

Response Styles in Cross-Cultural Perspective: TIMSS and PIRLS

Fons J. R. van de Vijver

Jamis He

Alena Kulikova



IEA

Researching education, improving learning

**7th IEA International
Research Conference**

Theme

- Focus on non-cognitive part of LSA
 - Self-reports have known limitations
- Response Styles in TIMSS and PIRLS
 - Prevalence
 - Meaning

Structure

1. What are Response Styles?
 - Views: Nuisance or style?
 - One or many?
2. Response Styles in TIMSS and PIRLS
 - Focus on country-level analyses
3. Conclusion: Should we bother?

What is Novel?

- Response Styles have not been studied systematically in TIMSS and PIRLS
- Response Styles infrequently studied in Grade 4 and 8

Response Styles

- Systematic tendency to endorse certain response options on some basis other than the target construct (Paulhus, 1991)
- Study
- Main
 - Requested Response Style (RRS)
 - Midpoint response style (MRS)
 - Extremity Response Style (ERS)
 - Socially Desirable Responding (SDR)

MRS: Midpoint Response Style
ERS: Extremity Response Style

Views on Response Styles

- 1. Response Styles as *nuisance*, styles represent non-target variance and should be minimized or eliminated
 - Taken view in psychology and survey research
 - Example: Eysenck: social desirability measured by Lie Scale (derived from Marlowe-Crowne Social Desirability Scale): scores on personality scales only to be interpreted if Lie Scale score below a certain criterion score

- 2. Response Style as part of personality (Costa & McCrae, 1983) or communication style (Smith, 2004)
 - Response styles related to agreeableness (fitting in)
 - Systematic cross-cultural differences
 - Low affluence/high collectivism/low HDI (Human Development Index) higher scores on response styles

Note on Operationalization

- ARS, ERS and MRS typically derived from
 - Items using a single response format (e.g., four-point agreement)
 - Items from heterogeneous domains
 - Makes style scores less dependent on substantive constructs measured
 - ARS, ERS and MRS derived from different items

- Example:
 - Likert 5-point scale
 - Scores 1 and 5 become 1 in ERS and 0 in ARS and MRS
 - Score 3 becomes 1 in MRS and 0 in ARS and ERS
 - Score 4 becomes 1 in ARS and 0 in MRS and ERS
 - Score 2 becomes 0
- Score vector: 2 4 5 3 1

Would be rescored as

- MRS: 0 0 0 1 0
- ARS: 0 1 0 0 0
- ERS: 0 0 1 0 1

Assessment and Statistical Correction: Integrated Response Style



Contents lists available at [SciVerse ScienceDirect](#)

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

A general response style factor: Evidence from a multi-ethnic study
in the Netherlands

Jia He^{a,*}, Fons J.R. van de Vijver^{a,b,c}

^a Tilburg University, The Netherlands

^b North-West University, South Africa

^c University of Queensland, Australia

Research questions:

1. Is there a General Response Style?
2. Cross-cultural differences in means and associations with personality

LISS Immigrant Panel Samples

➤ **Sample and procedure:** LISS Immigrant Panel <http://www.lissdata.nl/lissdata/>

Ethnic Group	Sample Size	Age (M)	% Male	Education
Dutch Mainstreamers	547	47.58	51%	3.78
2nd Generation Western Immigrants	253	51.84	42%	4.13
1st Generation Western Immigrants	346	43.74	47%	3.61
2nd Generation Non-Western Immigrants	345	47.20	46%	3.58
1st Generation Non-Western Immigrants	172	31.42	40%	3.56

Note. 1. Education was scored from 1 (primary school) to 6 (university)

Instruments

➤ Two Measures of Response Styles

• Indirect measures of acquiescence, extremity, and midpoint responding

(extracted from various scales in previous waves of data)

e.g., Acquiescence as the frequency of choosing 4 (agree) in a 5 point scale

• Direct measures

Acquiescence

Extremity

Midpoint responding

Social desirability

Newly developed scales

e.g., Do you prefer strong opinions to no opinion at all?

Adapted items from Marlowe-Crowne Scale and Balanced Inventory of Desirable Responding

e.g., I am jealous of others with good fortune.

