Assigning Adaptive NAEP Booklets Based on State Assessment Scores: A Simulation Study of the Impact on Standard Errors

Bob Linn
Don McLaughlin
Tao Jiang
Larry Gallagher

Commissioned by the NAEP Validity Studies (NVS) Panel
May 2004

George W. Bohrnstedt, Panel Chair
Frances B. Stancavage, Project Director

The NAEP Validity Studies Panel was formed by the American Institutes for Research under contract with the National Center for Education Statistics. Points of view or opinions expressed in this paper do not necessarily represent the official positions of the U.S. Department of Education or the American Institutes for Research.

The NAEP Validity Studies (NVS) Panel was formed in 1995 to provide a technical review of NAEP plans and products and to identify technical concerns and promising techniques worthy of further study and research. The members of the panel have been charged with writing focused studies and issue papers on the most salient of the identified issues.

The authors would like to thank Beth Scarloss for her help in preparing this manuscript.

For Information:

NAEP Validity Studies (NVS)
American Institutes for Research
1791 Arastradero Road
Palo Alto, CA 94304-1337
Phone: 650/ 493-3550
Fax: 650/ 858-0958

Table of Contents

Purpose

The purpose of this simulation was to assess the improvements in estimates of standard errors that could be expected if students participating in NAEP were pre-assigned to test booklets that were adapted to their level of performance based on their state assessment scores.[1] Students in extreme quartiles would receive one regular NAEP block and one block adapted to their performance quartile. For their adapted block, students in the lowest quartile would receive an easier block, students in the highest quartile would receive a more difficult block, and students in the middle two quartiles would receive a second "regular" NAEP block. We also considered the impact of assigning adapted blocks only to the top and bottom deciles of students, rather than to full quartiles.

Procedure

Students

The student samples were drawn from databases of linked NAEP and state assessment scores for four states. All fourth-grade students with complete data were selected, resulting in four files of 2,234, 2,444, 2,345, and 2,103 students. Each state data set was analyzed separately. Our procedure involved first grouping the students into quartiles by their state assessment scores. The lowest quartile of students was designated the "low" group, the highest quartile was designated the "high" group, and the remainder were combined into a "middle" group. A similar procedure was conducted using the two extreme deciles to define the high and low groups, leaving a larger residual "middle" group. Due to discrete gaps in the upper tail of one state's score distribution (State #4), the top "quartile" in that state refers to the top 16 percent of the distribution and the top decile refers to the top 6 percent of the distribution. All other quartiles and deciles are within 1 percent of nominal values.

Item Blocks

In prior years, NAEP mathematics assessments used three 15-minute blocks for each booklet. Recently, NAEP has moved to a common block design in all subject areas, in which all subjects are tested using two 25-minute blocks. For the purposes of this simulation, it was necessary to construct synthetic mathematics blocks fitting these new parameters. Eight synthetic item blocks were assembled using item parameters from the 1996 grade 4 NAEP mathematics assessment.

NAEP item blocks are made up of dichotomous (right/wrong) items in combination with polytomous (constructed response) items. The overall item counts and the numbers of dichotomous and polytomous items in each the synthesized blocks are shown in Table 1.

[1] We will refer to this process as "adaptive testing."

Table 1: Distribution of item type by synthetic block number

Block Number	*Dichotomous Items*	*Polytomous Items*	*Total Items*
1	16	0	16
2	14	4	18
3	14	4	18
4 (Easy)	8	9	17
5 (Difficult)	20	1	21
6	12	6	18
7	14	2	16
8	19	1	20
Total	117	27	144

After establishing baseline simulation runs, item difficulty (b-parameters) for Block #4 and Block #5 were adjusted by adding or subtracting a constant to synthesize an easier block (#4) and a more difficult block (#5). Three different constants were used to adjust block difficulty. In the most extreme case the parameters were shifted by ± 1.65, which corresponds roughly to the 5^{th}/95^{th} percentile of the ability distribution. This created a pair of synthetic item blocks in which the b-parameter distributions were centered within the extreme deciles of the ability distribution. Two other, more moderate, transformations were created by shifting the original b-parameters ± 0.83 or ± 0.55. These adjustments centered the resultant b-parameter distributions at roughly the 20^{th}/80^{th} and the 30^{th}/70^{th} percentiles of the ability distribution, respectively.

