Appendix methods for identifying CPL needs in four countries

Rikkert M. van der Lans

2023-08-29

Contents

Software Information

The following packages were loaded and mostly used for analysis. In addition several Latex packages not listed below were loaded. These were: *titlesec*, *floatrow*, *subfig*, *tikz*, *positioning*, *arrows.meta*, *arrows*, and *pdflscape*.

```
##packages for circular statistics
if(!require("circular")) install.packages("circular"); library("circular")
if(!require("DescTools")) install.packages("DescTools"); library("DescTools")
if(!require("bpnreg")) install.packages("bpnreg"); library("bpnreg")

# package for multilevel correlations
if(!require("correlation")) install.packages("correlation"); library("correlation")

# pacakages for linear IRT and unfolding scaling
if(!require("ltm")) install.packages("ltm"); library("ltm")
if(!require("mirt")) install.packages("mirt"); library("mirt")

## packages for datamanagement
if(!require("plyr")) install.packages("plyr"); library("plyr")
if(!require("dplyr")) install.packages("dplyr"); library("dplyr")
if(!require("data.table")) install.packages("data.table"); library("data.table")
if(!require("rstatix")) install.packages("rstatix"); library("rstatix")

# packages for data visualization
if(!require("kableExtra")) install.packages("kableExtra"); library("kableExtra")
if(!require("pryr")) install.packages("pryr"); library("pryr")
if(!require("tidyverse")) install.packages("tidyverse"); library("tidyverse")
if(!require("ggplot2")) install.packages("ggplot2"); library("ggplot2")
if(!require("WrightMap")) install.packages("WrightMap"); library("WrightMap")

## packages for reading in data
if(!require("foreign")) install.packages("foreign"); library("foreign")

## packages for data imputation
if(!require("mice")) install.packages("mice"); library("mice")
if(!require("mitools")) install.packages("mitools"); library("mitools")
if(!require("miceadds")) install.packages("miceadds"); library("miceadds")
```

Information about the R and Rstudio version used is printed below.

```
## R version 4.2.2 (2022-10-31 ucrt)
```

```
## Rstudio version 2023.06.0+421
```

Data and sample descriptives

The data available on the TALIS 2018 website and can be found here. Loaded datasets were:

1. ATGAUST3 *Australian primary school teacher data*
2. ATGENGT3 *English primary school teacher data*
3. ATGJPNT3 *Japanese primary school teacher data*
4. ATGNLDT3 *Dutch primary school teacher data*

```
## Teacher self-report data Primary Education.
TAL18.PO.AUS = read.spss("ATGAUST3.sav", use.value.labels = FALSE,
                         to.data.frame = TRUE)
TAL18.PO.ENG = read.spss("ATGENGT3.sav", use.value.labels = FALSE,
                         to.data.frame = TRUE)
TAL18.PO.JPN = read.spss("ATGJPNT3.sav", use.value.labels = FALSE,
                         to.data.frame = TRUE)
TAL18.PO.NLD = read.spss("ATGNLDT3.sav", use.value.labels = FALSE,
                         to.data.frame = TRUE)
```

Sample Australia

The Australian dataset contains self-reports of 3030 teachers within 213 schools for primary education. The majority of the participants is female (n = 2601), with 14.53 (SD = 11.1) years of teaching experience. Teachers report to spent a mean number of 23.83 (SD = 10.07) hours on teaching and a mean number of 1.77 (SD = 2.49) hours on professional development. The range in hours spent on teaching and hours spent on PD is however, considerable [0 – 95 hours] and [0 – 50 hours] respectively.

The employment status varies much less between teachers. The majority of the teachers n = 2153 have a permanent position. The remaining 892 teachers have a temporary position. Then majority of the teachers has a full-time position.

Sample England

The English dataset contains self-reports of 2009 teachers within 152 schools for primary education. The majority of the participants is female (n = 1661), with 12.11 (SD = 8.9) years of teaching experience. Teachers report to spent a mean number of 22.43 (SD = 9.81) hours on teaching and a mean number of 1.32 (SD = 2.01) hours on professional development. The range in hours spent on teaching and hours spent on PD is however, considerable [0 – 90 hours] and [0 – 40 hours] respectively.

The employment status varies much less between teachers. The majority of the teachers n = 1755 have a permanent position. The remaining 254 teachers have a temporary position. Judging by the eye, teachers with permanent and those with temporary positions are fairly equally distributed over the different contract types, i.e. Full-time, more than 70%, more than 50%, less than 50%.