➤ Personality: Big Five, using IPIP items

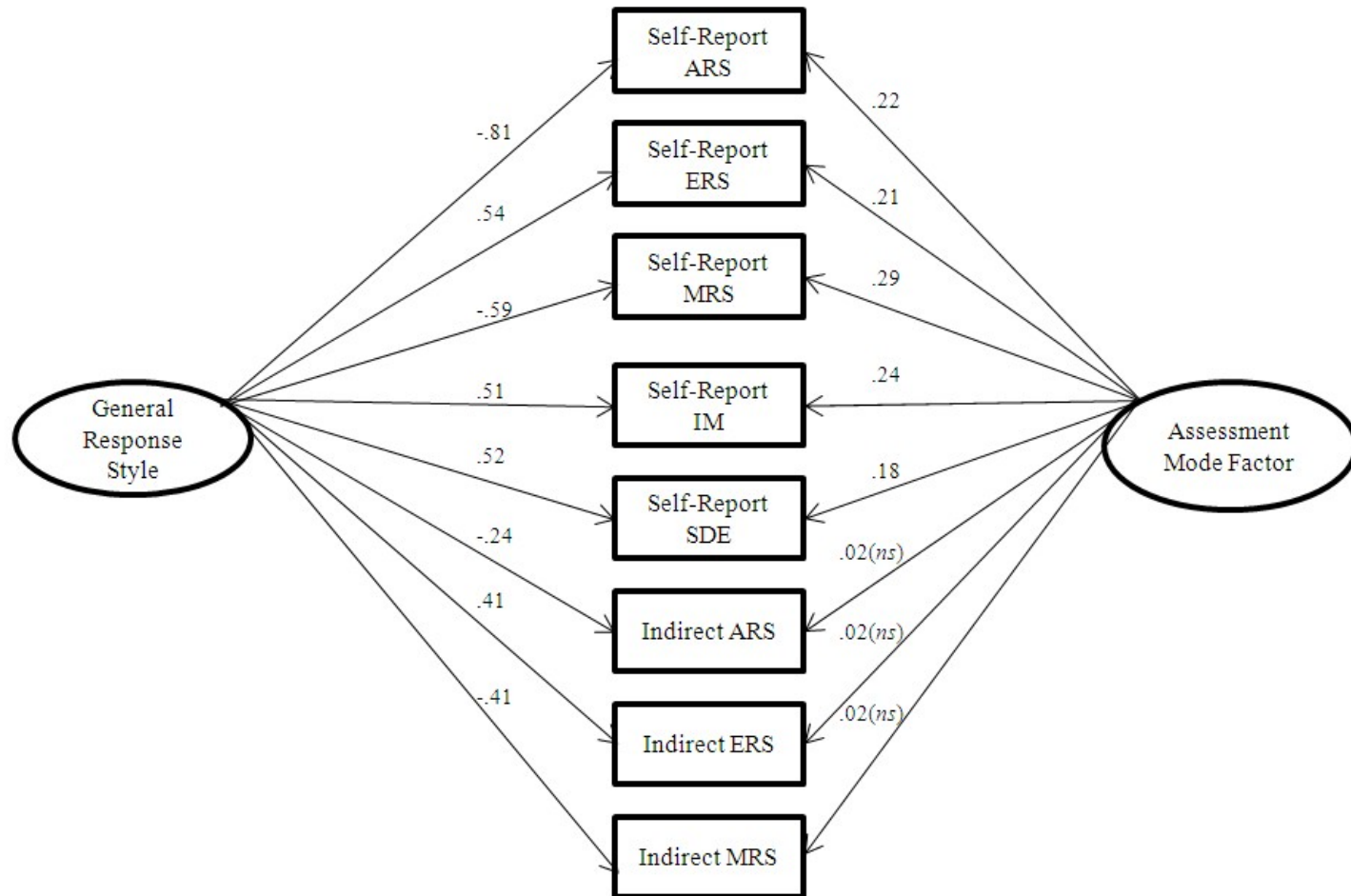
Big One factor: factor score of the one-factor solution of the Big Five

(4.11% variance explained)

Measures of Response Styles

- ARS scores
 - *From Self-Esteem and Positive Attitude towards Survey*,
 - 16 items
 - Both with half positively and half negatively worded items
 - 7-point *disagree* to *agree* response options.
 - Operationalized as the proportion of the responses of 5 (*somewhat agree*) and 6 (*agree*)
 - Note: Responses of 7 (*strongly agree*) were excluded from the ARS computation due to the fact that it may also be triggered by ERS.
- Similar procedure for MRS and ERS
- Social Desirability:
 - 17 items from the Marlowe-Crowne Social Desirability Scale and the Balanced Inventory of Desirable Responding
 - 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*)
 - 2-factor solution: impression management and self-deception

General Response Style Factor



Multigroup CFA: Structural covariances model,
 $\chi^2/df = 2.82$, CFI = .91, TLI = .92, RMSEA = .03

Cross-Group Similarities and Differences

Latent Means in the Multigroup CFA model:

Ethnic Group	General Response	
	Style	Mode factor
Dutch Mainstreamers	0 (fixed)	0 (fixed)
2nd Generation Western Immigrants	.01 (.04)	.02 (.11)
1st Generation Western Immigrants	-.05 (.04)	.06 (.13)
2nd Generation Non-Western Immigrants	-.16 (.05)**	.33 (.15)*
1st Generation Non-Western Immigrants	-.10 (.04)**	.56 (.12)**