Score Estimation

Synthetic booklets consisting of two NAEP item blocks were assigned to each student. Students with the lowest state test performance were always assigned Block #4 plus another random block (excluding Block #5), while students with the highest state test performance were assigned Block #5 plus another random block (excluding Block #4).[2] Students in the middle group were assigned two random blocks from the set {#1, #2, #3, #6, #7, and #8}.

Using the operational NAEP item parameters, and then the revised parameters for Blocks #4 and #5, simulated item responses for each student were generated based on the student's posterior mean theta on NAEP. Item skips and did-not-reach scores were also simulated. These simulated item responses were processed through the software program PARSCALE to generate score estimates and standard errors for each student. The simulation was replicated 250 times for each state and each combination of item parameters. The overall distribution of standard errors across replications was then computed for the low, middle, and high groups of students.

[2] Selected analyses were repeated using blocks #1 and #3 as the easier block, with no change in results.

Procedural Summary

In summary, this simulation uses two different grouping criteria for students (quartiles and deciles) based on state assessment scores. We also used four variations of the b-parameter consisting of:

1. the original parameters,
2. adjustments of ± 1.65 to the item parameters for easy and difficult blocks,
3. adjustments of ±0.83 to the item parameters for easy and difficult blocks, and
4. adjustments of ± 0.55 to the item parameters for easy and difficult blocks.

This gave us 8 sets of standard errors to report for each state.

Results

Results from this simulation indicate that, as predicted, assigning students "easier" test booklets based on their prior test performance can reduce standard errors in measuring students' ability levels. The reverse, however, was found to be true for "harder" booklets; in this simulation, assigning "harder" test booklets actually increased standard error estimates. Results were consistent across the four states included in this study.

Note that it is entirely feasible to adapt the assessment at only one end of the distribution (e.g., create easier booklets for students at the low end of the distribution, but randomly assign regular booklets to students in the middle *and* at the high end of the distribution.). In fact, this condition was also simulated in our study. The results are not reported because they essentially replicate the corresponding results from the other simulations. That is, in the "one-sided" adaptation, students in the lowest quartile performed the same as in condition 2, and students in the top quartile performed the same as in condition 1.

Standard Error Reductions

The key results are presented for each of the four states in Tables 2, 3, 4, and 5. Each table presents simulated results for quartiles and deciles of students for each of the four variations in item parameters discussed above. Recall that the original goal of this study was to determine whether it would be possible to use adaptive testing to generate estimates of student performance at the extremes of the distribution that are at least as accurate as estimates from the middle of the ability distribution. Using the initial item parameters and quartile grouping, we see that the mean estimated standard errors for the low-group students were .38, .34, .32, and .31 in the four states. When the b-parameters for Block #4 were shifted by -1.65, the standard errors in the low groups were reduced to .29, .28, .28, and .28. These values represent reductions of 23 percent, 18 percent, 14 percent, and 12 percent, and bring the standard errors for the low group in line with those observed for the middle and high groups.

The impact of smaller shifts in the b-parameters was also explored. A shift of -.83 (creating a synthetic block with a difficulty distribution that corresponds roughly to the performance of the bottom 40 percent of students) resulted in reductions in the original mean standard errors of 15 percent, 14 percent, 13 percent, and 12 percent for lowest-quartile students. A shift of -.55 (creating a difficulty distribution that corresponds roughly to the performance of the bottom 60 percent of students) resulted in reductions of 10 percent, 10 percent, 9 percent, and 9 percent. Thus we see that reductions in standard errors among the lowest performing students, while greater with larger shifts in difficulty parameters, did occur for each of the three modifications tested.