Sample Japan

The Japanese dataset contains self-reports of 3308 teachers within 197 schools for primary education. The majority of the participants is female (n = 2038), with 16.7 (SD = 11.92) years of teaching experience. Teachers report to spent a mean number of 23.2 (SD = 8.97) hours on teaching and a mean number of 0.67 (SD = 2.08) hours on professional development. The range in hours spent on teaching and hours spent on PD is however, considerable [0 – 90 hours] and [0 – 74 hours] respectively.

The employment status varies much less between teachers. The majority of the teachers n = 2436 have a permanent position. The remaining 872 teachers have a temporary position. Judging by the eye, teachers with permanent and those with temporary positions are fairly equally distributed over the different contract types, i.e. Full-time, more than 70%, more than 50%, less than 50%.

Sample the Netherlands

The Dutch dataset contains self-reports of 1504 teachers within 130 schools for primary education. The majority of the participants is female (n = 1274), with 16.26 (SD = 10.66) years of teaching experience. Teachers report to spent a mean number of 19.39 (SD = 7.83) hours on teaching and a mean number of 1.66 (SD = 2.46) hours on professional development. The range in hours spent on teaching and hours spent on PD is however, considerable [0 – 85 hours] and [0 – 40 hours] respectively.

The employment status varies much less between teachers. The majority of the teachers n = 1357 have a permanent position. The remaining 147 teachers have a temporary position. Judging by the eye, teachers with permanent and those with temporary positions are fairly equally distributed over the different contract types, i.e. Full-time, more than 70%, more than 50%, less than 50%.

Table 1: Teaching experience and spending of hours

	N	missing responses	M	Mdn	SD	Min	Max
Australia							
teacher experience (in years)	3030	108	14.53	12	11.1	0	52
weekly hours spent on teaching	3030	149	23.83	24	10.07	0	95
weekly hours spent on Professional Development	3030	189	1.77	1	2.49	0	50
England							
teacher experience (in years)1	2009	24	12.11	10	8.9	0	48
weekly hours spent on teaching1	2009	56	22.43	24	9.81	0	90
weekly hours spent on Professional Development1	2009	73	1.32	1	2.01	0	40
Japan							
teacher experience (in years)2	3308	17	16.7	14	11.92	0	55
weekly hours spent on teaching2	3308	40	23.2	24	8.97	0	90
weekly hours spent on Professional Development2	3308	100	0.67	0	2.08	0	74
Netherlands							
teacher experience (in years)3	1504	18	16.26	14	10.66	0	45
weekly hours spent on teaching3	1504	34	19.39	19	7.83	0	85
weekly hours spent on Professional Development3	1504	49	1.66	1	2.46	0	40

Table 2: Distribution of teacher experience by country ISED-1 Teachers

	0-5 years	6-10 years	11-15 years	15-20 years	20-30 years	31 years or more
Austrlia						
Number	757	590	412	357	459	323
percentage	0.26	0.2	0.14	0.12	0.16	0.11
England						
Number	567	415	323	294	288	70
percentage	0.29	0.21	0.17	0.15	0.15	0.04
Japan						
Number	685	621	463	272	637	577
percentage	0.21	0.19	0.14	0.08	0.2	0.18
Netherlands						
Number	218	279	292	243	246	187
percentage	0.15	0.19	0.2	0.17	0.17	0.13

Table 3: Descriptives Employment status by country ISED-1 Teachers

	full-time	part-time 70-90%	part-time 50-70%	part-time <50%
Austrlia				
permanent	1662	264	170	57
fixed-term contract > 1 year	148	31	19	5
fixed-term contract =< 1 year	339	83	42	29
England				
permanent	1395	157	151	52
fixed-term contract > 1 year	42	2	2	2
fixed-term contract =< 1 year	77	5	15	11
Japan				
permanent	2381	34	17	4
fixed-term contract > 1 year	162	10	2	7
fixed-term contract =< 1 year	363	48	48	53
Netherlands				
permanent	449	330	446	132
fixed-term contract > 1 year	17	6	5	5
fixed-term contract =< 1 year	36	23	14	13

Share of teaching hours relative to hours on the job

The questions 14, 15 and 16 provide informations concerning the share of teaching hours relative to non-teaching activities. These were:

16. *During your most recent complete calendar week, approximately how many 60-minute hours did you spend in total on tasks related to your job at this school?*
a) ... hour