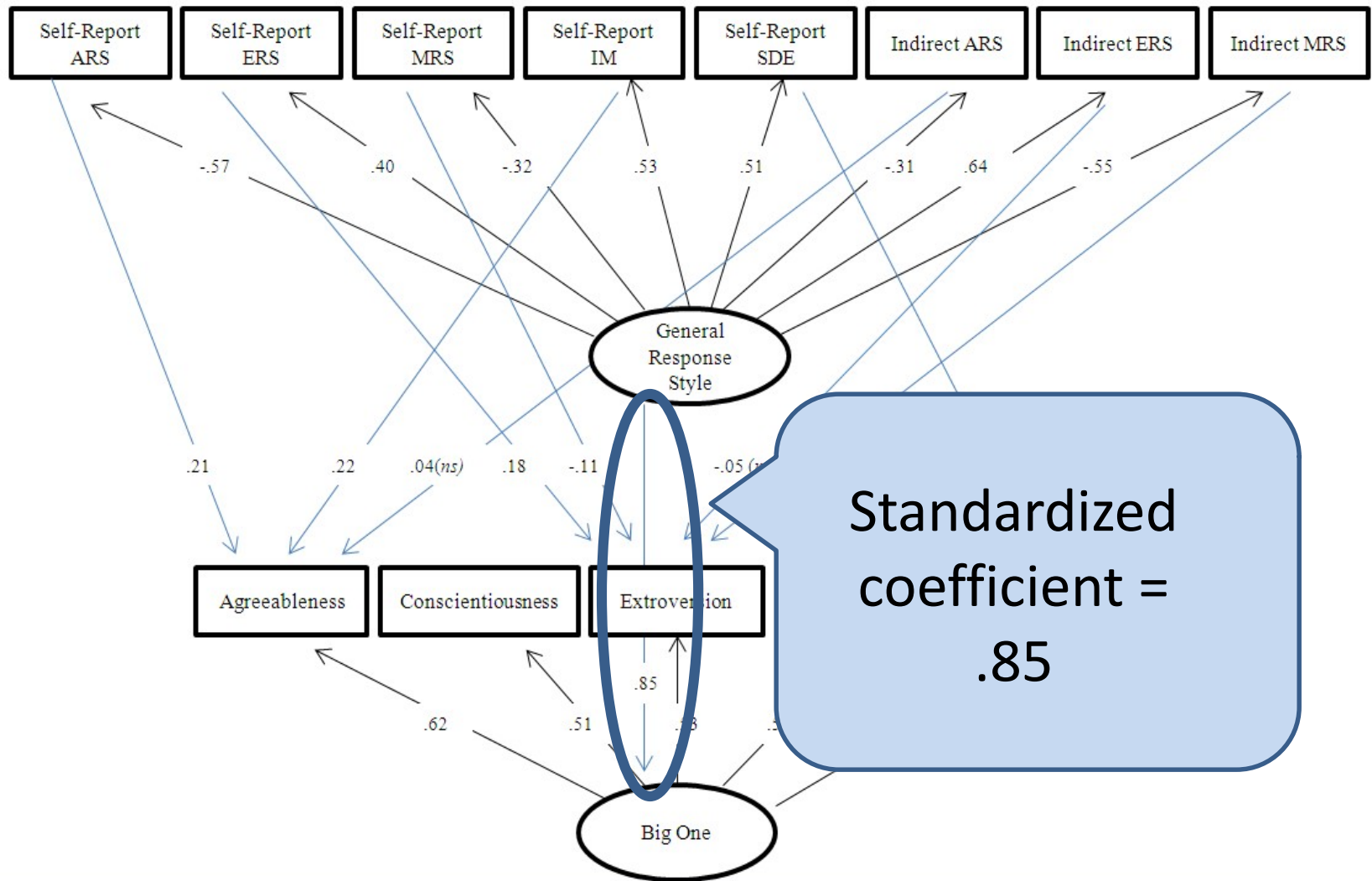
* $p < .05$

** $p < .01$

Correlates of the Response Style Factor

Personality Traits	Correlation with the Response Style Factor
Agreeableness	.27**
Conscientiousness	.33**
Extroversion	.37**
Openness	.35**
Neuroticism	-.35**
“Big One”	.52**

Response Style Factor and Personality Traits



Multigroup CFA: Structural residual model $\chi^2(328, N = 1301) = 798.16, p < .01$,
TLI = .87, CFI = .89, and RMSEA = .03

Correction for Response Styles in TALIS 2013

Public Opinion Quarterly, Vol. 79, Special Issue, 2015, pp. 267–290

**EFFECTS OF A GENERAL RESPONSE STYLE ON
CROSS-CULTURAL COMPARISONS
EVIDENCE FROM THE TEACHING AND LEARNING
INTERNATIONAL SURVEY**

JIA HE*

FONS J. R. VAN DE VIJVER

The Teaching and Learning International Survey 2013

- A 10-item social desirability scale in 18 countries (shortened, adapted Marlowe-Crowne scale)
- Various items available to construct indirect response style measures
- 17 core constructs with Likert-scale responses (e.g., teacher efficacy, job satisfaction, professional development, teacher-student relationship)

Samples in TALIS

<i>Country</i>	<i>Sample Size</i>
Abu Dhabi (UAE)	2187
Brazil	12490
Chile	1460
Croatia	3478
Estonia	2981
Finland	5890
France	2911
Iceland	1176
Korea	2783
Latvia	4103
Malaysia	3839
Mexico	5015
Poland	3694
Portugal	6809
Serbia	3549
Slovak Rep.	3404
Spain	9214
United States	1904
<i>Total</i>	<i>76887</i>

Items Social Desirability

<i>Item</i>	<i>Positive Impression Management</i>	<i>Negative Impression Management</i>
1. I always listen carefully to students	X	
2. I am confident about my judgments about students	X	
3. I have doubts about my ability to succeed as a teacher		X
4. I have always been honest with myself about my teaching qualities	X	
5. I feel threatened by teachers who are very successful		X
6. I have said things that hurt colleagues' or students' feelings		X
7. I feel angry when colleagues express ideas different from my own		X
8. I help students and colleagues in trouble	X	
9. I admit when I do not know something if a student asks a question in class	X	
10. I am irritated by students who ask for favours		X

