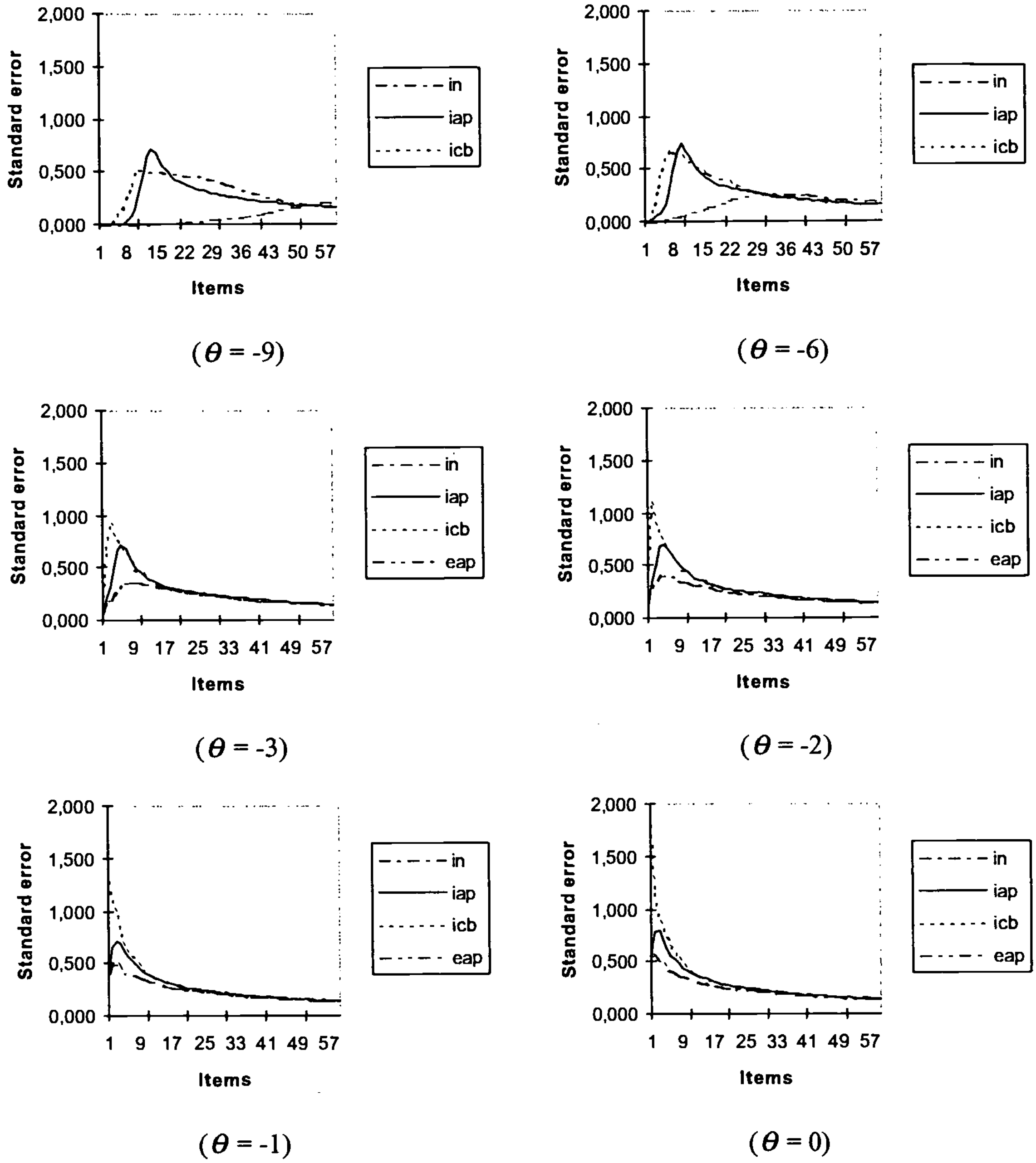

Figure 2. Standard error of the proficiency estimate according to the number of items administered

