# AERA 2010 Conference
# Denver, Colorado, USA

THE MODELING OF RESPONSE STYLE BIAS: AN ANSWER TO THE ATTITUDE-ACHIEVEMENT PARADOX?

Eva Van de gaer

Ray Adams

Australian Council for Educational Research

Table of Contents

# THE MODELING OF RESPONSE STYLE BIAS: AN ANSWER TO THE ATTITUDE-ACHIEVEMENT PARADOX?

## INTRODUCTION

1. One of the major challenges of an international study such as PISA is the cross-cultural validity and applicability of all instruments. In this context, a phenomenon has been of concern which has continued to appear across all PISA cycles whereby for a number of attitudinal student context constructs have shown to be linked to performance in unexpected ways. More specifically, at the between-country level, countries that demonstrate higher performance in a subject show less positive attitudes towards that subject whereas more positive attitudes are recorded for lower-performing countries. We will refer to this phenomenon as the attitude-achievement paradox.
2. Table 1 gives an overview of within- and between-country correlations between student context constructs and mathematics achievement in PISA 2003.

**Table 1 Correlations between student context constructs and mathematics achievement at the within-country and the between-country level, PISA 2003**

| Construct name | Construct label | B'ween country corr. | Within-country correlation | Item format (4 or 5 point Likert scale) |
|---|---|---|---|---|
| **Student background** | | | | |
| HEDRES | Educational resources at home | 0.67 | 0.26 | Tick box if resource at home |
| CULTPOSS | Cultural possessions at home | 0.25 | 0.25 | Tick box if possession at home |
| COMPHOME | Possession of a computer at home | 0.77 | 0.28 | Tick box if computer at home |
| *Self-concept* | | | | |
| **SCMAT** | **Mathematics self-concept** | **-0.20** | **0.34** | **Str. agree – Str. disagree** |
| *Self-efficacy* | | | | |
| MATHEFF | Mathematics self-efficacy | 0.45 | 0.49 | Very confident – Not at all confident |
| *ICT self-efficacy* | | | | |
| HIGHCONF | ICT: Confidence in high-level tasks | 0.38 | 0.10 | I can do this very well by myself– I don't know what this means |
| INTCONF | ICT: Confidence in internet tasks | 0.67 | 0.21 | I can do this very well by myself – I don't know what this means |
| ROUTCONF | ICT: Confidence in routine tasks | 0.63 | 0.31 | I can do this very well by myself – I don't know what this means |
| *Interest and motivation* | | | | |
| **ATSCHL** | **Attitudes towards school** | **-0.72** | **0.05** | **Str. agree – Str. disagree** |
| ATTCOMP | ICT: Attitudes towards computers | -0.04 | 0.03 | Str. agree – Str. disagree |
| **INSTMOT** | **Instrumental motivation in mathematics** | **-0.57** | **0.13** | **Str. agree – Str. disagree** |
| **INTMAT** | **Interest in mathematics** | **-0.74** | **0.16** | **Str. agree – Str. disagree** |
| *Learning strategies* | | | | |
| **CSTRAT** | **Control Strategies** | **-0.57** | **0.06** | **Str. agree – Str. disagree** |
| **ELAB** | **Elaboration Strategies** | **-0.80** | **0.02** | **Str. agree – Str. disagree** |
| **MEMOR** | **Memorisation Strategies** | **-0.77** | **-0.02** | **Str. agree – Str. disagree** |
| *School climate* | | | | |
| BELONG | Sense of belonging at school | -0.07 | 0.06 | Str. agree – Str. disagree |
| DISCLIM | Disciplinary climate in math lessons | 0.20 | 0.19 | Every lesson – never or hardly ever |

| Construct name | Construct label | B'ween country corr. | Within-country correlation | Item format (4 or 5 point Likert scale) |
|---|---|---|---|---|
| **STUREL** | **Student-teacher relations at school** | **-0.58** | **0.02** | **Str. agree – Str. disagree** |
| **TEACHSUP** | **Teacher support in math lessons** | **-0.59** | **-0.01** | **Every lesson – never or hardly ever** |
| *Engagement in learning activities* | | | | |
| INTUSE | ICT: Internet/entertainment use | 0.39 | 0.04 | Almost every day - Never |
| PRGUSE | ICT: Programs/software use | -0.28 | -0.02 | Almost every day - Never |
| *Other attitudes* | | | | |
| ANXMAT | Mathematics anxiety | -0.57 | -0.33 | Str. agree – Str. disagree |
| **COMPLRN** | **Competitive learning** | **-0.65** | **0.07** | **Str. agree – Str. disagree** |
| **COOPLRN** | **Co-operative learning** | **-0.51** | **-0.01** | **Str. agree – Str. disagree** |

Note: Constructs displaying positive correlations with achievement at the between-student within-country level and negative correlations with achievement at the between-country level are given in bold. Likewise, constructs showing correlations with achievement at the between-student within-country level that are not substantially different from zero which become considerably negative at the between-country level are given in bold.

3. A central observation can be made from Table 1 is that constructs which are based on items with a more concrete reference point do not show the paradox whereas those based on items with more general and abstract reference points tend to show the phenomenon. Thus, none of the constructs related to ICT show the phenomenon as the explicitly relate to computers, software and the internet. Likewise, HEDRES, COMPHOME and CULPOSS are constructed from questions in which students have to indicate whether or not specific items are available at home. It is also interesting to note that while self-concept shows the phenomenon, self-efficacy does not show it. Again, it can be argued not to emerge for self-efficacy as items measuring that construct ask students how comfortable they feel in answering specific mathematics problems (e.g. solving a specific equation, telling the price of a TV set after a 30% reduction) whereas self-concept items students about mathematics more generally (e.g. learning mathematics quickly, understanding the most difficult work easily).
4. Cross-cultural differences in response styles are considered to be a serious source of bias in international surveys using Likert items. Several types of response styles have been described (e.g. Greenleaf, 1992, Clarke, 2000; Johnson & al., 2005). All of them can make it difficult to distinguish authentic cultural differences from "stylistic" biases in respondent behaviour (Van de Vijver & Poortinga, 1997; van Hemert, Poortinga & van de Vijver, 2007).
   Previous analyses of the attitudinal constructs used in PISA2006 (Buckley, 2008) have provided evidence of cross-national differences in these response styles.
5. The present study wants to explore whether the paradox is a result of country or cultural differences in response styles.

## RESEARCH QUESTION

6. We hypothesize that a response style bias is at least partially responsible for the attitude-achievement paradox (see Figure 1). In order to test this hypothesis, we assume that attitudinal indices that show the phenomenon described in this paper are not only determined by the 'real' attitudes of students but also by an overarching "superfactor" (Lie and Turmo, 2005) of response style bias.

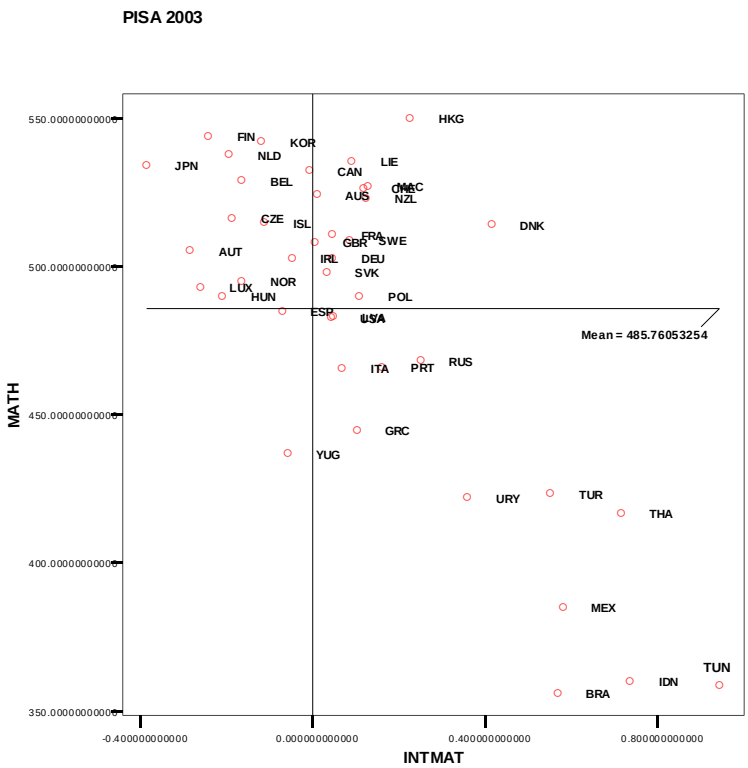


**Figure 1 Between-country level correlations between interest in mathematics and mathematics performance in PISA 2003**

## MODEL AND METHOD

In a first step, we selected 11 countries which participated in PISA 2003. These are AUS, BRA, FIN, FRA, DEU, HKG, KOR, IDN, IRL, JPN, and TUN. We based the selection of these countries on the position they hold in Figure 1.