At the upper end of the ability distribution, the original standard errors were more similar to those of the middle group of students than were the estimates for the lowest performing students. The original standard errors for the middle performance group were approximately .32, .30, .29, and .28, while original standard errors for the highest quartile were .28, .29, .28, and .28. Mean standard errors for the high group were *increased* by the addition of 1.65 to the b-parameter, or item difficulty. Thus, rather than improving standard errors, increasing item difficulty *decreased* the accuracy of ability estimates among the highest performing group. This is an indication that a shift of this magnitude pushes the peak of the test information function well past the mean ability level for the group. Note that as the b-parameter shift is reduced to .83 and .55, the standard errors of the high group approach those of the original condition.

The *lack* of difference across conditions in mean standard errors for the middle group is also noteworthy. Since, in each case, middle-group students were assigned unmodified item blocks (blocks with original item parameters), we did not expect an effect on their standard errors. However, the high level of consistency (identical to the fourth decimal place in several of the conditions) offer reassurance that the differences across conditions found in the high and, particularly, in the low groups are more likely to be real and less likely to be due to random error.

A similar, but stronger, pattern of reductions in the standard errors of students in the low group was found when the adjusted booklets were administered to only the lowest decile. Most notably, a b-parameter shift of 1.65 reduced the means of the low groups' standard errors by 25 percent, 23 percent, 20 percent, and 19 percent in the four states. At the high end, restricting analysis to the top decile improves standard errors slightly, but the parameter adjustments still produce higher standard errors than the original condition.

Table 2: Mean standard errors and standard deviations over 250 simulations, State #1

State Quartile Grouping									
	Lowest Quartile			*Middle*			*Highest Quartile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.3801	0.0020	100.0%	0.3167	0.0011	100.0%	0.2841	0.0010	100.0%
± 1.65	0.2927	0.0008	77.0%	0.3166	0.0010	100.0%	0.3256	0.0012	114.6%
± .83	0.3232	0.0015	85.0%	0.3166	0.0010	100.0%	0.3024	0.0010	106.4%
± .55	0.3409	0.0016	89.7%	0.3167	0.0011	100.0%	0.2952	0.0013	103.9%
State Decile Grouping									
	Lowest Decile			*Middle*			*Highest Decile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.4028	0.0032	100.0%	0.3221	0.0010	100.0%	0.2834	0.0016	100.0%
± 1.65	0.3023	0.0015	75.0%	0.3221	0.0009	100.0%	0.3170	0.0018	111.8%
± .83	0.3426	0.0024	85.1%	0.3221	0.0009	100.0%	0.2945	0.0016	103.9%
± .55	0.3620	0.0027	89.9%	0.3220	0.0009	100.0%	0.2865	0.0014	101.1%

Table 3: Mean standard errors and standard deviations over 250 simulations, State #2

State Quartile Grouping									
	Lowest Quartile			*Middle*			*Highest Quartile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.3444	0.0018	100.0%	0.3038	0.0010	100.0%	0.2925	0.0010	100.0%
± 1.65	0.2816	0.0006	81.8%	0.3038	0.0010	100.0%	0.3393	0.0014	116.0%
± .83	0.2963	0.0011	86.0%	0.3038	0.0010	100.0%	0.3151	0.0014	107.7%
± .55	0.3098	0.0014	90.0%	0.3037	0.0010	100.0%	0.3074	0.0013	105.1%
State Decile Grouping									
	Lowest Decile			*Middle*			*Highest Decile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.3696	0.0030	100.0%	0.3090	0.0008	100.0%	0.2888	0.0015	100.0%
± 1.65	0.2849	0.0011	77.1%	0.3089	0.0008	100.0%	0.3305	0.0020	114.5%
± .83	0.3121	0.0021	84.4%	0.3090	0.0007	100.0%	0.3070	0.0019	106.3%
± .55	0.3293	0.0024	89.1%	0.3089	0.0008	100.0%	0.2989	0.0019	103.5%

Table 4: Mean standard errors and standard deviations over 250 simulations, State #3