CFA

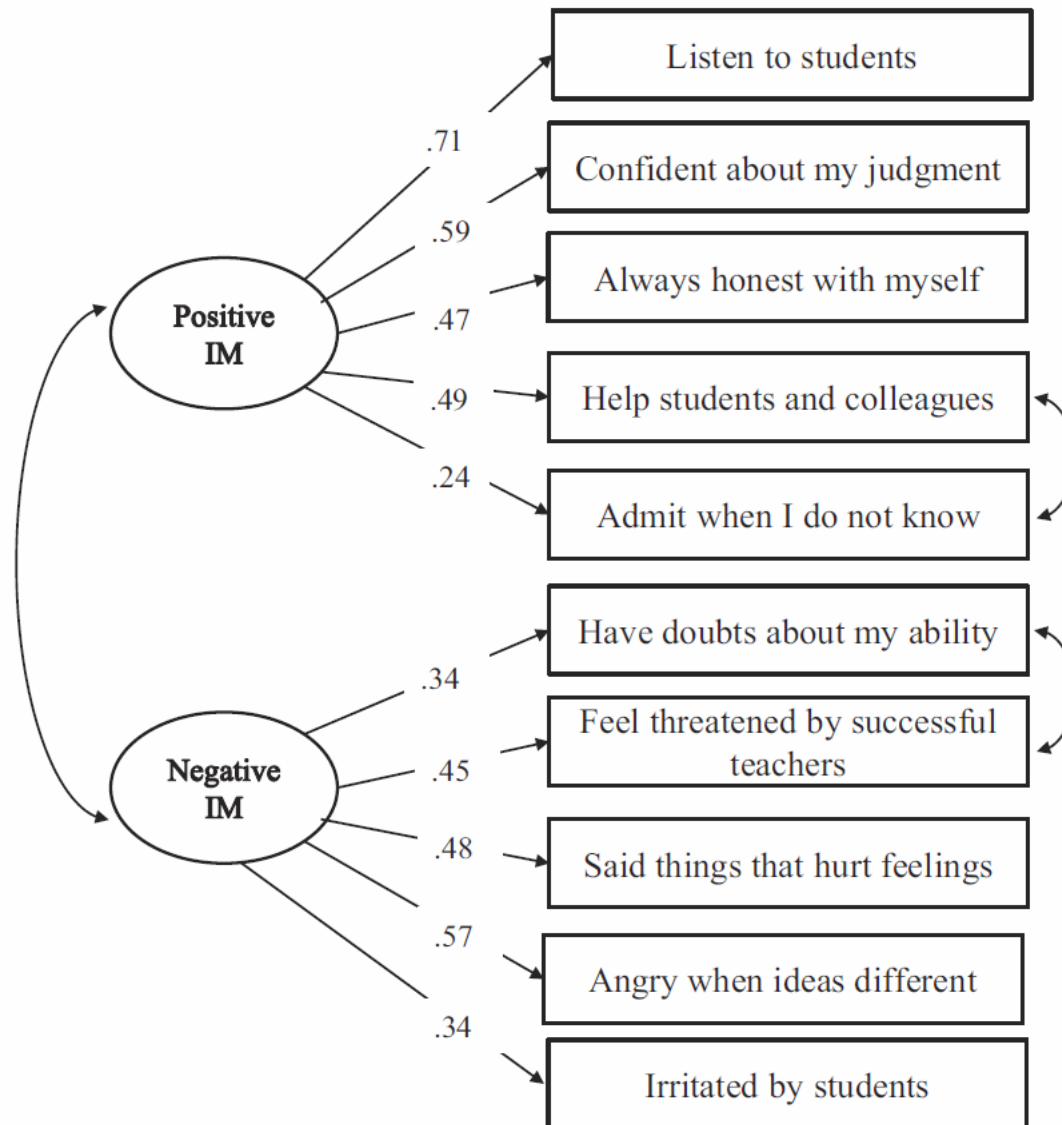


Table 3. Model Fit of the Multigroup Confirmatory Factor Analyses

Model	χ^2/df	CFI	TLI	RMSEA	$\Delta\chi^2/\Delta df$	ΔCFI	$\Delta RMSEA$
Configural							
invariance	10.46**	.93	.91	.01			
<i>Metric</i>							
<i>invariance</i>	<i>10.97**</i>	<i>.92</i>	<i>.90</i>	<i>.01</i>	<i>13.10**</i>	<i>-.01</i>	<i>.00</i>
Scalar							
invariance	76.98**	.20	.26	.03	353.46**	-.72	-.02
Partial scalar							
invariance	39.31**	.63	.63	.02	183.37**	-.29	-.01

NOTE.—Most restrictive model with an adequate fit is printed in italics. CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation.

** $p < .01$

Positive IM		Negative IM	
Country*	Scale Mean	Country*	Scale Mean
Abu Dhabi (UAE)	6.34	Abu Dhabi (UAE)	1.69
Mexico	6.24	Mexico	1.87
Portugal	6.20	Croatia	1.97
Serbia	6.19	Spain	2.01
Brazil	6.19	Brazil	2.08
Spain	6.18	Iceland	2.11
Iceland	6.18	Slovak Rep.	2.18
Croatia	6.09	United States	2.25
United States	6.07	Serbia	2.25
Chile	6.01	Chile	2.28
Poland	5.91	Malaysia	2.39
Latvia	5.91	Finland	2.49
Estonia	5.89	Estonia	2.52
Slovak Rep.	5.88	Portugal	2.57
France	5.88	France	2.57
Finland	5.84	Poland	2.76
Korea	5.40	Latvia	2.84
Malaysia	5.25	Korea	2.97

Response Style Measures in TALIS

- **Social Desirability**

- Positive: I always listen carefully to students
- Negative: I feel threatened by teachers who are very successful

- **Extremity and Midpoint Responding**

- Extracted with non-overlapping items of 4-point Likert anchors from various items (15 items each, average inter-item correlation .03 and .02)

- **The General Response Style**

- Factor score from the factor analysis of social desirability (.51), extremity (.91) and midpoint responding (-.91)

Correlation at Individual- and Country Level

Scale	Correlation with GRS at Individual Level	Correlation with GRS at Country Level
Efficacy in classroom management	.33**	.49*
Efficacy in instruction	.41**	.73**
Efficacy in student engagement	.35**	.47*
Overall teacher efficacy	.39**	.61**
Satisfaction with current work environment	.41**	.77**
Satisfaction with teaching profession	.36**	.56*
Overall job satisfaction	.42**	.72**
Exchange and coordination for teaching	.13**	.15
Professional collaboration	.10**	.17
Overall teacher cooperation	.13**	.22
Participation among stakeholders	.15**	-.34
Teacher-student relations	.42**	.50*
Classroom disciplinary climate	.18**	-.01
Constructivist beliefs	.29**	.29
Effectiveness in professional development	-.14**	-.33
Professional development in subject matter	.14**	.44
Professional development for diversity	-.02**	.04

GRS Correction in ANCOVA

- DV: each of the core constructs in TALIS
- IV: Country
- Covariate: GRS

Central question: How large are cross-cultural differences *before* and *after* statistical correction for GRS?

Scale	Effect Size Before Correction	Effect Size After Correction
Efficacy in classroom management	.11	.11
Efficacy in instruction	.21	.17
Efficacy in student engagement	.27	.26
Overall teacher efficacy	.18	.17
Satisfaction with current work environment	.06	.03
Satisfaction with teaching profession	.14	.13
Overall job satisfaction	.10	.07
Exchange and coordination for teaching	.24	.24
Professional collaboration	.30	.27
Overall teacher cooperation	.13	.12
Participation among stakeholders	.04	.05
Teacher-student relations	.05	.06
Classroom disciplinary climate	.10	.11
Constructivist beliefs	.07	.06
Effectiveness in professional development	.05	.05
Professional development in subject matter	.21	.19
Professional development for diversity	.23	.24