Three groups of countries can be identified:

1) countries that show a high mean interest in mathematics but low performance. Examples are TUN, IDN, and BRA.
2) countries that show a low mean interest in mathematics but a high performance. Examples are KOR, JPN, HKG, and FIN. Note: HKG shows a higher mean interest in mathematics than the other countries in this group.
3) countries that show a mean interest in mathematics and a mean performance. Examples are AUS, FRA, DEU, and IRL.

From these 3 groups of countries, we selected a number of countries for our analyses.

7. In the second step, we selected a number of constructs to be included in our conceptual model. In the first analyses, we included the constructs INMAT, SCMAT, and MATHEFF in our model. Figure 2 presents the statistical model. We used Mplus to analyse the data. We performed a confirmatory factor analysis that involved the estimation of:

a) the loadings of the items ST30Q01, ST30Q03, ST30Q04, and ST30Q06 on the latent variable 'INTMAT',
b) the loadings of the items ST32Q02, ST32Q04, ST32Q06, ST32Q07, and ST32Q09 on the latent variable 'SCMATH',
c) the loadings of the items ST31Q01, ST31Q02, ST31Q03, ST31Q04, ST31Q05, ST31Q06, ST31Q07, and ST31Q08 on the latent variable 'MATHEFF".

d) the loadings of the items ST30Q01, ST30Q03, ST30Q04, ST30Q06, and, ST32Q02, ST32Q04, ST32Q06, ST32Q07, ST32Q09 on the latent variable 'RESPONSE BIAS' (we restricted the loadings of the items ST31Q01, ST31Q02, ST31Q03, ST31Q04, ST31Q05, ST31Q06, ST31Q07, and ST31Q08 – items loading on MATHEFF - to zero on the latent variable 'RESPONSE BIAS').
e) variances, and covariances between the latent constructs with the restriction that the covariances between the latent variable 'RESPONSE BIAS' and the other latent variables are set to zero
f) residual variances
g) factor scores.

8. We hypothesize that the attitude items are indicators of not only content constructs but also a response bias construct. A response bias factor may be responsible for the negative cross-country correlation of math attitudes such as interest in math and math self-concept with math performance. We expect that after taken into account the response bias factor, the negative between country correlations will become non-significant or – at least – will reduce in size.

9. Because math efficacy did not show the paradoxical pattern (i.e., switch in sign in correlations within and between countries with math performance), we restricted the loading of the items, which are indicators of the latent construct 'math efficacy', on the latent variable 'response bias' to zero. All the items scores (except ST32Q02) were inverted so that a higher item score corresponds to a higher attitude.

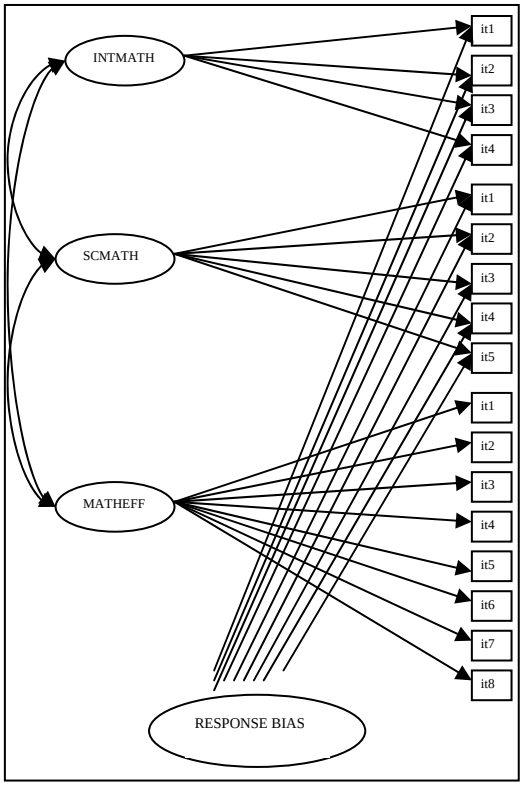

Figure 2. Conceptual Model

10. We analysed the data with multiple group confirmatory factor analyses using Mplus.

## RESULTS

### *WITHIN COUNTRY CORRELATIONS*

11. Tables 2 to 5 show the within country correlations between the constructs interest in math, math self-concept, math efficacy, and response bias with math performance. We also included the within country correlations between the original PISA 2003 constructs for comparison.
12. Overall, the results show that when we take into account response bias, the within country correlations between interest in math, math self-concept, math efficacy math performance become somewhat larger after taking into account response bias. For BRA and IDN, the within country correlation between intmat and pvmath even changes from negative to positive. The results seem to suggest that we are taken some bias into account.
13. The size of the increase in correlations seems to be larger for F_INTMATH than for F_SCMATH and F_MATHEFF suggesting that the interest in math index shows more bias than the other indices.
14. Table 4 shows that for MATHEFF the within-country correlations become somewhat larger after taking into account response style bias but the increase is very modest. This is what we expected as we hypothesized (see Conceptual model in Figure 2) that the response style factor would not affect MATHEFF.
15. Table 5 presents the within-country correlations between the response bias factor and pvmath. For most countries, the correlations were negative indicating that in these countries the lower achieving students show more response bias. The highest negative correlations were found for the lower achieving countries such as BRA, TUN, and IDN (with the exception of DEU). We only found positive correlations for FIN, HKG, and KOR. Except for FIN, these countries are all (East-) Asian countries and are all high performing countries. The positive correlation indicates that the higher performing students show more response bias.
16. All the correlations were estimated using final students weights and BRR.

Table 2. Correlations between interest in math and math performance in PISA 2003 for 11 countries

| CORR F_INTMATH with PVMATH (estimated with Model in Figure 2) | | | | | CORR INTMAT with PVMATH (original correlations) | | | |
|---|---|---|---|---|---|---|---|---|
| CNT | CORR | SE | NU_cases* | N_cases | CORR | SE | NU_cases* | N_cases |
| AUS | 0.34 | 0.01 | 12421 | 233576 | 0.19 | 0.01 | 12421 | 233576 |
| BRA | 0.39 | 0.03 | 4231 | 1849670 | -0.12 | 0.03 | 4231 | 1849670 |
| DEU | 0.23 | 0.02 | 4424 | 828450 | 0.12 | 0.02 | 4424 | 828450 |
| FIN | 0.49 | 0.01 | 5697 | 56789 | 0.33 | 0.02 | 5697 | 56789 |
| FRA | 0.33 | 0.02 | 4237 | 722862 | 0.22 | 0.02 | 4237 | 722862 |
| HKG | 0.38 | 0.01 | 4465 | 72237 | 0.30 | 0.01 | 4465 | 72237 |
| IDN | 0.09 | 0.03 | 10475 | 1927137 | -0.07 | 0.03 | 10475 | 1927137 |
| IRL | 0.32 | 0.02 | 3826 | 53946 | 0.20 | 0.02 | 3826 | 53946 |
| JPN | 0.28 | 0.02 | 4688 | 1234840 | 0.28 | 0.02 | 4688 | 1234840 |
| KOR | 0.49 | 0.01 | 5438 | 532850 | 0.39 | 0.01 | 5438 | 532850 |
| TUN | 0.28 | 0.02 | 4670 | 149221 | 0.10 | 0.02 | 4670 | 149221 |

* NU_cases: the number of unweighted cases; N_cases: the number of weighted cases.

Table 3. Correlations between math self-concept and math performance in PISA 2003 for 11 countries

| CORR F_SCMATH with PVMATH (estimated with Model in Figure 2) | | | | | CORR SCMATH with PVMATH (original correlations) | | | |
|---|---|---|---|---|---|---|---|---|
| CNT | CORR | SE | NU_cases* | N_cases | CORR | SE | NU_cases* | N_cases |
| AUS | 0.47 | 0.01 | 12403 | 233438 | 0.41 | 0.01 | 12403 | 233438 |
| BRA | 0.33 | 0.02 | 4248 | 1856797 | 0.21 | 0.02 | 4248 | 1856797 |
| DEU | 0.31 | 0.02 | 4416 | 826828 | 0.27 | 0.02 | 4416 | 826828 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| FIN | 0.60 | 0.01 | 5766 | 57599 | 0.58 | 0.01 | 5766 | 57599 |
| FRA | 0.36 | 0.02 | 4212 | 718772 | 0.32 | 0.02 | 4212 | 718772 |
| HKG | 0.40 | 0.02 | 4463 | 72204 | 0.35 | 0.02 | 4463 | 72204 |
| IDN | 0.01 | 0.03 | 10531 | 1933486 | -0.05 | 0.03 | 10531 | 1933486 |
| IRL | 0.43 | 0.02 | 3823 | 53919 | 0.38 | 0.02 | 3823 | 53919 |
| JPN | 0.23 | 0.02 | 4684 | 1233773 | 0.20 | 0.02 | 4684 | 1233773 |
| KOR | 0.52 | 0.01 | 5435 | 532582 | 0.46 | 0.01 | 5435 | 532582 |
| TUN | 0.31 | 0.02 | 4653 | 148696 | 0.28 | 0.02 | 4653 | 148696 |

* NU_cases: the number of unweighted cases; N_cases: the number of weighted cases.