State Quartile Grouping									
	Lowest Quartile			*Middle*			*Highest Quartile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.3206	0.0017	100.0%	0.2877	0.0009	100.0%	0.2848	0.0010	100.0%
± 1.65	0.2764	0.0006	86.2%	0.2877	0.0009	100.0%	0.3229	0.0011	113.4%
± .83	0.2801	0.0009	87.4%	0.2877	0.0009	100.0%	0.3003	0.0010	105.4%
± .55	0.2904	0.0011	90.6%	0.2877	0.0009	100.0%	0.2919	0.0013	102.5%
State Decile Grouping									
	Lowest Decile			*Middle*			*Highest Decile*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	0.3504	0.0031	100.0%	0.2925	0.0008	100.0%	0.2873	0.0016	100.0%
± 1.65	0.2785	0.0010	79.5%	0.2924	0.0008	100.0%	0.3190	0.0018	111.0%
± .83	0.2971	0.0019	84.8%	0.2925	0.0008	100.0%	0.2955	0.0019	102.9%
± .55	0.3126	0.0023	89.2%	0.2925	0.0008	100.0%	0.2887	0.0015	100.5%

Table 5: Mean standard errors and standard deviations over 250 simulations, State #4

State Quartile Grouping									
	Lowest Quartile (26%)			*Middle (58%)*			*High Quartile (16%)*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	.3144	.0021	100.0%	.2825	.0018	100.0%	.2849	.0019	100.0%
± 1.65	.2755	.0006	87.6%	.2834	.0008	100.3%	.3186	.0014	111.8%
± .83	.2769	.0011	88.1%	.2835	.0008	100.4%	.2966	.0012	104.1%
± .55	.2860	.0012	91.0%	.2835	.0007	100.4%	.2888	.0013	101.3%
State Decile Grouping									
	Lowest Decile (10%)			*Middle (84%)*			*High Decile (6%)*		
Shift in b-parameters	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*	*Mean SE*	*SD*	*% of Original*
Original b-parameters	.3439	.0035	100.0%	.2893	.0023	100.0%	.2900	.0029	100.0%
± 1.65	.2773	.0010	80.6%	.2902	.0007	100.3%	.3180	.0019	109.7%
± .83	.2936	.0020	85.4%	.2903	.0007	100.3%	.2949	.0023	101.7%
± .55	.3075	.0024	89.4%	.2901	.0007	100.3%	.2898	.0021	99.9%

Not-Reached Items

An important factor influencing standard errors in the NAEP context is the presence of missing data due to not-reached items.[3] The simulation included simulation of not-reached and omitted items. An examination of the items not reached in each simulation condition shed some light on the reasons for the observed changes in standard error. We examined the not-reached items in one state (State #4). Recall that Block #4, which became the easy block, had a total of 17 items; Block #5, which became the difficult block, had 21 items. As shown in Table 6, with quartile grouping and original item parameters, we see that in Block #4, 4.0 percent of the low group students fail to reach item 16, and 7.9 percent do not reach item 17. When the b-parameter is shifted by 1.65 for this block, 3.7 percent fail to reach item 16 or item 17. While this does represent a decrease in the proportion not reaching the last item in the block, this would not be expected to have a large impact on standard errors given that there are only 17 items in the block and the "not reached" simulation assumes that all students respond through item 15.

The effect is more dramatic in the difficult block (Block #5). Under the baseline condition, 10.6 percent of high group students fail to reach item 17 or beyond in this 21-item block. With a b-parameter shift of 1.65, this number jumps to 19.0 percent. Further, while under baseline conditions only 11 percent of students fail to reach the last item (item 21), under a high b-parameter shift fully 34.4 percent of respondents fail to reach the last item. Fewer items reached for these respondents can be expected to result in a higher standard error for this group, as the ability estimate is based on less information.

[3] Not-reached items are defined as unanswered items at the ends of item blocks and are treated by NAEP as missing data. Omitted items are unanswered items in the middle of item response strings, and these are treated by NAEP as incorrect responses.