Scale	Correlation of Individual Corrected- Uncorrected Scores	Correlation of Country Corrected-Uncorrected Scores
Efficacy in classroom management	.95	.96
Efficacy in instruction	.91	.97
Efficacy in student engagement	.93	.98
Overall teacher efficacy	.92	.96
Satisfaction with current work environment	.91	.90
Satisfaction with teaching profession	.93	.95
Overall job satisfaction	.91	.92
Exchange and coordination for teaching	.99	1.00
Professional collaboration	1.00	1.00
Overall teacher cooperation	.99	.99
Participation among stakeholders	.99	.98
Teacher-student relations	.91	.88
Classroom disciplinary climate	.98	.98
Constructivist beliefs	.96	.93
Effectiveness in professional development	.99	.99
Professional development in subject matter	.99	1.00
Professional development for diversity	1.00	1.00

Conclusions

- Correlations of response styles with TALIS core constructs differ
 - Teacher efficacy and job satisfaction most related to response styles
 - Domain dependency: response styles most triggered in personal domains when evaluation apprehension is strongest
- Correction effects negligible
 - The correction leaves the country differences intact (cf. Ones et al., 1996)
- Should we bother about score corrections?
 - Substantial cross-cultural differences in response styles
 - Statistical correction does not have much impact

Response Styles in TIMSS and PIRLS

Data: Student Data

- Focus on background questionnaires
- PIRLS
 - Years: 2001, 2006, 2011
- TIMSS
 - Years
 - 1995, 2003, 2007, and 2011 (Grade 4)
 - 1995, 1999, 2003, 2007, and 2011 (Grade 8)

Examples of Item Categorizations

PIRLS	Response Format	Number of Categories	Number of items		
			2011	2006	2001
Student	{1, AGREE A LOT}...	4	29	16	15
	{1, AT LEAST ONCE A WEEK}...	4	1	11	
	{1, EVERY DAY}...	4	18		
	{1, LESS THAN 30 MINUTES}...	4	1		
	{1, ONCE A WEEK}...	4	6		
	{1, 5 HOURS OR MORE}...	5		4	
	{1, EVERY DAY OR ALMOST EVERY DAY}...	4		32	41
	{1, ALWAYS OR ALMOST ALWAYS}...	3			2

Indicators (only four-point scales):

- First set: most common response formats: (a) Likert; (b) every day...
- Second set: less common formats: could be Likert or frequency

TIMSS - grade 4

TIMSS - grade 4			Number of items					
	Response Format	Number of Categories	2011	2007	2003	Response Format	Number of Categories	1995
Student	{1, AGREE A LOT}...	4	39	19	15	{1, strongly agree}...	4	24
	{1, ALWAYS OR ALMOST ALWAYS}...	3	1	1				
	{1, EVERY DAY}...	4	7	4	3,2	{1, about every day}...	4	13
	{1, ONCE A WEEK}...	4	6					
	{1, NONE OR FEW (0-10)}...	5	1	1				
	{1, AT LEAST ONCE A WEEK}...	4		10	8			
	{1, EVERY OR ALMOST EVERY LESSON}...	4		11	10	{1, most lessons}...	3	30
	{1, NO TIME}...	5		8	8			9
	{1, EVERY OR ALMOST EVERY DAY}...	4			2			
						{1, like a lot}...	4	2
						{1, strongly negative}...	4	2
	{1, ALWAYS}...	4						

TIMSS - grade 8			Number of items						
	Response Format	Number of Categories	2011	2007	2003	Response Format	Number of Categories	1999*	1995*
Student BG	{1, AGREE A LOT}...	4	158	75	76	{1, strongly agree}...	4	111	24
	{1, TEACHER NEVER GIVES ME HOMEWORK}...	6	5						
	{1, NEVER GIVES ME HOMEWORK IN MAT/SCI}...	5	2						
	{1, EVERY DAY}...	5	6	13	15	{1, about every day}...	4	13	13
	{1, EVERY DAY OR ALMOST EVERY DAY}...	4	7	0	7				
	{1, AT LEAST ONCE A WEEK}...	4	6	0	0	{1, at least once a week}...	4	4	
	{1, ALWAYS}...	4	1	1	1	{1, always or almost always}...	3	1	
	{1, 0-10 BOOKS}...	5	1	1	1				
	{1, EVERY OR ALMOST EVERY LESSON}...	4		97	84	{1, almost always}...	4	166	
	{1, NO TIME}...	5		9	9	{1, no time}...	5	13	9
	{1, ZERO MINUTES}...	6		6					
	{1, FEWER THAN 15 MINUTES}...	5			7				
						{1, dislike a lot}...	4	6	
						{1, like a lot}...	4	6	2
						{1, low}...	3	12	
						{1, never}...	4	6	
						{1, not at all}...	4	6	
						{1, most lessons}...	3		30
						{1, strongly negative}...	4		2

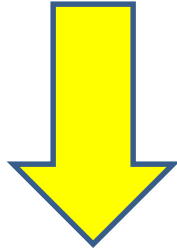
[illegible]

[illegible]

	PIRLS			TIMSS					TIMSS				
				Grade 8					Grade 4				
	2001	2006	2011	1995	1999	2003	2007	2011	1995	2003	2007	2011	<i>N</i>
Portugal									x			x	2
Qatar		x	x				x	x				x	5
Romania	x	x	x								x	x	5
Russian Federation												x	1
Saudi Arabia			x			x	x	x			x	x	6
Serbia												x	1
Singapore	x	x	x	x	x	x	x	x	x	x		x	11
Slovak Republic	x										x	x	3
Slovenia	x		x						x	x	x	x	6
South Africa		x				x		x			x		4
Spain		x	x	x								x	4
Sweden	x	x		x								x	4
Switzerland				x							x		2
Thailand				x			x	x	x			x	5
Tunisia					x	x	x	x		x		x	6
Turkey					x		x	x			x	x	5
United Arab Emirates								x				x	2
United Kingdom	x	x		x		x	x	x	x	x		x	9
United States of America				x	x	x	x	x	x	x	x	x	9
Yemen										x	x	x	3

Rank Order of Countries (ERS and MRS Combined)