Table 4. Correlations between math efficacy and math performance in PISA 2003 for 11 countries

| | CORR F_MATHEFF with PVMATH (estimated with Model in Figure 2) | | | | CORR MATHEFF with PVMATH (original correlations) | | | |
|---|---|---|---|---|---|---|---|---|
| CNT | CORR | SE | NU_cases* | N_cases | CORR | SE | NU_cases* | N_cases |
| AUS | 0.55 | 0.01 | 12418 | 233600 | 0.52 | 0.01 | 12418 | 233600 |
| BRA | 0.33 | 0.03 | 4292 | 1875654 | 0.31 | 0.04 | 4292 | 1875654 |
| DEU | 0.53 | 0.01 | 4419 | 827415 | 0.51 | 0.02 | 4419 | 827415 |
| FIN | 0.58 | 0.01 | 5697 | 56752 | 0.52 | 0.01 | 5697 | 56752 |
| FRA | 0.53 | 0.01 | 4210 | 718444 | 0.50 | 0.01 | 4210 | 718444 |
| HKG | 0.59 | 0.01 | 4464 | 72198 | 0.56 | 0.02 | 4464 | 72198 |
| IDN | 0.11 | 0.03 | 10524 | 1935425 | 0.11 | 0.03 | 10524 | 1935425 |
| IRL | 0.54 | 0.01 | 3828 | 53968 | 0.53 | 0.01 | 3828 | 53968 |
| JPN | 0.59 | 0.02 | 4687 | 1234469 | 0.59 | 0.02 | 4687 | 1234469 |
| KOR | 0.61 | 0.01 | 5436 | 532692 | 0.58 | 0.01 | 5436 | 532692 |
| TUN | 0.40 | 0.02 | 4516 | 144259 | 0.37 | 0.02 | 4516 | 144259 |

* NU_cases: the number of unweighted cases; N_cases: the number of weighted cases.

Table 5. Correlations between response bias and math performance in PISA 2003 for 11 countries

| | CORR F_RESPBIAS with PVMATH | | | |
|---|---|---|---|---|
| CNT | CORR | SE | NU_cases | N_cases |
| AUS | -0.02 | 0.01 | 12460 | 234233 |
| BRA | -0.27 | 0.02 | 4344 | 1899581 |
| DEU | -0.10 | 0.02 | 4440 | 831613 |
| FIN | 0.01 | 0.02 | 5792 | 57831 |
| FRA | -0.02 | 0.02 | 4259 | 727263 |
| HKG | 0.01 | 0.01 | 4466 | 72259 |
| IDN | -0.13 | 0.03 | 10723 | 1966952 |
| IRL | -0.03 | 0.02 | 3843 | 54207 |
| JPN* | | | 4691 | 1235779 |
| KOR | 0.12 | 0.01 | 5441 | 533182 |
| TUN | -0.11 | 0.02 | 4704 | 150315 |

* We restricted the variance of the response bias factor to zero for Japan due to convergence problems (i.e., negative variance for Japan).

### *BETWEEN COUNTRY CORRELATIONS*

17. In a next step, we will look at the cross-country correlations between math achievement and attitudes. We aggregated the factor scores that were estimated using Mplus by country and

correlated them with each of the five aggregated PV's for math. Then we calculated the mean of the five correlations.

18. Table 6 shows the cross-country correlations between math achievement and the country mean factor scores CNT_F_INTMATH, CNT_F_SCMATH, CNT_F_MATHEFF, and CNT_F_RESPBIAS (estimated with Model in Figure 2) whereas Table 7 shows the cross-country correlations between math achievement and the original country mean attitude indices.
19. When we compare Table 6 with Table 7, we conclude that the original large negative correlations between interest in math, math self-concept and math performance with math achievement change to small positive correlations. We found the largest change in correlation between interest in math and math achievement indicating that this index seems to be not only at the within country level but also at the between country level the most affected by response bias. Overall, the cross-country correlations between the three attitude indices are all positive and larger in size after taken into account response bias.

Table 6. Cross-country correlations between interest in math, math self-concept, math efficacy, and response bias and math performance for 11 countries in PISA 2003.

| Between country correlations (estimated with Model in Figure 2) | CNT_ PVMATH* | CNT_ F_INTMATH | CNT_ F_SCMATH | CNT_ F_MATHEFF | CNT_ F_RESPBIAS |
|---|---|---|---|---|---|
| CNT_PVMATH | 1 | | | | |
| CNT_F_INTMATH | 0.32 | 1 | | | |
| CNT_F_SCMATH | 0.32 | 0.93 | 1 | | |
| CNT_F_MATHEFF | 0.31 | 0.47 | 0.58 | 1 | |
| CNT_F_RESPBIAS | -0.64 | -0.91 | -0.86 | -0.37 | 1 |

*The correlations between each of the 5 PV's and the constructs were averaged

Table 7. Cross-country correlations between interest in math, math self-concept, math efficacy, and math performance for 11 countries in PISA 2003.

| Between country correlations (original correlations) | CNT_ PVMATH* | CNT_ INTMAT | CNT_ SCMAT | CNT_ MATHEFF |
|---|---|---|---|---|
| CNT_PVMATH | 1 | | | |
| CNT_INTMAT | -0.90 | 1 | | |
| CNT_SCMAT | -0.54 | 0.57 | 1 | |
| CNT_MATHEFF | 0.37 | -0.11 | 0.41 | 1 |

*The correlations between each of the 5 PV's and the constructs were averaged

20. Table 8 and 9 present the mean factor scores for each country with and without controlling for response bias, respectively. The results show that DEU and FRA are among the countries with the most positive mean score on the 'new' math interest index F_INTMATH (see Table 8) whereas these countries—especially JPN and KOR—were among the countries with lower than average mean interest in math on the original math index (see Table 9). IDN, JPN, KOR, and TUN were among the countries with the lowest mean interest in math on the 'new' index F_INTMATH (see Table 8) whereas IDN and TUN were among the countries with the highest mean score on the original index (see Table 9).
21. Overall, the largest shifts in relative country position regarding interest in math seem to appear for for the (East-)Asian countries that are also among the highest performing countries (JPN, HKG, and KOR) and for the lower performing countries (BRA, TUN, and IDN). In Table 9, JPN and KOR were among the countries with the lowest interest in math whereas in Table 8 the opposite is true. In Table 9, BAR, TUN, and IDN were among the countries with the highest interest in math scores whereas they are among the countries with low interest levels in math.

Table 8. Countries mean factor scores (estimated with Model in Figure 2) and pv.

| CNT | F_INTMATH | F_SCMATH | F_MATHEFF | CNT_RESPBIAS | CNT_PVMATH1 |
|---|---|---|---|---|---|
| AUS | 0.01 | 0 | 0.01 | 0 | 524.08 |
| BRA | 0.16 | -0.17 | -0.27 | 0.26 | 355.52 |
| DEU | 0.49 | 0.26 | 0.05 | -0.53 | 503.08 |
| FIN | -0.01 | -0.02 | -0.12 | -0.17 | 544.17 |
| FRA | 0.47 | 0.01 | -0.04 | -0.52 | 511.47 |
| HKG | -0.3 | -0.54 | 0.05 | 0.56 | 549.43 |
| IDN | -0.93 | -0.8 | -0.19 | 1.73 | 360.09 |
| IRL | -0.07 | -0.14 | -0.04 | 0.05 | 503.48 |
| JPN | -0.33 | -0.58 | -0.41 | 0.17 | 533.64 |
| KOR | -0.44 | -0.58 | -0.29 | 0.47 | 541.63 |
| TUN | -0.51 | -0.6 | -0.21 | 1.35 | 358.92 |
| Mean | -0.13 | -0.29 | -0.13 | 0.31 | 480.50 |
| SD | 0.43 | 0.34 | 0.15 | 0.71 | 80.10 |

Note. The results are based on CFA multigroup analyses where AUS is the reference category with a country mean of zero. The other country means are estimated as deviations from the reference group.