Table 6: Mean percent of not-reached responses (SD) for each item in modified blocks, State #4

State Quartile Grouping								
	Easy Block #4		*Difficult Block #5*					
Shift in b-parameters	*Item 16*	*Item 17*	*Item 16*	*Item 17*	*Item 18*	*Item 19*	*Item 20*	*Item 21*
Original b-parameters	4.0% (0.9%)	7.9% (1.3%)	0.0% (0.0%)	10.6% (1.7%)	10.6% (1.7%)	10.6% (1.7%)	11.0% (2.6%)	11.0% (2.6%)
± 1.65	3.7% (1.0%)	3.7% (1.0%)	10.0% (1.8%)	19.0% (2.2%)	19.0% (2.2%)	27.0% (2.3%)	34.4% (2.6%)	34.4% (2.6%)
± .83	4.0% (0.9%)	7.9% (1.0%)	8.7% (3.3%)	18.1% (3.3%)	18.1% (3.3%)	25.1% (4.2%)	32.9% (4.3%)	32.9% (4.3%)
± .55	3.9% (0.8%)	7.7% (1.2%)	0.0% (0.0%)	10.4% (1.6%)	10.4% (1.6%)	10.4% (1.6%)	19.0% (2.2%)	19.0% (2.2%)
State Decile Grouping								
	Easy Block #4		*Difficult Block #5*					
Shift in b-parameters	*Item 16*	*Item 17*	*Item 16*	*Item 17*	*Item 18*	*Item 19*	*Item 20*	*Item 21*
Original b-parameters	3.7% (1.2%)	7.6% (1.4%)	0.0% (0.0%)	10.4% (2.9%)	10.4% (2.9%)	10.4% (2.9%)	10.8% (3.8%)	10.8% (3.8%)
± 1.65	4.0% (1.4%)	4.0% (1.4%)	9.9% (2.9%)	18.5% (3.7%)	18.5% (3.7%)	26.8% (4.6%)	34.2% (4.8%)	34.2% (4.8%)
± .83	4.0% (1.3%)	7.8% (1.9%)	4.2% (5.2%)	14.4% (5.5%)	14.4% (5.5%)	15.7% (6.4%)	25.2% (5.8%)	25.2% (5.8%)
± .55	3.9% (1.4%)	7.7% (2.1%)	0.0% (0.0%)	10.0% (3.1%)	10.0% (3.1%)	10.0% (3.1%)	18.6% (3.5%)	18.6% (3.5%)

Discussion

In summary, the use of an easier item block for lower-ability students appears to have the desired effect of reducing standard errors for this population. However, it should be noted that the standard error estimates reported for this simulation are based on PARSCALE results, which do not incorporate "conditioning."[4] Conditioning is designed to reduce standard errors. It may be that the advantage of the adaptive assignment of an easier block is not as great when applied in the context of conditioning. To explore this matter further, an analysis is currently under way of the relationship between block difficulty and standard errors for low-group students (identified by means of their state assessment scores) who took the 2003 operational NAEP assessment. The analysis is feasible because the item blocks used in the 2003 assessment show variation in difficulty, although the assignment of students to item blocks was random.

Improving the precision of measurement among low performing groups of students has the potential to make NAEP more useful to state policymakers in an era in which educational policy has been increasingly concerned with reducing performance gaps between low-performing and other students. Particularly with state samples, the sizes of

[4] See the 1996 NAEP Technical Report for details.

the current standard errors (which combine both sampling error and measurement error) have often precluded the possibility of meaningful statistical comparisons between many of the groups of interest. If the finding of an improvement in standard errors through adapting the difficulty of the item blocks is sustained even after accounting for conditioning, then we recommend that NAEP seriously consider the use of adaptive block assignment based on state assessment scores. In order to allow data from all students to be scaled together, however, it is necessary to maintain an overlap between the items presented to the students in the adapted condition and all other students. Thus, an appropriate design must be similar to the one that was used in the simulation, in which only one of the two item blocks presented to a student was purposefully assigned on the basis of difficulty. The other item block represented a random draw from the full NAEP item pool.