Table 7.
Empirical standard error of the proficiency estimate according to the number of items administered ($\theta = -9$)

ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}	ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}
1	0.000	0.000	0.000	0.000	31	0.047	0.284	0.403	0.001
2	0.000	0.000	0.000	0.000	32	0.051	0.271	0.389	0.000
3	0.000	0.000	0.000	0.000	33	0.056	0.268	0.379	0.000
4	0.000	0.000	0.000	0.000	34	0.060	0.262	0.367	0.000
5	0.000	0.000	0.048	0.000	35	0.063	0.254	0.354	0.000
6	0.000	0.000	0.094	0.000	36	0.067	0.249	0.342	0.000
7	0.000	0.000	0.148	0.000	37	0.071	0.240	0.328	0.000
8	0.000	0.036	0.252	0.000	38	0.078	0.235	0.319	0.000
9	0.000	0.086	0.380	0.000	39	0.084	0.228	0.307	0.000
10	0.000	0.154	0.496	0.000	40	0.089	0.221	0.295	0.000
11	0.000	0.277	0.518	0.000	41	0.093	0.219	0.285	0.000
12	0.000	0.469	0.515	0.000	42	0.101	0.219	0.275	0.000
13	0.000	0.655	0.528	0.000	43	0.108	0.211	0.264	0.000
14	0.000	0.710	0.511	0.000	44	0.114	0.208	0.255	0.000
15	0.000	0.677	0.513	0.007	45	0.121	0.204	0.248	0.000
16	0.013	0.605	0.501	0.007	46	0.129	0.198	0.239	0.000
17	0.014	0.554	0.500	0.006	47	0.139	0.197	0.232	0.000
18	0.015	0.502	0.496	0.005	48	0.147	0.193	0.224	0.000
19	0.016	0.461	0.490	0.004	49	0.155	0.187	0.216	0.000
20	0.017	0.427	0.488	0.004	50	0.161	0.186	0.212	0.000
21	0.018	0.410	0.480	0.003	51	0.168	0.184	0.206	0.000
22	0.019	0.390	0.470	0.003	52	0.174	0.181	0.201	0.000
23	0.023	0.370	0.462	0.002	53	0.183	0.177	0.196	0.000
24	0.026	0.356	0.460	0.002	54	0.190	0.176	0.190	0.000
25	0.028	0.342	0.455	0.001	55	0.199	0.172	0.185	0.000
26	0.031	0.330	0.451	0.001	56	0.204	0.172	0.183	0.000
27	0.037	0.318	0.445	0.001	57	0.209	0.171	0.180	0.000
28	0.039	0.308	0.436	0.001	58	0.214	0.168	0.176	0.000
29	0.041	0.299	0.426	0.001	59	0.218	0.166	0.172	0.000
30	0.044	0.294	0.416	0.001	60	0.219	0.166	0.170	0.000

Table 8.

Empirical standard error of the proficiency estimate according to the number of items administered ($\theta = -6$)

ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}	ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}
1	0.000	0.000	0.000	0.000	31	0.266	0.262	0.264	0.008
2	0.000	0.000	0.000	0.000	32	0.267	0.251	0.254	0.006
3	0.000	0.037	0.153	0.000	33	0.269	0.246	0.248	0.006
4	0.000	0.059	0.295	0.000	34	0.266	0.241	0.240	0.005
5	0.000	0.086	0.493	0.000	35	0.265	0.235	0.233	0.004
6	0.000	0.159	0.641	0.000	36	0.264	0.231	0.230	0.003
7	0.020	0.289	0.663	0.020	37	0.262	0.224	0.225	0.003
8	0.039	0.487	0.660	0.038	38	0.263	0.223	0.221	0.002
9	0.046	0.678	0.627	0.044	39	0.258	0.216	0.215	0.002
10	0.052	0.742	0.611	0.049	40	0.252	0.209	0.209	0.002
11	0.059	0.689	0.587	0.054	41	0.252	0.206	0.206	0.001
12	0.068	0.613	0.566	0.059	42	0.253	0.206	0.204	0.001
13	0.081	0.549	0.541	0.063	43	0.246	0.200	0.200	0.001
14	0.092	0.499	0.523	0.066	44	0.243	0.198	0.195	0.001
15	0.103	0.461	0.504	0.065	45	0.236	0.196	0.192	0.001
16	0.115	0.426	0.477	0.061	46	0.230	0.191	0.189	0.001
17	0.126	0.404	0.463	0.055	47	0.229	0.189	0.187	0.001
18	0.140	0.381	0.439	0.051	48	0.224	0.184	0.184	0.001
19	0.153	0.362	0.418	0.044	49	0.219	0.180	0.180	0.000
20	0.168	0.348	0.400	0.038	50	0.218	0.179	0.178	0.000
21	0.183	0.339	0.385	0.033	51	0.217	0.178	0.176	0.000
22	0.197	0.326	0.369	0.028	52	0.214	0.176	0.174	0.000
23	0.212	0.312	0.349	0.025	53	0.211	0.171	0.171	0.000
24	0.222	0.310	0.335	0.021	54	0.207	0.169	0.170	0.000
25	0.232	0.299	0.320	0.018	55	0.203	0.165	0.167	0.000
26	0.241	0.293	0.308	0.015	56	0.203	0.164	0.166	0.000
27	0.251	0.287	0.297	0.013	57	0.199	0.162	0.164	0.000
28	0.263	0.281	0.288	0.012	58	0.195	0.161	0.162	0.000
29	0.269	0.274	0.280	0.010	59	0.194	0.159	0.160	0.000
30	0.269	0.267	0.272	0.009	60	0.195	0.160	0.160	0.000

Table 9.
Empirical standard error of the proficiency estimate according to the number of items administered ($\theta = -3$)

ITEM	s_{in}	s_{iap}	s_{icb}	s_{eap}	ITEM	s_{in}	s_{iap}	s_{icb}	s_{eap}
1	0.072	0.072	0.225	0.072	31	0.231	0.230	0.235	0.231
2	0.156	0.181	0.689	0.156	32	0.223	0.222	0.227	0.223
3	0.200	0.322	0.925	0.200	33	0.222	0.220	0.225	0.222
4	0.248	0.508	0.803	0.248	34	0.218	0.219	0.220	0.218
5	0.289	0.663	0.742	0.289	35	0.215	0.214	0.215	0.215
6	0.325	0.709	0.685	0.325	36	0.210	0.212	0.211	0.210
7	0.348	0.678	0.624	0.348	37	0.203	0.208	0.206	0.203
8	0.357	0.602	0.578	0.357	38	0.200	0.206	0.204	0.200
9	0.358	0.539	0.530	0.358	39	0.193	0.201	0.199	0.193
10	0.362	0.490	0.485	0.362	40	0.188	0.196	0.193	0.188
11	0.356	0.449	0.449	0.356	41	0.186	0.194	0.192	0.186
12	0.347	0.415	0.420	0.347	42	0.185	0.192	0.190	0.185
13	0.342	0.392	0.394	0.342	43	0.183	0.188	0.185	0.183
14	0.329	0.369	0.373	0.329	44	0.181	0.185	0.181	0.181
15	0.321	0.354	0.356	0.321	45	0.180	0.184	0.179	0.180
16	0.312	0.336	0.338	0.312	46	0.176	0.180	0.176	0.176
17	0.316	0.332	0.331	0.316	47	0.176	0.178	0.174	0.176
18	0.302	0.314	0.315	0.302	48	0.174	0.174	0.171	0.174
19	0.293	0.305	0.303	0.293	49	0.171	0.169	0.169	0.171
20	0.289	0.296	0.295	0.289	50	0.170	0.169	0.167	0.170
21	0.283	0.290	0.286	0.283	51	0.169	0.167	0.165	0.169
22	0.275	0.282	0.276	0.275	52	0.165	0.165	0.163	0.165
23	0.269	0.271	0.268	0.269	53	0.162	0.161	0.159	0.162
24	0.266	0.270	0.267	0.266	54	0.162	0.159	0.158	0.162
25	0.258	0.261	0.265	0.258	55	0.160	0.156	0.155	0.160
26	0.251	0.254	0.260	0.251	56	0.158	0.156	0.154	0.158
27	0.247	0.250	0.255	0.247	57	0.155	0.154	0.153	0.155
28	0.244	0.246	0.251	0.244	58	0.153	0.153	0.151	0.153
29	0.239	0.241	0.244	0.239	59	0.152	0.152	0.150	0.152
30	0.237	0.237	0.241	0.237	60	0.152	0.151	0.149	0.152