Table 9. Countries mean original indices scores and pv.

| CNT | CNT_INTMAT | CNT_SCMAT | CNT_MATHEFF | CNT_PVMATH1 |
|---|---|---|---|---|
| AUS | -0.27 | 0.22 | 0.41 | 524.08 |
| BRA | 0.29 | 0.11 | -0.12 | 355.52 |
| DEU | -0.23 | 0.24 | 0.46 | 503.08 |
| FIN | -0.52 | 0.09 | 0.13 | 544.17 |
| FRA | -0.23 | -0.11 | 0.29 | 511.47 |
| HKG | -0.05 | -0.21 | 0.42 | 549.43 |
| IDN | 0.46 | 0.19 | -0.04 | 360.09 |
| IRL | -0.33 | 0.04 | 0.26 | 503.48 |
| JPN | -0.66 | -0.50 | -0.28 | 533.64 |
| KOR | -0.40 | -0.31 | -0.16 | 541.63 |
| TUN | 0.66 | 0.24 | -0.03 | 358.92 |
| Mean | -0.12 | 0.00 | 0.12 | 480.50 |
| SD | 0.40 | 0.24 | 0.25 | 80.10 |

22. Table 10 shows the correlations between the original attitude indices, the response bias factor score, and the attitude factor scores after taken into account response bias. The correlations among the 'new' attitude factor scores are higher and all positive than those among the original attitude indices suggesting that we indeed successfully controlled for bias (that may have attenuated the correlations among the original attitude indices).

Table 10. Correlations between country mean original indices and country mean factor scores for 11 countries in PISA 2003 (estimated with Model in Figure 2).

| | CNT_INTMAT | CNT_SCMAT | CNT_MATHEFF | CNT_FINTMATH | CNT_FSCMATH | CNT_FMATHEFF | CNT_FRESPBIAS |
|---|---|---|---|---|---|---|---|
| CNT_INTMAT | 1 | | | | | | |
| CNT_SCMAT | 0.57 | 1 | | | | | |
| CNT_MATHEFF | -0.11 | 0.41 | 1 | | | | |
| CNT_FINTMATH | -0.38 | 0.14 | 0.53 | 1 | | | |
| CNT_FSCMATH | -0.38 | 0.38 | 0.64 | 0.93 | 1 | | |
| CNT_FMATHEFF | -0.04 | 0.45 | 0.99 | 0.47 | 0.58 | 1 | |
| CNT_FRESPBIAS | 0.73 | 0.15 | -0.45 | -0.91 | -0.86 | -0.37 | 1 |

## RESPONSE STYLE 'SIMPLE' INDICES

Next to the Mplus analyses, we also calculated a number of simple indices that measure response bias on the PISA 2003 data. Table 11 gives an overview of the indices and clarifies how they were measured.

Table 11. An overview of response bias indices

| *Response Style* | | *Definition* | *Measurement* |
|---|---|---|---|
| ARS0_all_PRP | Acquiescence or agreement response style | The tendency to agree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on all the 75 Likert type items in PISA 2003.<br>Coding:<br>strongly agree: score 1; agree/disagree/strongly disagree: score 0<br>(following definition of Buckley, 2009) |
| ARS1_all_PRP | Acquiescence or agreement response style | The tendency to agree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on all the 75 Likert type items in PISA 2003.<br>Coding:<br>strongly agree: score 2, agree: score 1, disagree/strongly disagree: score 0<br>(following definition of Cheung & Rensvold, 2001) |
| ARS0_5it_PRP | Acquiescence or agreement response style | The tendency to agree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on 5 most heterogeneous Likert type items in PISA 2003.<br>Coding:<br>strongly agree: score 1; agree/disagree/strongly disagree: score 0<br>(following definition of Buckley, 2009) |

| | | | |
|---|---|---|---|
| ARS1_5it_PRP | Acquiescence or agreement response style | The tendency to agree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on 5 most heterogeneous Likert type items in PISA 2003.<br>Coding:<br>strongly agree: score 2, agree: score 1, disagree/strongly disagree: score 0<br>(following definition of Cheung & Rensvold, 2001) |
| ARS_BD_PRP | Acquiescence or agreement response style | The tendency to agree with items regardless of the construct measured. | The acquiescence or agreement response style was measured using balanced data. Balanced data are items belonging to the same scale that have the same meaning but are negative and positively worded. We calculated the proportion of agreement on 5 pairs of balanced items.<br>Coding for each pair of items:<br>Strongly agree – strongly agree: score 3<br>Strongly agree – agree: score 2<br>agree – Strongly agree: score 2<br>agree – agree: score 1<br>other combinations: score 0. |
| DARS0_all_PRP | Disacquiescence of disagreement response style | The tendency to disagree with items regardless of the construct measured. | Proportion of 'strongly disagree' responses on all the 75 Likert type items in PISA 2003.<br>Coding:<br>strongly disagree: score 1;<br>/disagree/agree/strongly disagree: score 0<br>(following definition of Buckley, 2009) |
| DARS1_all_PRP | Disacquiescence of disagreement response style | The tendency to disagree with items regardless of the construct measured. | Proportion of 'strongly disagree' responses on all the 75 Likert type items in PISA 2003.<br>Coding:<br>strongly disagree: score 2, disagree: score 1, agree/strongly agree: |

| | | | |
|---|---|---|---|
| | | | score 0 (following definition of Cheung & Rensvold, 2001) |
| DARS0_5it_PRP | Disacquiescence of disagreement response style | The tendency to disagree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on 5 most heterogeneous Likert type items in PISA 2003. Coding: strongly disagree: score 1; /disagree/agree/strongly disagree: score 0 (following definition of Buckley, 2009) |
| DARS1_5it_PRP | Disacquiescence of disagreement response style | The tendency to disagree with items regardless of the construct measured. | Proportion of 'strongly agree' responses on 5 most heterogeneous Likert type items in PISA 2003. Coding: strongly disagree: score 2, disagree: score 1, agree/strongly agree: score 0 (following definition of Cheung & Rensvold, 2001) |
| ERS0_all_PRP | Extreme response style | The tendency to endorse the most extreme responses regardless of the construct measured. | Proportion of extreme responses 'strongly agree' or 'strongly disagree' on all the 75 Likert type items in PISA 2003. Coding: strongly agree/strongly disagree: score 1; disagree/agree: score 0 |
| ERS0_5it_PRP | Extreme response style | The tendency to endorse the most extreme responses regardless of the construct measured. | Proportion of extreme responses 'strongly agree' or 'strongly disagree' on 5 most heterogeneous Likert type items in PISA 2003. Coding: strongly agree/strongly disagree: score 1; disagree/agree: score 0 |
| NCR_PRP | Non contingent response style. | The tendency to respond to items carelessly, randomly, or non-purposefully | Average absolute difference between responses of 5 pairs of items, where the items in each pair are maximally correlated. |

23. In order to reduce the number of missing values on the response style indices, we calculated a response style index whenever less than 50% of the responses to the items were missing. When more than 50% of the responses to the items were missing, a response style index was not calculated.

24. For the measurement of ARS0_all_PRP, ARS1_all_PRP, DARS0_all_PRP, DARS1_all_PRP, and ERS0_all_PRP, we used 75 Likert items. These are all the items belonging to the attitude indices in PISA 2003 (ATSCHL, STUREL, BELONG, INTMAT, INSTMOT, MATHEFF, ANXMAT, SCMAT, CSTRAT, ELAB, MEMOR, COMPLRN, COOPLRN, TEACHSUP, DISCLIM) except for the ICT-indices as the ICT questionnaire was a national option (and not all countries filled out this questionnaire). All items are Likert type bipolar items with four categories 'Strongly Agree', 'Agree', 'Disagree', and 'Strongly Disagree'. There are, however, some exceptions. The items belonging to the scale MATHEFF and DISCLIMA are also bipolar with four categories but they used different wordings 'Very confident', 'Confident', 'Not very confident', 'Not at all confident' and 'Every lesson', 'Most lessons', 'Some lessons', 'Never or hardly any lesson', respectively. We used the raw student scores on each of the 75 Likert items to calculate the response style indices. This means that although some items were positively whereas others were negatively orientated, we simply calculated the proportion of agreement with, disagreement with, or extreme responses over all the 75 Likert items.

25. For the measurement of ARS0_5it_PRP, ARS1_5it_PRP, DARS0_5it_PRP, DARS1_5it_PRP, and ERS0_5it_PRP, we used a subset of five Likert type items (out of the 75 Likert items) that are heterogeneous[1]. These items are ST24Q01, ST26Q02, ST32Q02, ST37Q06, and ST37Q10. In a first step, we identified the attitude constructs that showed the lowest intercorrelations (ATSCHL, SCMATH, COOPLRN, STUREL, and COMPLRN) and, in a second step, we selected items that showed the lowest intercorrelations among each other. The average correlation between the 5 items was 0.32 with a minimum of 0.25 and a maximum of 0.62. Regardless the orientation of the items (ST24Q01 and ST32Q02 were positively orientated meaning a higher item score indicates a more positive attitude whereas the other items were negatively orientated), we used the raw scores on each of the item questions to calculate the proportion agreement, disagreement, or extreme responses.

26. For the measurement of ARS_BD_PRP, we selected five pairs of items belonging to the same scale that were positively and negatively worded. These pairs are (ST24Q01, ST24Q03), (ST24Q02, ST24Q04), (ST27Q04, ST27Q03), (ST27Q02, ST27Q06) and (ST32Q02, ST32Q06).