Low



Korea	United States of America	Botswana
Japan	Israel	Lithuania
Malaysia	Iceland	Qatar
Hong Kong-China	Canada	Kuwait
Switzerland	Finland	Romania
China	Bulgaria	Saudi Arabia
Australia	Chile	Tunisia
Thailand	Greece	Ghana
Indonesia	Macedonia	Serbia
Singapore	Turkey	Austria
United Kingdom	Slovak Republic	Portugal
Italy	Slovenia	Oman
New Zealand	Spain	United Arab Emirates
Philippines	Occupied Palestinian Territory	Yemen
Norway	El Salvador	Russian Federation
Moldova, Rep,	Germany	Poland
Latvia	South Africa	Morocco
Sweden	Egypt	Croatia
Argentina	Iran, Islamic Republic of	Malta
Belize	Colombia	Kazakhstan
Cyprus	Jordan	Honduras
Netherlands	Hungary	Armenia

High

Interpretation (Tentative)

- East Asia low
- Protestant Europe low
- Central and South America high
- Muslim countries high
- Former Soviet Union high

How Large Are Country Differences in Response Styles?: Effect Size of Response Styles

- Proportion of Variance accounted for in ANOVA with Country as Independent Variable and Response Styles as Dependent Variables

Effect Sizes

TIMSS 4 grade						
	1995	1999	2003	2007	2011	Median
ERS Likert	.077		.189	.147	.197	.168
MRS Likert	.114				.195	.155
ERS Frequency	.086		.103	.194	.109	.106
MRS Frequency	.067					.067

TIMSS 8 grade						
	1995	1999	2003	2007	2011	Median
ERS Likert	.077	.037	.177	.203	.211	.177
MRS Likert	.114	.049	.229	.180	.193	.180
ERS Frequency	.086	.038	.088	.116	.067	.086
MRS Frequency	.067	.043		.141		.067

PIRLS				
	2001	2006	2011	Median
ERS Likert	.103	.131	.184	.131
MRS Likert			.166	.166
ERS Frequency	.060	.046	.052	.052
MRS Frequency	.076	.033		.055

Conclusion

- Effect sizes are substantial and larger for attitudinal questions (Likert) than for behavioral questions (frequency)
 - Large effect sizes of attitudes, medium effect sizes for behaviors

Stability of Response Style Scores across Years: PIRLS

Correlations

		ERSP1	ERSP6	ERSP11	ERS_FREQP 1	ERS_FREQP 6	ERS_FREQP 11
ERSP1	Pearson Correlation	1	.702*	.841**	.149	.602*	.681*
	Sig. (2-tailed)		.016	.001	.518	.050	.015
	N	21	11	12	21	11	12
ERSP6	Pearson Correlation	.702*	1	.803**	.090	.471*	.522
	Sig. (2-tailed)	.016		.002	.791	.036	.081
	N	11	20	12	11	20	12
ERSP11	Pearson Correlation	.841**	.803**	1	.163	.414	.675**
	Sig. (2-tailed)	.001	.002		.613	.181	.000
	N	12	12	26	12	12	26
ERS_FREQP1	Pearson Correlation	.149	.090	.163	1	.423	.566
	Sig. (2-tailed)	.518	.791	.613		.194	.055
	N	21	11	12	21	11	12
ERS_FREQP6	Pearson Correlation	.602*	.471*	.414	.423	1	.885**
	Sig. (2-tailed)	.050	.036	.181	.194		.000
	N	11	20	12	11	20	12
ERS_FREQP11	Pearson Correlation	.681*	.522	.675**	.566	.885**	1
	Sig. (2-tailed)	.015	.081	.000	.055	.000	
	N	12	12	26	12	12	26

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Response Style Scores Stable across Years: PIRLS

Correlations

		ERSP1	ERSP6	ERSP11	ERS_FREQP 1	ERS_FREQP 6	ERS_FREQP 11
ERSP1	Pearson Correlation	1	.702*	.841**	.149	.602*	.681*
	Sig. (2-tailed)		.016	.001	.518	.050	.015
	N	21	11	12	21	11	12
ERSP6	Pearson Correlation	.702*	1	.803**	.090	.471*	.522
	Sig. (2-tailed)	.016		.002	.791	.036	.081
	N	11	20	12	11	20	12
ERSP11	Pearson Correlation	.841**	.803**	1	.163	.414	.675**
	Sig. (2-tailed)	.001	.002		.613	.181	.000
	N	12	12	26	12	12	26
ERS_FREQP1	Pearson Correlation	.149	.090	.163	1	.423	.566
	Sig. (2-tailed)	.518	.791	.613		.194	.055
	N	21	11	12	21	11	12
ERS_FREQP6	Pearson Correlation	.602*	.471*	.414	.423	1	.885**
	Sig. (2-tailed)	.050	.036	.181	.194		.000
	N	11	20	12	11	20	12
ERS_FREQP11	Pearson Correlation	.681*	.522	.675**	.566	.885**	1
	Sig. (2-tailed)	.015	.081	.000	.055	.000	
	N	12	12	26	12	12	26

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Response Style Scores Stable across Years: MRS TIMSS Grade 8

Correlations

		MRS8G95	MRS8G99	MRS8G3	MRS8G7	MRS8G11
MRS8G95	Pearson Correlation	1	.940 ^{**}	.814 ^{**}	.832 ^{**}	.809 ^{**}
	Sig. (2-tailed)		.001	.002	.001	.003
	N	19	8	11	12	11
MRS8G99	Pearson Correlation	.940 ^{**}	1	.956 ^{**}	.914 ^{**}	.934 ^{**}
	Sig. (2-tailed)	.001		.000	.000	.000
	N	8	15	12	11	12
MRS8G3	Pearson Correlation	.814 ^{**}	.956 ^{**}	1	.940 ^{**}	.962 ^{**}
	Sig. (2-tailed)	.002	.000		.000	.000
	N	11	12	25	19	22
MRS8G7	Pearson Correlation	.832 ^{**}	.914 ^{**}	.940 ^{**}	1	.946 ^{**}
	Sig. (2-tailed)	.001	.000	.000		.000
	N	12	11	19	27	22
MRS8G11	Pearson Correlation	.809 ^{**}	.934 ^{**}	.962 ^{**}	.946 ^{**}	1
	Sig. (2-tailed)	.003	.000	.000	.000	
	N	11	12	22	22	27

^{**}. Correlation is significant at the 0.01 level (2-tailed).