Table 10.

Empirical standard error of the proficiency estimate according to the number of items administered (θ = -2)

ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}	ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}
1	0.151	0.151	0.474	0.151	31	0.212	0.230	0.229	0.212
2	0.275	0.330	1.107	0.275	32	0.206	0.221	0.221	0.206
3	0.356	0.531	0.863	0.356	33	0.206	0.219	0.219	0.206
4	0.394	0.678	0.766	0.393	34	0.204	0.213	0.214	0.204
5	0.404	0.698	0.709	0.404	35	0.199	0.208	0.209	0.199
6	0.414	0.664	0.641	0.413	36	0.197	0.205	0.206	0.197
7	0.401	0.599	0.577	0.400	37	0.193	0.199	0.200	0.193
8	0.385	0.548	0.527	0.384	38	0.192	0.199	0.199	0.192
9	0.364	0.497	0.481	0.363	39	0.186	0.193	0.193	0.186
10	0.349	0.465	0.452	0.348	40	0.181	0.188	0.188	0.181
11	0.340	0.433	0.424	0.340	41	0.181	0.185	0.185	0.180
12	0.323	0.407	0.398	0.323	42	0.181	0.184	0.185	0.180
13	0.314	0.381	0.384	0.314	43	0.178	0.181	0.182	0.178
14	0.305	0.365	0.371	0.305	44	0.176	0.179	0.180	0.175
15	0.301	0.353	0.351	0.301	45	0.174	0.178	0.179	0.174
16	0.290	0.332	0.330	0.289	46	0.172	0.175	0.176	0.172
17	0.287	0.324	0.329	0.286	47	0.171	0.173	0.174	0.170
18	0.272	0.307	0.310	0.271	48	0.168	0.169	0.170	0.167
19	0.264	0.297	0.301	0.264	49	0.165	0.165	0.166	0.165
20	0.256	0.287	0.291	0.256	50	0.164	0.164	0.166	0.164
21	0.250	0.283	0.285	0.250	51	0.162	0.162	0.164	0.162
22	0.244	0.275	0.274	0.244	52	0.160	0.160	0.161	0.160
23	0.239	0.264	0.262	0.238	53	0.158	0.156	0.157	0.158
24	0.237	0.261	0.263	0.237	54	0.157	0.155	0.156	0.157
25	0.231	0.254	0.257	0.231	55	0.155	0.153	0.154	0.155
26	0.229	0.251	0.248	0.229	56	0.154	0.153	0.153	0.154
27	0.226	0.249	0.246	0.225	57	0.152	0.151	0.151	0.152
28	0.222	0.246	0.243	0.221	58	0.149	0.149	0.149	0.149
29	0.217	0.240	0.238	0.217	59	0.148	0.148	0.148	0.148
30	0.216	0.236	0.236	0.216	60	0.148	0.148	0.149	0.148

Table 11.
Empirical standard error of the proficiency estimate according to the number of items administered (θ = -1)

ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}	ITEM	S_{in}	S_{iap}	S_{icb}	S_{eap}
1	0.405	0.405	1.274	0.405	31	0.210	0.214	0.222	0.210
2	0.489	0.652	1.095	0.488	32	0.204	0.206	0.215	0.204
3	0.484	0.715	0.946	0.483	33	0.202	0.205	0.213	0.202
4	0.443	0.688	0.760	0.443	34	0.199	0.201	0.211	0.199
5	0.414	0.618	0.651	0.414	35	0.194	0.198	0.206	0.194
6	0.409	0.567	0.601	0.409	36	0.192	0.197	0.203	0.192
7	0.394	0.527	0.547	0.394	37	0.187	0.192	0.198	0.187
8	0.385	0.490	0.512	0.385	38	0.187	0.192	0.198	0.187
9	0.369	0.453	0.471	0.369	39	0.183	0.187	0.193	0.183
10	0.361	0.420	0.440	0.361	40	0.178	0.183	0.189	0.178
11	0.344	0.394	0.414	0.344	41	0.178	0.181	0.187	0.178
12	0.328	0.378	0.392	0.328	42	0.178	0.180	0.188	0.178
13	0.319	0.360	0.369	0.319	43	0.174	0.176	0.184	0.174
14	0.308	0.344	0.351	0.307	44	0.171	0.174	0.181	0.171
15	0.296	0.328	0.341	0.296	45	0.170	0.173	0.180	0.170
16	0.283	0.312	0.320	0.282	46	0.167	0.171	0.177	0.167
17	0.280	0.309	0.312	0.280	47	0.165	0.169	0.174	0.165
18	0.268	0.294	0.298	0.268	48	0.162	0.166	0.172	0.162
19	0.259	0.285	0.289	0.259	49	0.159	0.164	0.169	0.159
20	0.253	0.273	0.281	0.253	50	0.158	0.162	0.167	0.158
21	0.248	0.267	0.273	0.248	51	0.157	0.161	0.165	0.157
22	0.243	0.261	0.266	0.243	52	0.155	0.158	0.162	0.155
23	0.236	0.252	0.256	0.236	53	0.152	0.155	0.159	0.152
24	0.237	0.249	0.255	0.237	54	0.151	0.153	0.158	0.151
25	0.232	0.243	0.250	0.232	55	0.149	0.152	0.156	0.149
26	0.228	0.238	0.246	0.228	56	0.148	0.151	0.155	0.148
27	0.227	0.233	0.242	0.226	57	0.147	0.149	0.153	0.147
28	0.223	0.228	0.236	0.223	58	0.145	0.147	0.151	0.145
29	0.217	0.222	0.230	0.217	59	0.144	0.146	0.150	0.144
30	0.215	0.218	0.227	0.215	60	0.144	0.146	0.150	0.143

Table 12.
Empirical standard error of the proficiency estimate according to the number of items administered ($\theta = 0$)

ITEM	s_{in}	s_{iap}	s_{icb}	s_{eap}	ITEM	s_{in}	s_{iap}	s_{icb}	s_{eap}
1	0.566	0.566	1.783	0.566	31	0.213	0.225	0.231	0.213
2	0.550	0.788	1.131	0.550	32	0.205	0.217	0.222	0.205
3	0.511	0.796	0.883	0.511	33	0.202	0.213	0.220	0.202
4	0.469	0.711	0.791	0.469	34	0.198	0.209	0.216	0.198
5	0.435	0.619	0.717	0.435	35	0.192	0.203	0.209	0.192
6	0.420	0.557	0.631	0.420	36	0.190	0.200	0.207	0.190
7	0.393	0.513	0.564	0.392	37	0.186	0.195	0.203	0.186
8	0.380	0.481	0.517	0.380	38	0.186	0.194	0.202	0.186
9	0.361	0.445	0.473	0.361	39	0.181	0.190	0.196	0.181
10	0.351	0.418	0.436	0.351	40	0.178	0.185	0.191	0.178
11	0.333	0.393	0.409	0.333	41	0.177	0.184	0.189	0.177
12	0.324	0.374	0.395	0.324	42	0.175	0.183	0.187	0.175
13	0.312	0.356	0.376	0.312	43	0.171	0.179	0.184	0.171
14	0.298	0.343	0.359	0.298	44	0.168	0.176	0.181	0.168
15	0.293	0.336	0.346	0.293	45	0.166	0.175	0.179	0.166
16	0.279	0.319	0.327	0.279	46	0.164	0.172	0.176	0.164
17	0.277	0.312	0.322	0.277	47	0.162	0.170	0.174	0.162
18	0.263	0.293	0.305	0.263	48	0.160	0.167	0.171	0.160
19	0.257	0.282	0.294	0.257	49	0.157	0.164	0.168	0.157
20	0.252	0.273	0.285	0.252	50	0.156	0.162	0.168	0.156
21	0.247	0.270	0.278	0.247	51	0.154	0.159	0.165	0.154
22	0.242	0.263	0.270	0.242	52	0.152	0.157	0.162	0.152
23	0.236	0.254	0.261	0.236	53	0.149	0.154	0.160	0.149
24	0.236	0.252	0.258	0.236	54	0.149	0.153	0.159	0.149
25	0.231	0.248	0.253	0.231	55	0.148	0.151	0.157	0.148
26	0.229	0.244	0.250	0.229	56	0.146	0.150	0.156	0.146
27	0.226	0.241	0.247	0.226	57	0.145	0.149	0.155	0.145
28	0.223	0.237	0.243	0.223	58	0.144	0.147	0.153	0.144
29	0.219	0.231	0.239	0.219	59	0.142	0.146	0.152	0.142
30	0.218	0.229	0.236	0.217	60	0.142	0.145	0.153	0.142