27. For the measurement of NCR_PRP, we selected five pairs of items that were highly correlated. These are (ST30Q03, ST30Q04), (ST30Q02, ST30Q05), (ST31Q05, ST31Q07), (ST37Q05, ST37Q07), and (ST38Q06, ST38Q08). The average correlation was 0.66 ranging from 0.603 to 0.744.

28. In the literature, we found two other response styles that may be interesting. MPR or the tendency to respond towards the midpoint of the scale (this is actually equivalent to 1-ERS in PISA 2003) and NARS or the tendency to show greater acquiescence than disacquiescence (=ARS-DARS). We didn't include them because they are linear combinations from ERS, ARS, and DARS. (Note: another response style that was mentioned by Baumgartner and Steenkamp (2001) is RR or the tendency to use a narrow or wide range of response categories around the mean response and is usually measured by the standard deviation of a person's responses across many heterogeneous items. We did not calculate RR.)

[1] In the literature it is recommended to calculate response styles using a heterogeneous set of items.

*Countries mean math performance and mean response styles*

29. Figures 3 to 5 present some examples of correlations between different response styles and math performance for *all* 41 countries in PISA 2003. Table 12 shows the actual cross-country correlations between math performance and all of the response styles for all 41 countries in PISA 2003.
30. Figures 3 and 4 show a positive and negative relationship between ARS and DARS with PV1MATH across countries, respectively. This means that the higher performing countries show fewer tendencies to agree with statements (ARS) but a higher tendency to disagree with statements (DARS) than the lower performing countries.
31. The between-country relationship between ERS and NCR with PV1MATH shown in Figures 5 and 6 are both negative but the correlation is much smaller in size compared to Figures 3 and 4. The results indicate that poor performing countries show a higher tendency to show extreme and non-contingent responses than the higher performing countries. Interestingly, the German speaking countries such as LIE, DEU, CHE, AUT, and LUX show high scores on the extreme response style index ERS0_all_PRP.

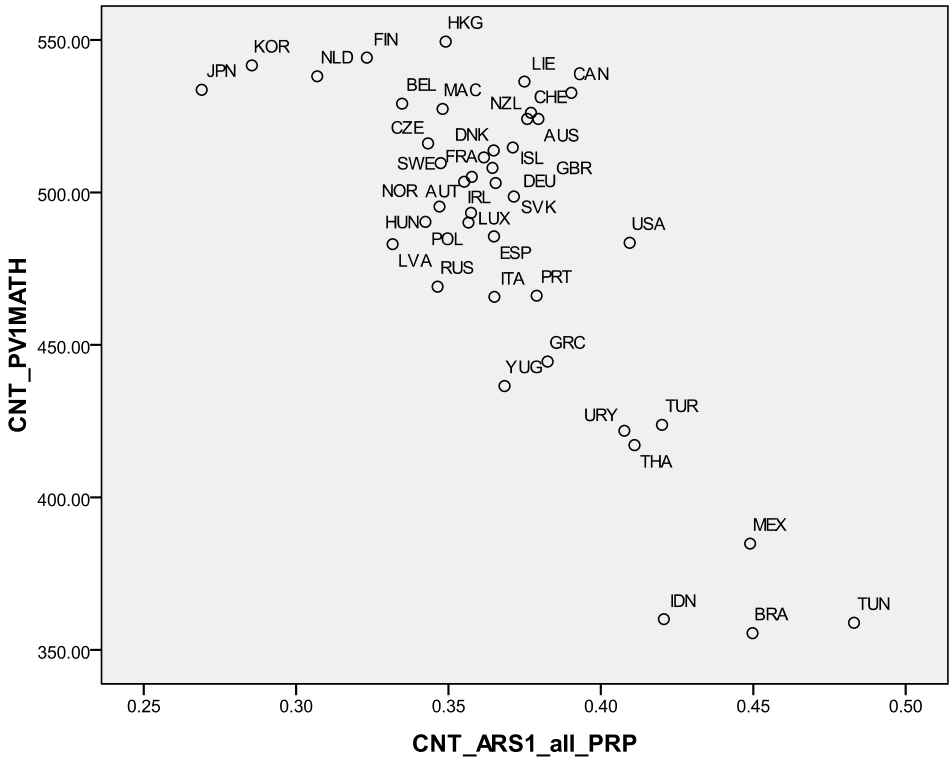


Figure 3. Country mean acquiescence response styles by country mean PV1MATH for all 41 countries in PISA 2003

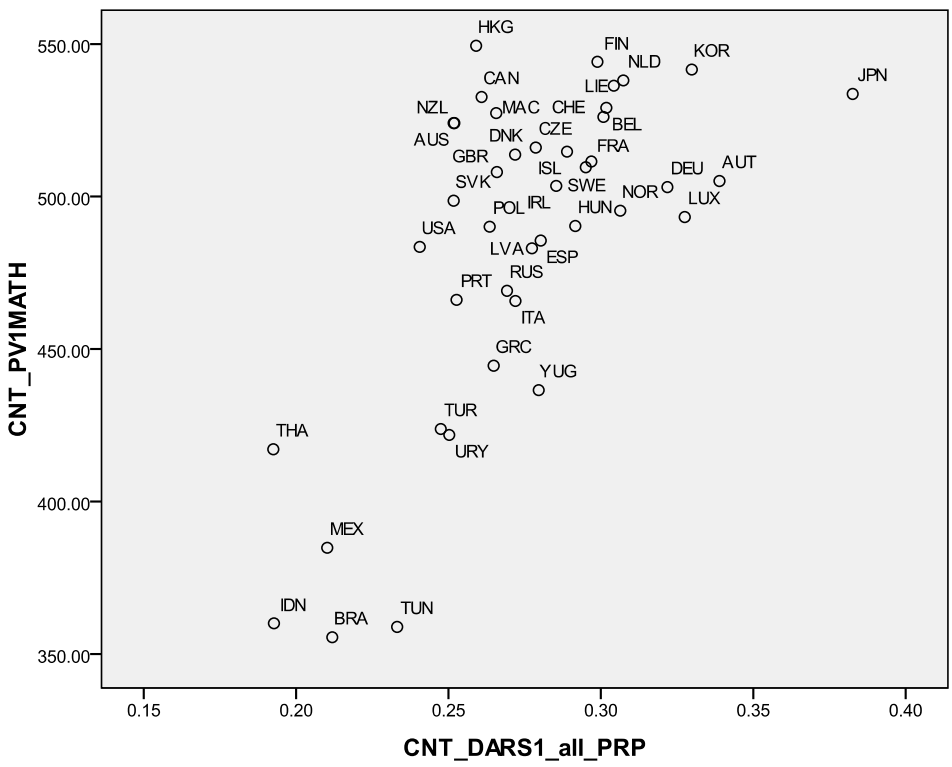


Figure 4. Country mean disacquiescence response style by country mean PV1MATH for all 41 countries in PISA 2003

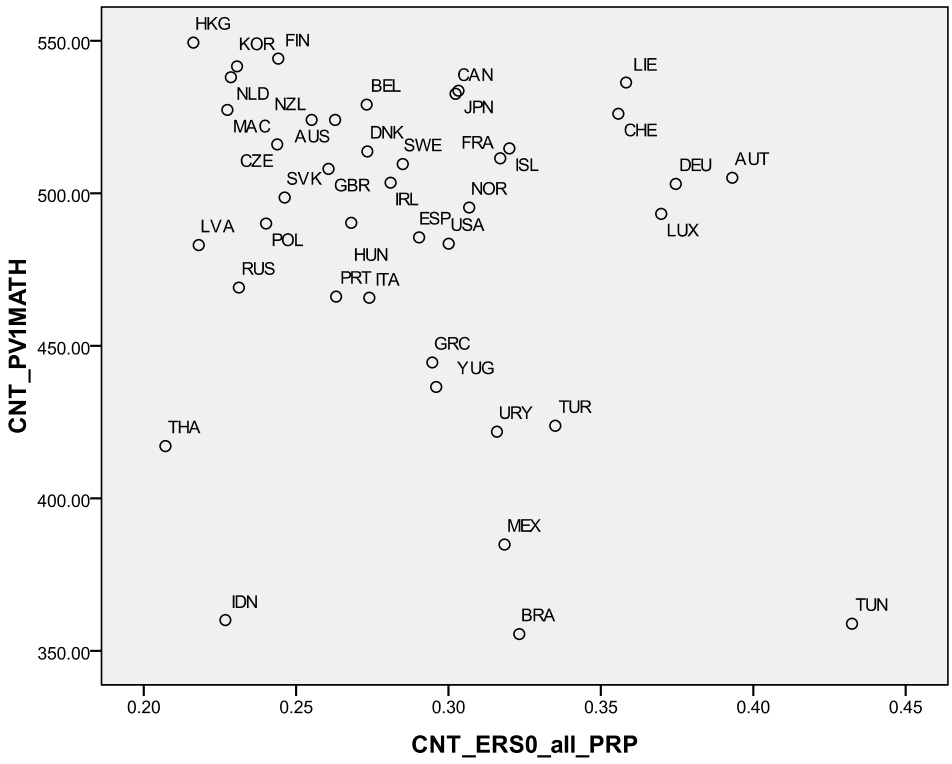


Figure 5. Country mean extreme response style by country mean PV1MATH for all 41 countries in PISA 2003