With regard to the use of more difficult item blocks for students at the top of the distribution, the results of the simulation suggest that such a practice would be counterproductive, since the use of an equivalently more difficult block for higher-ability students *increased* standard errors. From a pure measurement error perspective, adaptive testing for high-achieving students is, in any event, unnecessary since the operational item parameters generate standard errors for this group that are as good or better than for students in the middle of the distribution. If, however, one wishes to take advantage of adaptive testing to introduce more challenging items that are better aligned with advanced achievement level expectations, then the parameters for such item blocks will need to be tailored more carefully to students' expected abilities. The simulation reported here used extant NAEP items, as written, and shifted their difficulty parameters. Further improvements for either low- or high-performing students may be observed by carefully constructing specific item blocks for these groups.

References

Allen, N.L., Carlson, J.E., and Zelenak, C.A. (1999). *The NAEP 1996 Technical Report.* Washington, DC: National Center for Education Statistics.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychologic

Appendix

For the reader's reference, the original item parameters for Blocks #4 and #5 are presented below. The ***a***, ***b***, and ***c parameters*** are from the familiar three-parameter logistic (3PL) model. The parameters $\boldsymbol{d_0}$ through $\boldsymbol{d_4}$ represent offsets to the ***b***-parameter in a generalized partial credit model[5].

Table A-1 – Original item parameters for block #4

Item	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
1	0.74	-0.125	0.274					
2	1.412	0.422	0					
3	0.59	-0.14	0.191					
4	1.037	-0.018	0.09					
5	1.026	-0.098	0					
6	1.241	0.761	0.287					
7	0.793	0.962	0.359					
8	1.014	-0.387	0					
9	0.417	-0.407	0	0	-5.308	5.308		
10	0.763	0.803	0	0	-0.196	0.197		
11	0.746	0.455	0	0	0.253	-0.253		
12	0.548	1.55	0	0	-1.186	1.186		
13	0.505	0.719	0	0	0.69	-0.69		
14	0.598	0.431	0	0	-0.012	0.012		
15	0.58	0.906	0	0	-0.233	0.233		
16	0.693	1.463	0	0	-0.952	0.952		
17	0.563	0.693	0	0	1.946	-0.362	-1.007	-0.577

[5] See the 1996 NAEP Technical Report for details. Equation 11.3 and text are excerpted below:

The polytomous items were scaled using a generalized partial credit model (Muraki, 1992). The fundamental equation of this model is the probability that a person with proficiency θ_k on scale k will have, for the jth item, a response x_j that is scored in the ith of m_j ordered score categories:

$$P(x_j = i \mid \theta_k, a_j, b_j, d_{j,1}, \mathrm{K}\ , d_{j,m_j-1}) = \frac{\exp\left(\sum_{v=0}^{i} 1.7a_j(\theta_k - b_j + d_{j,v})\right)}{\sum_{g=0}^{m_j-1} \exp\left(\sum_{v=0}^{g} 1.7a_j(\theta_k - b_j + d_{j,v})\right)} \equiv P_{ji}(\theta_k) \qquad (11.3)$$

Where

- m_j is the number of categories in the response to item j
- x_j is the response to item j, with possibilities 0,1,...,m_{j-1}
- a_j is the slope parameter;
- b_j is the item location parameter characterizing overall difficulty; and
- $d_{j,i}$ is the category i threshold parameter.

Table A-1 – Original item parameters for block #5

Item	*a*	*b*	*c*	*d0*	*d1*	*d2*	*d3*	*d4*
1	0.766	-1.617	0.209					
2	0.899	1.694	0.134					
3	0.992	0.29	0					
4	0.777	0.448	0.322					
5	1.005	1.632	0					
6	0.493	-1.277	0.162					
7	0.792	-0.217	0					
8	0.824	0.778	0					
9	0.821	-0.194	0.164					
10	0.826	-0.285	0.139					
11	0.634	0.689	0.093					
12	1.396	1.245	0.204					
13	0.928	1.817	0.366					
14	0.359	-3.271	0.209					
15	1.39	1.102	0.174					
16	1.414	0.767	0.223					
17	1.052	2.723	0.184					
18	0.307	-0.719	0.254					
19	1.383	0.539	0.115					
20	1.056	1.472	0.235					
21	0.444	-1.026	0	0	-0.822	-1.91	2.733	