Conclusion

The results of our study demonstrate that bias can be reduced more effectively by using an adaptive adjustment of the estimation procedure rather than considering only the final correction for bias proposed by Bock and Mislevy. These findings suggest the use of these adaptive estimation strategies in adaptive testing, especially the adaptive a priori combined with the adaptive integration interval.

The results indicate a substantial reduction in the empirical bias, compared to the standard EAP strategy, especially with ACB and AAP conditions respectively combined with the the IN condition, and an increase of the standard error. It would appear that at extreme values of the proficiency level, the ICB condition allowed the largest increase of the empirical standard error, while the IAP condition diminished more significantly the level of bias while at the same time less increase the empirical standard error. Considering these results, the AAP condition combined with the IN condition seems to be a good strategy to incorporate to the CAT for bias reduction of the proficiency estimate.

Additional research is needed to determine the influence of the use of more or less q quadrature points and of variation of the range of integration [*min*, *max*] according to the number of items administered. WLE estimation is also attractive because of is unbiasedness to order n-1 (van der Linden and Glas, 2000, p. 207 ; van der Linden and Pashley, 2000, p. 5 ; Warm, 1989). Research on adaptive WLE estimators combined with IAP or ICB strategies would also be of interest.

We have also noticed biaseness of the theoretical standard error of the proficiecy estimate: positive for $\hat{\theta}_{iap}$ and $\hat{\theta}_{icb}$, and negative for $\hat{\theta}_{eap}$ and $\hat{\theta}_{in}$. Can we apply a correction for the bias of the theoretical standard error similar to the one used by Bock and Mislevy?

Finally, links should be made with sequential analysis (Wald, 1947), EM estimation (Demster, Laird and Rubin, 1983) or Stein's estimator and its competitors and generalizations, like the empirical bayesian estimation (Braun, 1989 ; Efron and Morris, 1973, 1975), being that they seem to be established on similar bases.

References

Bock, R.D., and Mislevy, R.J. (1982). Adaptive EAP estimation of ability in a micro computer environment. *Applied Psychological Measurement, 6*(4), 431-444.

Braun, H.I. (1989). Empirical bayes methods: A tools for exploratory analysis. *In* R.D. Bock (Ed.): *Multilevel analysis of educational data.* New York: Academic press.

Demster, A.P., Laird, N.M. and Rubin, D.B. (1983). Maximum likelihood estimation from incomplete data via the EM algorithm. *The Journal of the Royal Statistical Society, Series B, 39*(1), 1-38.

Efron, B. and Morris, C.N. (1973). Stein's estimation and its competitors - An empirical Bayes approach. *Journal of the American Statistical Association, 68*, 117-130.