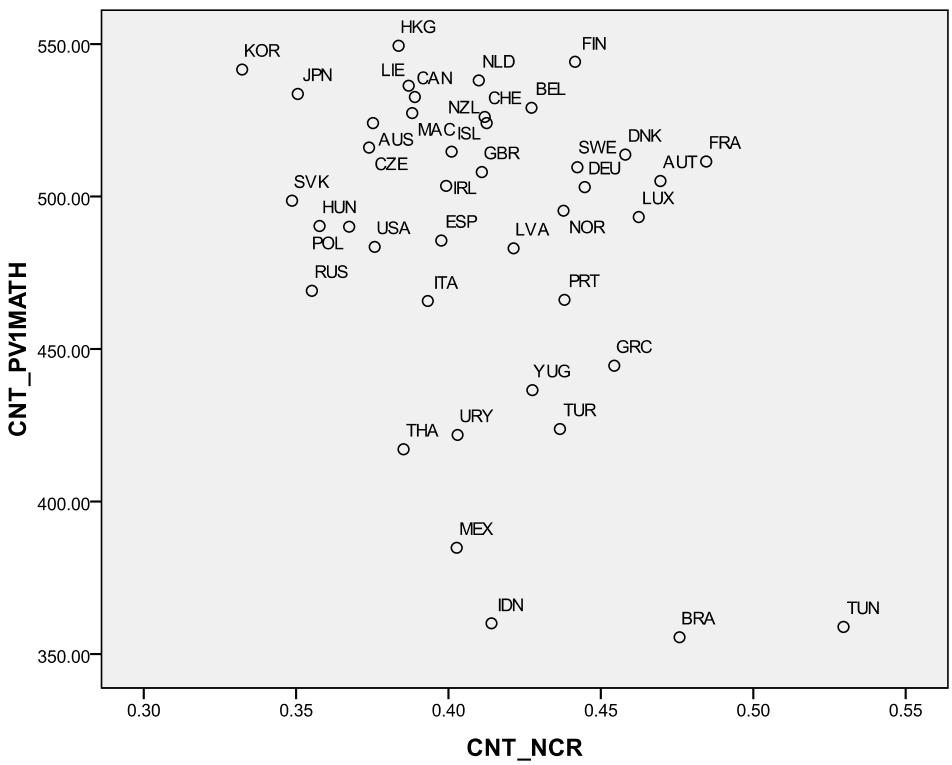


**Figure 6. Country mean non-contingent response style by country mean PV1MATH for all 41 countries in PISA 2003**

32. In Appendix A, Table 15 shows the country mean math performance in descending order and the country mean response style scores.

33. Next to the cross-country correlations between math performance and the response styles for all 41 countries in PISA 2003, Table 12 also shows the correlations between the different response styles. The results show that the different measurements of ARS are highly correlated (correlations ranging from .605 to .837) as well as the different measurements of DARS (correlations ranging from .594 to .875). As can be expected ERS is positively correlated both with ARS and DARS measures. NCR shows the largest correlation with ERS. Note that when comparing ARS0 and ARS1, DARS0 and DARS1 (both for the 75 items and the 5 items), ARS1 and DARS1 always have a stronger correlation with math achievement suggesting that the way of measuring and scoring acquiescence and disacquiescence by Cheung and Rensvold (2001) is more efficient than the definition proposed by Buckley (2009). In addition, the results suggest that using all 75 Likert items or only a subset of 5 heterogeneous items (or balanced data for ARS) to measure ARS, DARS, and ERS doesn't seem to matter too much.

**Table 12. Cross-country correlations between math performance and response styles for all 41 countries in PISA 2003**

| | CNT_PVMATH | CNT_ARS0_all_PRP | CNT_ARS1_all_PRP | CNT_ARS0_5it_PRP | CNT_ARS1_5it_PRP | CNT_DARS0_all_PRP | CNT_DARS1_all_PRP | CNT_DARS0_5it_PRP | CNT_DARS1_5it_PRP | CNT_ERS0_all_PRP | CNT_ERS0_5it_PRP | CNT_ARS_BD_PRP | CNT_NCR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CNT_PVMATH | 1.00 | | | | | | | | | | | | |
| CNT_ARS0_all_PRP | -0.63 | 1.00 | | | | | | | | | | | |
| CNT_ARS1_all_PRP | -0.79 | 0.86 | 1.00 | | | | | | | | | | |
| CNT_ARS0_5it_PRP | -0.69 | 0.87 | 0.77 | 1.00 | | | | | | | | | |
| CNT_ARS1_5it_PRP | -0.72 | 0.62 | 0.79 | 0.84 | 1.00 | | | | | | | | |
| CNT_DARS0_all_PRP | 0.42 | 0.08 | -0.40 | 0.02 | -0.43 | 1.00 | | | | | | | |
| CNT_DARS1_all_PRP | 0.68 | -0.37 | -0.78 | -0.36 | -0.70 | 0.88 | 1.00 | | | | | | |
| CNT_DARS0_5it_PRP | 0.23 | 0.38 | -0.06 | 0.17 | -0.33 | 0.84 | 0.59 | 1.00 | | | | | |
| CNT_DARS1_5it_PRP | 0.57 | -0.17 | -0.54 | -0.42 | -0.84 | 0.77 | 0.80 | 0.78 | 1.00 | | | | |
| CNT_ERS0_all_PRP | -0.23 | 0.80 | 0.41 | 0.67 | 0.22 | 0.66 | 0.24 | 0.79 | 0.33 | 1.00 | | | |
| CNT_ERS0_5it_PRP | -0.31 | 0.82 | 0.48 | 0.77 | 0.34 | 0.56 | 0.14 | 0.76 | 0.23 | 0.95 | 1.00 | | |
| CNT_ARS_BD_PRP | -0.64 | 0.53 | 0.61 | 0.71 | 0.79 | -0.28 | -0.49 | -0.28 | -0.65 | 0.23 | 0.29 | 1.00 | |
| CNT_NCR | -0.36 | 0.57 | 0.43 | 0.52 | 0.30 | 0.25 | -0.06 | 0.47 | 0.07 | 0.58 | 0.65 | 0.05 | 1.00 |

*Country mean attitude and mean response styles*

34. Table 13 shows the correlations between country mean attitude and country mean response styles. We highlighted the attitude indices in yellow when they show a positive within-country but a negative between country correlations with achievement (cf. paradox). Positive high correlations (> .6) are shown in green font and high negative correlations (< -.6) are shown in red font. There seems to be a pattern among the attitude indices that show the paradox. They all show positive correlations with ARS and negative with DARS, except for CNT_SCMATH (We suspect that this index is also affected by other type of bias such as the BFLPE). The indices that do not show the paradox (CNT_MATHEFF, CNT_BELONG, CNT_ANXMAT, and CNT_DISCLIM) do not seem to show a clear pattern in their correlations with response styles. This suggests that they are indeed less susceptible to response bias and that this may explain why they do not show the paradox.
35. In the literature it is often suggested that using balanced scales (i.e., including positively and negatively orientated items measuring the same construct) is an efficient way of reducing acquiescence bias. However, the results do not seem to support this. The only scales that are balanced are ATSCHL (2 items positively and 2 items negatively orientated) and BELONG (3 items positively and 3 items negatively orientated)[2] and at least ATSCHL doesn't not show lower correlations with acquiescence than other non-balanced scales.