Conclusion

- So, Response Styles are stable country characteristics
 - Note that items and countries were not exactly the same across years

Associations of Response Style Scores with Country Characteristics

- Computed correlations of country-level Response Style scores with various characteristics
- Procedure for GDP (ppp) per capita is explained and results of the other variables summarized

		GDP per capita		
Style	Year	TIMSS Grade 4	TIMSS Grade 8	PIRLS
ERS Likert	1995	.20	-.47	
	1999		-.55	
	2003	-.36	-.43	
	2007	-.37	-.36	
	2011		-.33	
ERS Frequency	1995	-.08	-.40	
	1999		-.02	
	2003	-.58	-.57	
	2007	-.49	-.38	
	2011		.01	
MRS Likert	1995		.47	
	1999		.47	
	2003		.58	
	2007		.24	
	2011		.33	
ERS Likert	2001			-.56
	2006			-.32
	2011			-.45
ERS Frequency	2001			-.35
	2006			-.32
	2011			-.37

- Median correlations between GDP per capita and Response Style scores stable at country level:

	PIRLS	TIMSS
ERS Likert	-.37	-.45
ERS Frequency	-.39	-.35
MRS Likert	.47	

- PIRLS and TIMSS rather similar pattern
- So, more affluent countries lower on extremity and higher on midpoint responding
 - Later more about interpretation

- We did the same procedure for many country-level variables
- We established the significance of the correlations by bootstrapping

Correlations of Response Styles with Affluence-Related Indicators

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
GDP per capita	-.36*	-.43*		.38*	.12
Expenditure on education (ppp)	-.42*	-.49*		.49*	.54*
Enrolment rate secondary education	-.40*	-.29*		.47*	-.10
Completion rate secondary education	-.38*	.06		.40*	.57*

Conclusion:

- Affluence-related factors negatively related with ERS and positively with MRS
- Pattern clearer for Likert than for frequency scales

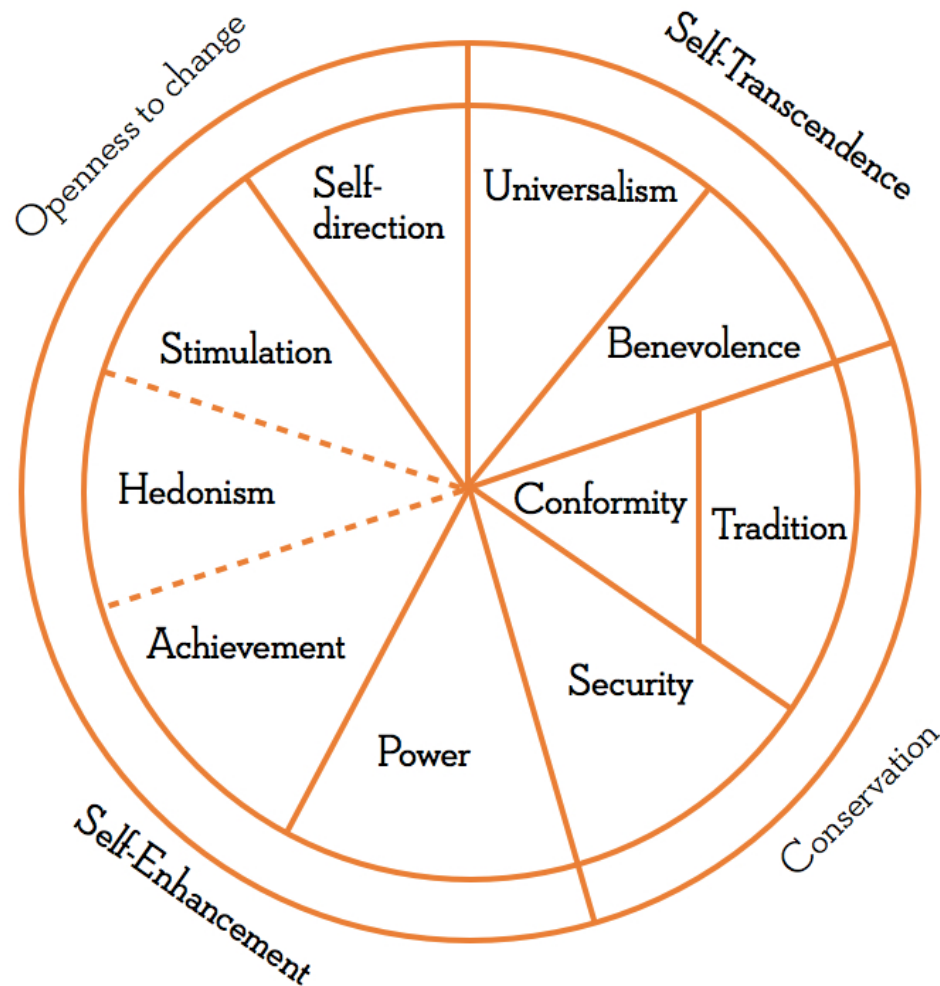
Correlations of Response Styles with Hofstede's Indicators

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
Power Distance	-.10	.09		.04	-.08
Individualism	.05	-.05		-.14	.11
Masculinity	-.15	.06		.10	-.12
Uncertainty avoidance	.14	.54*		-.19	.07

Conclusion:

- Remarkable absence of systematic relationships

Schwartz's Value Model



Correlations of Response Styles with Hofstede's Indicators

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
Harmony	-.12	.11		.16	.28*
Embeddedness (Conservation)	.48*	.33		-.51*	-.38*
Hierarchy	-.06	-.09		-.01	-.12
Masculinity	-.18*	-.12		.14	-.02
Affective autonomy	-.29*	-.34*		.33*	.21*
Intellectual autonomy	-.36*	-.13		.31*	.35*
Egalitarianism	.06	.02		-.02	.12

Conclusion:

- Systematic relationships, notably with embeddedness vs. autonomy

Correlations with Personality

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
Neuroticism McCrae	.07	.23*		-.06	.03
Neuroticism John	-.61*	-.17		.65*	.27*
Extroversion McCrae	.14*	-.20*		-.10	.24
Extroversion John	.46*	.01		-.48*	-.23*
Openness McCrae	-.22*	.00		.26*	.26
Openness John	.50*	.28*		-.53*	-.07
Agreeableness McCrae	.08	-.07		-.07	.22
Agreeableness John	.78*	.13		-.87*	-.25
Conscientiousness McCrae	.06	.13*		-.11	.16*
Conscientiousness John	.73*	.39*		-.76*	-.26

- Conclusion
 - Analysis troubled by strong discrepancies in country scores on personality
 - John's instrument has intuitively more appealing results, such as strong link with conscientiousness and agreeableness

Correlations of Response Styles with Subjective Well-Being

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
Positive emotionality	.64*	.25*		-.63*	-.21
Subjective well-being	.15*	.07*		-.23*	-.02
Happiness index	.37*	.13		-.45*	-.37*
Beck Depression Score (reversed)	.48*	.46*		-.46*	-.14

Conclusion:

- Systematic relationships, despite variety of measures
- Expressiveness as underlying construct?

Correlations of Response Styles with Teacher Indicators (TALIS)

	ERS Likert	ERS Frequency		MRS Likert	MRS Frequency
Teacher Efficacy	.55*	.19		-.59*	-.29*
Job Satisfaction	.45*	.02		-.45*	.02
Classroom Discipline	-.22	.06		.15	-.10
Constructive Beliefs	.10	.10		-.12*	-.09*
Teacher Cooperation	.29*	-.26*		-.34*	-.02
Effective Professional Development	.12*	.01		-.23*	-.30*

Conclusion:

- Some relationships, notably with domains of more personal involvement, pointing again to the role of expressiveness

Correlations of Response Styles with Educational Achievement

	ERS Likert	ERS Frequency		MRS Likert	MRS Fequency
Math	-.75*	-.43*		.81*	.13
Science	-.73*	-.39*		.75*	.12

Conclusion:

- Strong relationships, notably with Likert measures
- Note: correlation is negative

Conclusion

- Response styles are correlated at country level with:

		ERS	MRS
Affluence		-	+
Values	Embeddedness	+	-
	Autonomy	-	+
Well-being		+	-
Teacher characteristics	Efficacy	+	-
	Job satisfaction	+	-
Educational Achievement		-	+

Unpacking These Correlations

- Patterning of the correlations (+, -) is similar for adults
- Background of the correlations may be multifaceted

My Interpretation: (1) Role of Affluence

- Affluence positively linked with educational quality, autonomy, individualism
- Affluence negatively linked to conformity
- More conformity (embeddedness) associated with more extremity and less midpoint responding

(2) Role of Expressiveness, General Response Style

- Teacher job satisfaction positively linked to affluence, also possible that link is due to shared preferences in response styles
 - Preference for extremity/moderation as cultural characteristic
 - Moderation in East Asia, amplification in Central and South America and Muslim countries
 - Possible explanation of motivation achievement paradox

Correction for Response Styles

- What is the correlation of raw and corrected scores on various scales?
- Is the size of country differences influenced by corrections for Response Styles?
 - Design: ANOVA on Attitude Scales with Response Style as covariate

Correlations Uncorrected— Corrected Scores

TIMSS Grade 4	2003	2007	2011
Self-confidence in Math	.98	.99	.89
Self-confidence in Science	.96	.97	.91
Students bullied at school		1.00	1.00

TIMSS Grade 8	1999	2003	2007	2011
Self-confidence in Math	1.00	.98	.99	.96
Self-confidence in Science	1.00	.97	.97	.92
Students bullied at school		1.00	1.00	1.00
Value learning Math		.96	.98	.95
Value learning Science		.96	.94	.93

PIRLS	2001	2006	2011
Self-confidence in reading	.97	.96	.96
Motivation to read	.96	.97	.90
Students bullied at school		1.00	1.00

Conclusion: correlations corrected—uncorrected scores very high

Effect Sizes Before and After Correction

TIMSS Grade 4	2003		2007		2011	
	Before	After	Before	After	Before	After
Self-confidence in Math	.040	.030	.040	.040	.069	.034
Self-confidence in Science	.032	.025	.026	.031	.074	.043
Students bullied at school			.058	.062	.050	.053

TIMSS Grade 8	1999		2003		2007		2011	
	Before	After	Before	After	Before	After	Before	After
Self-confidence in Math	.083	.069	.055	.052	.062	.066	.077	.068
Self-confidence in Science	.080	.081	.097	.070	.105	.070	.132	.074
Students bullied at school			.133	.142	.076	.058	.088	.093
Value learning Math			.160	.143	.122	.094	.126	.108
Value learning Science			.185	.148	.144	.094	.151	.098

PIRLS	2001		2006		2011	
	Before	After	Before	After	Before	After
Self-confidence in reading	.021	.025	.030	.026	.024	.041
Motivation to read	.044	.047	.051	.054	.051	.010
Students bullied at school			.049	.049	.064	.074

Conclusion:

- Country differences tend to become smaller after correction (about 11%), effect is larger for TIMSS than for PIRLS

Overall Conclusion

- Usage of Response Styles is a stable, replicable country characteristic
 - Stronger in agreement format than in frequency format
- Related to affluence and cultural norms regarding self-expression
- Corrections for Response Styles tend to reduce cross-cultural differences by about 10%
- Individual differences within a single country and country differences should be interpreted in a different manner
- Response Styles: We should bother...

Thank you for your
attention

- A priori procedures (anchoring vignettes and direct measures of response styles (i.e., acquiescence, extremity, midpoint responding, and social desirability))
- A posteriori procedures focusing on data transformations prior to analysis (ipsatization and item parceling)
- Two data modeling procedures (treating data as continuous vs. as ordered categories)

Methodology of Educational Measurement and Assessment

Susanne Kuger
Eckhard Klieme
Nina Jude
David Kaplan *Editors*

Assessing Contexts of Learning

An International Perspective

 Springer