Efron, B. and Morris, C.N. (1975). Data analysis using Stein's estimatior and its generalizations. *Journal of the American Statistical Association, 70*, 311-319.

Raîche, G. (2000). *La distribution d'échantillonnage en testing adaptatif en fonction de deux règles d'arrêt : selon l'erreur type et selon le nombre d'items administrés* [*Sampling distribution of the proficiency estimate in computerized adaptive testing according to two stopping rules: the standard error of the estimate and the number of administered items*]. Doctoral thesis. Montréal: University of Montréal.

Raîche, G. (2002, submitted). La simulation d'un test adaptatif basé sur le modèle de Rasch [*Simulation of a Rasch-based adaptive test*]. *Mesure et évaluation en éducation*.

Raîche, G. (2002, accepted). Le testing adaptatif [Adaptive testing]. *In* R. Bertrand and J.G. Blais (Eds) : *Les théories modernes de la mesure* [*Modern theories of measurement*]. Sainte-Foy: Presses de l'Université du Québec.

Raîche, G., and Blais, J.G. (2002, accepted). Étude de la distribution d'échantillonnage de l'estimateur du niveau d'habileté en testing adaptatif en fonction de deux règles d'arrêt dans le contexte de l'application du modèle de Rasch [Study of the sampling distribution of the proficiecy estimate in Rasch-based adaptive testing according to two stopping rules]. *Mesure et évaluation en éducation.*

Van der Linden, W.J. and Glas, C.A.W. (2000). Cross-validating item parameter estimation in adaptive testing. *In* A. Boomsma, M.A.J. van Duijn and T.A.B. Snijders (Eds): *Essays on item response theory*. New York: Springer.

Van der Linden, W.J. and Pashley, P.J. (2000). Item selection and ability estimation in adaptive testing. *In* W.J. van der Linder and C.A.W. Glas (Eds): *Computerized adaptive testing: theory and practice*. Dordrecht: Kluwer.

Wald, A. (1947). *Sequential analysis*. New York: Wiley.

Warm, T.A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427-450.

Acknowlegments

The authors express their gratitude to Jacques Saindon, professor at the University of Moncton, and Vera Ryan, teacher in New-Foundland, for their helpful comments.

Gilles Raîche	**Jean-Guy Blais**
Faculté des sciences de l'éducation	Faculté des sciences de l'éducation
Pavillon Jeanne-de-Valois	Université de Montréal
Université de Moncton	C.P. 6128, Succursale Centre-Ville
Moncton, NEW-BRUNSWICK, CANADA E1A 3E9	Montréal, QUÉBEC, CANADA H3C 3J7
Téléphone : (506) 858-4425	Téléphone: (514) 342-7527
Télécopie : (506) 858-4317	Télécopie : (514) 343-2497
email : raicheg@umoncton.ca	email : blaisjg@scedu.umontreal.ca
http://www.umoncton.ca/raicheg/	

* Report available at *http://www.umoncton.ca/raicheg/recherche/*

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM033791

I. DOCUMENT IDENTIFICATION:

Title: PRACTICAL CONSIDERATIONS ABOUT EXPECTED A POSTERIORI ESTIMATION IN ADAPTIVE TESTING: ADAPTIVE APRIORI, ADAPTIVE CORRECTION FOR BIAS, AND ADAPTIVE INTEGRATION INTERVAL	
Author(s): RAICHE, GILLES ; BLAIS, JEAN-GUY	
Corporate Source:	Publication Date: 4 APRIL 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 [X]	**Level 2A** [ ]	**Level 2B** [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: [signature]	Printed Name/Position/Title: GILLES RAÎCHE, PROF.	
Organization/Address: UNIVERSITY OF MONCTON EDUCATION MONCTON, NB E1A 3E9	Telephone: 506 858-4425	FAX: (506) 858-4317
	E-Mail Address: RAICHEG@UMONCTON.CA	Date: april 4, 2002

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org