Table 13. Cross-country correlations between math performance and attitude indices for all 41 countries in PISA 2003

| | CNT_ INTMAT | CNT_ SCMAT | CNT_ MATHEFF | CNT_ ATSCHL | CNT_ STUREL | CNT_ BELONG | CNT_ INSTMOT |
|---|---|---|---|---|---|---|---|
| CNT_ARS0_all_PRP | .573 | .525 | .069 | .595 | .483 | .389 | .483 |
| CNT_ARS1_all_PRP | .808 | .574 | -.053 | .774 | .716 | .230 | .784 |
| CNT_ARS0_5it_PRP | .622 | .225 | -.275 | .460 | .466 | .158 | .450 |
| CNT_ARS1_5it_PRP | .786 | .151 | -.387 | .473 | .621 | -.069 | .644 |
| CNT_DARS0_all_PRP | -.598 | -.115 | .228 | -.392 | -.498 | .434 | -.635 |
| CNT_DARS1_all_PRP | -.808 | -.387 | .185 | -.676 | -.715 | .135 | -.844 |
| CNT_DARS0_5it_PRP | -.329 | .296 | .302 | -.014 | -.183 | .585 | -.276 |
| CNT_DARS1_5it_PRP | -.698 | .041 | .408 | -.318 | -.532 | .352 | -.597 |
| CNT_ERS0_all_PRP | .077 | .329 | .188 | .217 | .069 | .551 | -.014 |
| CNT_ERS0_5it_PRP | .201 | .340 | .012 | .296 | .191 | .482 | .121 |
| CNT_ARS_BD_PRP | .644 | .078 | -.292 | .265 | .410 | -.223 | .373 |
| CNT_NCR | .294 | .332 | -.176 | .453 | .269 | .382 | .235 |

| | CNT_ ANXMAT | CNT_ CSTRAT | CNT_ ELAB | CNT_ MEMOR | CNT_ COMPLRN | CNT_ COOPLRN | CNT_ TEACHSUP | CNT_ DISCLIM |
|---|---|---|---|---|---|---|---|---|
| CNT_ARS0_all_PRP | .243 | .812 | .594 | .556 | .603 | .679 | .378 | -.271 |
| CNT_ARS1_all_PRP | .354 | .768 | .865 | .821 | .801 | .838 | .705 | -.361 |
| CNT_ARS0_5it_PRP | .495 | .652 | .569 | .498 | .632 | .563 | .410 | -.278 |
| CNT_ARS1_5it_PRP | .588 | .525 | .728 | .645 | .722 | .647 | .660 | -.312 |
| CNT_DARS0_all_PRP | -.376 | -.078 | -.693 | -.577 | -.515 | -.420 | -.685 | .162 |
| CNT_DARS1_all_PRP | -.405 | -.430 | -.896 | -.822 | -.751 | -.713 | -.834 | .311 |
| CNT_DARS0_5it_PRP | -.446 | .116 | -.382 | -.316 | -.206 | -.149 | -.381 | .032 |
| CNT_DARS1_5it_PRP | -.586 | -.250 | -.678 | -.595 | -.582 | -.510 | -.668 | .229 |
| CNT_ERS0_all_PRP | -.040 | .569 | .037 | .077 | .150 | .263 | -.122 | -.109 |
| CNT_ERS0_5it_PRP | .040 | .507 | .131 | .126 | .286 | .277 | .027 | -.163 |
| CNT_ARS_BD_PRP | .542 | .452 | .585 | .499 | .568 | .405 | .480 | -.028 |
| CNT_NCR | -.008 | .449 | .252 | .180 | .214 | .425 | .065 | -.466 |

[2] SCMAT cannot be considered as a balanced scale as it only includes 1 item that is positively and 4 items that are negatively orientated.

### COMPARING THE RESPONSE STYLE FACTOR ESTIMATED WITH MODEL IN FIGURE 2 WITH THE 'SIMPLE' MEASURES OF RESPONSE STYLES ACROSS 11 COUNTRIES IN PISA 2003

36. In order to give an answer to the question what the response style factor that was extracted using the Model shown in Figure 2 using Mplus actually measures, we calculated the correlations between this factor and the simple response style measures for 11 countries in PISA 2003. Table 14 presents the results. These are only preliminary results because the correlations are based on only 11 countries.

37. First, we found positive correlations between the response style factor and the 'simple' acquiescence (CNT_ARS_BD_PRP showed the highest correlation) and a negative correlation of similar size between the response style factor and the 'simple' disacquiescence (CNT_DARS1_5it_PRP showed the highest negative correlation) response style measure. These results indicate that the response style factor that was extracted using the Model shown in Figure 2 represent an acquiescence response style.
38. A second interesting finding also seems to support the conclusion that the response style factor represents an acquiescence response style. When we look at the correlations between the simple ARS and DARS response style measures and CNT_FINTMATH, CNT_FSCMATH and compare these with the correlation between the simple ARS and DARS with the original country mean attitudes CNT_INTMAT and CNT_SCMATH, we notice that the correlations between CNT_FINTMATH, CNT_FSCMATH and ARS are near zero whereas they are high and positively correlated with the original attitude country mean constructs. This may suggest that ARS has been controlled for when estimating FINTMATH, FSCMATH. However, there correlations between DARS with CNT_FINTMATH, CNT_FSCMATH are relatively high and positive which may indicate that there still remains some disacquiescence response bias.

Table 14. Cross-country correlations between attitude indices, ‘simple’ response style measures, and response style factor for 11 countries in PISA 2003

| | CNT_PVMATH | CNT_INTMAT | CNT_SCMAT | CNT_MATHEFF | CNT_F_INTMATH | CNT_F_SCMATH | CNT_F_MATHEFF | CNT_F_RESPBIAS |
|---|---|---|---|---|---|---|---|---|
| CNT_PV1MATH | 1 | | | | | | | |
| CNT_INTMAT | -0.90 | 1 | | | | | | |
| CNT_SCMAT | -0.54 | 0.57 | 1 | | | | | |
| CNT_MATHEFF | 0.37 | -0.11 | 0.41 | 1 | | | | |
| CNT_F_INTMATH | 0.32 | -0.38 | 0.14 | 0.53 | 1 | | | |
| CNT_F_SCMATH | 0.32 | -0.38 | 0.38 | 0.64 | 0.93 | 1 | | |
| CNT_F_MATHEFF | 0.31 | -0.04 | 0.45 | 0.99 | 0.47 | 0.58 | 1 | |
| CNT_F_RESPBIAS | -0.64 | 0.73 | 0.15 | -0.45 | -0.91 | -0.86 | -0.37 | 1 |
| CNT_ARS0_all_PRP | -0.78 | 0.80 | 0.65 | 0.01 | 0.08 | 0.06 | 0.02 | 0.29 |
| CNT_ARS1_all_PRP | -0.87 | 0.92 | 0.77 | 0.07 | -0.08 | -0.04 | 0.12 | 0.47 |
| CNT_ARS0_5it_PRP | -0.76 | 0.74 | 0.33 | -0.25 | -0.06 | -0.17 | -0.26 | 0.38 |
| CNT_ARS1_5it_PRP | -0.81 | 0.87 | 0.37 | -0.15 | -0.25 | -0.34 | -0.12 | 0.58 |
| CNT_DARS0_all_PRP | 0.46 | -0.60 | -0.40 | -0.06 | 0.44 | 0.33 | -0.16 | -0.58 |
| CNT_DARS1_all_PRP | 0.75 | -0.85 | -0.68 | -0.10 | 0.29 | 0.20 | -0.18 | -0.59 |
| CNT_DARS0_5it_PRP | 0.04 | -0.12 | 0.17 | 0.06 | 0.50 | 0.48 | -0.02 | -0.42 |
| CNT_DARS1_5it_PRP | 0.56 | -0.68 | -0.19 | 0.07 | 0.45 | 0.49 | 0.00 | -0.64 |
| CNT_ERS0_all_PRP | -0.44 | 0.37 | 0.35 | -0.03 | 0.33 | 0.25 | -0.07 | -0.07 |
| CNT_ERS0_5it_PRP | -0.51 | 0.45 | 0.32 | -0.14 | 0.22 | 0.13 | -0.19 | 0.04 |
| CNT_ARS_BD_PRP | -0.65 | 0.70 | 0.10 | -0.32 | -0.48 | -0.58 | -0.31 | 0.68 |
| CNT_NCR | -0.61 | 0.62 | 0.60 | 0.11 | 0.30 | 0.26 | 0.13 | 0.06 |

## REFERENCES


Buckley, J. (2008). *Cross-national response styles in international educational assessments: Evidence from PISA 2006*. Last accessed 11/02/2010 https://edsurveys.rti.org/PISA/documents/Buckley_PISAresponsestyle.pdf

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*, 187-212.

Clarke, I. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior and Personality, 15*, 137-152.

Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly, 56*, 328-351.

Johnson, T., Kulesa, P., Young, I. C., & Shavitt, S. (2005). The relation between culture and response styles. Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 369*(2), 264-277.

Lie, S., & Turmo, A. (2005). *Cross-country comparability of students' self-reports*. Report to the PISA Technical Advisory Group.

van Hemert, D.A., Poortinga, Y.H. & van de Vijver, F.J.R. (2007). Emotion and culture: A meta-analysis. *Cognition and Emotion, 21*(5), 913-943.

van de Vijver, F. J. R. & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment, 13*(1), 29-37.

## APPENDIX A TABLE 15. COUNTRY MEAN MATH PERFORMANCE AND MEAN RESPONSE STYLES IN PISA 2003

| CNT | CNT_PVMATH* | CNT_ARS0_all_PRP | CNT_ARS1_all_PRP | CNT_ARS0_5it_PRP | CNT_ARS1_5it_PRP | CNT_DARS0_all_PRP | CNT_DARS1_all_PRP | CNT_DARS0_5it_PRP | CNT_DARS1_5it_PRP | CNT_ERS0_all_PRP | CNT_ERS0_5it_PRP | CNT_ARS_BD_PRP | CNT_NCR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HKG | 550.38 | .13 | .35 | .10 | .34 | .08 | .26 | .06 | .24 | .22 | .16 | .05 | .38 |
| FIN | 544.29 | .12 | .32 | .06 | .24 | .12 | .30 | .12 | .35 | .24 | .18 | .03 | .44 |
| KOR | 542.23 | .10 | .29 | .07 | .25 | .13 | .33 | .11 | .34 | .23 | .18 | .03 | .33 |
| NLD | 537.82 | .11 | .31 | .06 | .24 | .12 | .31 | .10 | .34 | .23 | .16 | .03 | .41 |
| LIE | 535.80 | .20 | .37 | .11 | .27 | .16 | .30 | .15 | .35 | .36 | .25 | .05 | .39 |
| JPN | 534.14 | .11 | .27 | .12 | .28 | .19 | .38 | .13 | .34 | .30 | .25 | .05 | .35 |
| CAN | 532.49 | .19 | .39 | .11 | .29 | .11 | .26 | .13 | .32 | .30 | .24 | .03 | .39 |
| BEL | 529.29 | .14 | .33 | .09 | .27 | .13 | .30 | .11 | .33 | .27 | .20 | .03 | .43 |
| MAC | 527.27 | .15 | .35 | .11 | .32 | .08 | .27 | .07 | .27 | .23 | .17 | .05 | .39 |
| CHE | 526.55 | .20 | .38 | .13 | .29 | .15 | .30 | .16 | .35 | .36 | .29 | .04 | .41 |
| AUS | 524.27 | .16 | .38 | .10 | .29 | .10 | .25 | .11 | .31 | .26 | .21 | .03 | .38 |
| NZL | 523.49 | .16 | .38 | .09 | .30 | .10 | .25 | .10 | .30 | .26 | .20 | .03 | .41 |
| CZE | 516.46 | .14 | .34 | .07 | .25 | .10 | .28 | .09 | .33 | .24 | .16 | .03 | .37 |
| ISL | 515.11 | .18 | .37 | .11 | .27 | .14 | .29 | .16 | .36 | .32 | .27 | .03 | .40 |
| DNK | 514.29 | .16 | .36 | .12 | .33 | .12 | .27 | .12 | .29 | .27 | .24 | .03 | .46 |
| FRA | 510.80 | .17 | .36 | .11 | .28 | .14 | .30 | .14 | .35 | .32 | .25 | .03 | .48 |
| SWE | 509.05 | .15 | .35 | .09 | .26 | .13 | .30 | .15 | .36 | .28 | .24 | .03 | .44 |
| GBR | 508.26 | .15 | .36 | .10 | .30 | .11 | .27 | .11 | .30 | .26 | .20 | .03 | .41 |
| AUT | 505.61 | .20 | .36 | .11 | .26 | .19 | .34 | .19 | .39 | .39 | .31 | .03 | .47 |
| DEU | 502.99 | .20 | .37 | .13 | .28 | .17 | .32 | .16 | .36 | .37 | .29 | .04 | .44 |
| IRL | 502.84 | .16 | .36 | .09 | .28 | .13 | .29 | .11 | .32 | .28 | .20 | .03 | .40 |
| SVK | 498.18 | .16 | .37 | .08 | .27 | .09 | .25 | .07 | .31 | .25 | .15 | .03 | .35 |
| NOR | 495.19 | .16 | .35 | .12 | .30 | .15 | .31 | .13 | .32 | .31 | .25 | .03 | .44 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LUX | 493.21 | .19 | .36 | .13 | .29 | .18 | .33 | .17 | .36 | .37 | .30 | .05 | .46 |
| POL | 490.24 | .14 | .36 | .09 | .28 | .10 | .26 | .10 | .31 | .24 | .19 | .05 | .37 |
| HUN | 490.01 | .15 | .34 | .09 | .28 | .12 | .29 | .09 | .31 | .27 | .17 | .04 | .36 |
| ESP | 485.11 | .16 | .36 | .13 | .32 | .13 | .28 | .11 | .29 | .29 | .24 | .04 | .40 |
| LVA | 483.37 | .12 | .33 | .07 | .26 | .10 | .28 | .10 | .33 | .22 | .18 | .03 | .42 |
| USA | 482.88 | .20 | .41 | .13 | .33 | .10 | .24 | .12 | .30 | .30 | .25 | .06 | .38 |
| RUS | 468.41 | .14 | .35 | .07 | .24 | .09 | .27 | .10 | .35 | .23 | .17 | .04 | .36 |
| PRT | 466.02 | .16 | .38 | .12 | .34 | .11 | .25 | .09 | .26 | .26 | .21 | .04 | .44 |
| ITA | 465.66 | .16 | .37 | .11 | .32 | .12 | .27 | .08 | .28 | .27 | .20 | .04 | .39 |
| GRC | 444.91 | .18 | .38 | .13 | .33 | .11 | .26 | .10 | .28 | .29 | .23 | .05 | .45 |
| YUG | 436.87 | .18 | .37 | .11 | .28 | .12 | .28 | .12 | .33 | .30 | .22 | .05 | .43 |
| TUR | 423.42 | .22 | .42 | .17 | .39 | .11 | .25 | .10 | .25 | .34 | .28 | .09 | .44 |
| URY | 422.20 | .21 | .41 | .13 | .32 | .11 | .25 | .11 | .30 | .32 | .24 | .05 | .40 |
| THA | 416.98 | .15 | .41 | .12 | .40 | .05 | .19 | .03 | .18 | .21 | .15 | .08 | .39 |
| MEX | 385.22 | .23 | .45 | .19 | .42 | .08 | .21 | .09 | .22 | .32 | .28 | .08 | .40 |
| IDN | 360.16 | .17 | .42 | .11 | .33 | .06 | .19 | .06 | .25 | .23 | .17 | .05 | .41 |
| TUN | 358.73 | .31 | .48 | .22 | .39 | .12 | .23 | .16 | .30 | .43 | .39 | .06 | .53 |
| BRA | 356.02 | .24 | .45 | .18 | .39 | .08 | .21 | .09 | .25 | .32 | .27 | .05 | .48 |

*The countries are presented in descending order by math performance

## APPENDIX B. ITEM COMPOSITION OF THE SCALES INTMAT, SCMAT, AND MATHEFF IN PISA 2003

Interest in mathematics (INTMAT)

Thinking about your views on mathematics: To what extent do you agree with the following statements?

(Please <tick> only one box in each row.)

| | | Strongly agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|---|
| ST30Q01 | a) I enjoy reading about mathematics | 1 | 2 | 3 | 4 |
| ST30Q03 | c) I look forward to my mathematics lessons | 1 | 2 | 3 | 4 |
| ST30Q04 | d) I do mathematics because I enjoy it | 1 | 2 | 3 | 4 |
| ST30Q06 | f) I am interested in the things I learn in mathematics | 1 | 2 | 3 | 4 |

Note: All items were inverted for scaling.

Self-concept in Mathematics (SCMAT)

Thinking about your views on mathematics: To what extent do you agree with the following statements?

(Please <tick> only one box in each row.)

| | | Strongly agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|---|
| ST32Q02 | b) I am just not good at mathematics. | 1 | 2 | 3 | 4 |
| ST32Q04 | d) I get good <marks> in mathematics. | 1 | 2 | 3 | 4 |
| ST32Q06 | f) I learn mathematics quickly. | 1 | 2 | 3 | 4 |
| ST32Q07 | g) I have always believed that mathematics is one of my best subjects. | 1 | 2 | 3 | 4 |
| ST32Q09 | i) In my mathematics class, I understand even the most difficult work. | 1 | 2 | 3 | 4 |

Note: Items ST32Q04, ST32Q06, ST32Q07, and ST32Q09 were inverted for scaling.

Mathematics self-efficacy (MATHEFF)

How confident do you feel about having to do the following calculations?

(Please <tick> only one box in each row.)

| | | Very confident | Confident | Not very confident | Not at all confident |
|---|---|---|---|---|---|
| ST31Q01 | a) Using a <train timetable>, how long it would take to get from Zedville to Zedtown | 1 | 2 | 3 | 4 |
| ST31Q02 | b) Calculating how much cheaper a TV would be after a 30 percent discount | 1 | 2 | 3 | 4 |
| ST31Q03 | c) Calculating how many square metres of tiles you need to cover a floor | 1 | 2 | 3 | 4 |
| ST31Q04 | d) Understanding graphs presented in newspapers | 1 | 2 | 3 | 4 |
| ST31Q05 | e) Solving an equation like $3x + 5 = 17$ | 1 | 2 | 3 | 4 |
| ST31Q06 | f) Finding the actual distance between two places on a map with a 1:10,000 scale | 1 | 2 | 3 | 4 |
| ST31Q07 | g) Solving an equation like $2(x+3) = (x + 3)(x - 3)$ | 1 | 2 | 3 | 4 |
| ST31Q08 | h) Calculating the petrol consumption rate of a car | 1 | 2 | 3 | 4 |

Note: All items were inverted for scaling.