17. *Of this total, how many 60-minute hours did you spend on teaching at this school during your most recent complete calendar week?*
a) ... hour

18. *Approximatwely how many 60-minute hours did you spend on the following tasks during your most recent complete calendar week, in your job at this school?*

a) ... Hours Individual planning or preparation of lessons either at school or out of school
b) ... Hours Team work and dialogue with colleagues within this school
c) ... Hours Marking/correcting of student work
d) ... Hours Counselling students (including student supervision, mentoring, virtual counselling, career guidance and behaviour guidance)
e) ... Hours Participation in school management
f) ... Hours General administrative work (including communication, paperwork and other clerical duties)
g) ... Hours Professional development activities
h) ... Hours Communication and co-operation with parents or guardians
i) ... Hours Engaging in extracurricular activities (e.g. sports and cultural activities after school)
j) ... Hours Other work tasks

The percentage of teaching hours relative to the total hours can be estimated by dividing the response to question 17 by question 16 or by dividing the response to question 16 by the sum of question 18. Both strategies were explored.

Australia

Based on the ratio of questions 17 and 16 the estimated number of hours spent on teaching is 0.5(SD = 0.17). Of the 3030 participating teachers, 513 report to spent above 100% hours on teaching, meaning that the reported hours spent on teaching (question 17) in the last calender week was higher than the reported hours spent on total on the job (question 16). It is unclear how to interpret this result. Another smaller group of 2945, 598 reported to spent less than 10% time on teaching.

A second strategy divides the hours spent on teaching by the hours spent (question 16) on other activities plus the hours on teaching (questions 16 + 18). By definition, this strategy cannot result in estimates of above 100% hours spent on teaching. The distribution of hours spent on teaching as estimated by this strategy is 0.49(SD = 0.17). The correlation between the two methods is high 0.7 but also indicates that the choice for any of the strategies likely impacts on the results.

England

Based on the ratio of questions 17 and 16 the estimated number of hours spent on teaching is 0.45(SD = 0.18). Of the 2009 participating teachers, 204 report to spent above 100% hours on teaching, meaning that the reported hours spent on teaching (question 17) in the last calender week was higher than the reported hours spent on total on the job (question 16). It is unclear how to interpret this result. Another smaller group of 1915, 298 reported to spent less than 10% time on teaching.

A second strategy divides the hours spent on teaching by the hours spent (question 16) on other activities plus the hours on teaching (questions 16 + 18). By definition, this strategy cannot result in estimates of above 100% hours spent on teaching. The distribution of hours spent on teaching as estimated by this strategy is 0.45(SD = 0.17). The correlation between the two methods is high 0.77 but also indicates that the choice for any of the strategies likely impacts on the results.

Japan

Based on the ratio of questions 17 and 16 the estimated number of hours spent on teaching is 0.43(SD = 0.16). Of the 3308 participating teachers, 197 report to spent above 100% hours on teaching, meaning that the reported hours spent on teaching (question 17) in the last calender week was higher than the reported hours spent on total on the job (question 16). It is unclear how to interpret this result. Another smaller group of 3210, 295 reported to spent less than 10% time on teaching.

A second strategy divides the hours spent on teaching by the hours spent (question 16) on other activities plus the hours on teaching (questions 16 + 18). By definition, this strategy cannot result in estimates of above 100% hours spent on teaching. The distribution of hours spent on teaching as estimated by this strategy is 0.46(SD = 0.17). The correlation between the two methods is high 0.7 but also indicates that the choice for any of the strategies likely impacts on the results.

Netherlands

Based on the ratio of questions 17 and 16 the estimated number of hours spent on teaching is 0.52(SD = 0.14). Of the 1504 participating teachers, 142 report to spent above 100% hours on teaching, meaning that the reported hours spent on teaching (question 17) in the last calender week was higher than the reported hours spent on total on the job (question 16). It is unclear how to interpret this result. Another smaller group of 1486, 160 reported to spent less than 10% time on teaching.

A second strategy divides the hours spent on teaching by the hours spent (question 16) on other activities plus the hours on teaching (questions 16 + 18). By definition, this strategy cannot result in estimates of above 100% hours spent on teaching. The distribution of hours spent on teaching as estimated by this strategy is 0.51(SD = 0.14). The correlation between the two methods is high 0.72 but also indicates that the choice for any of the strategies likely impacts on the results.

Descriptives for teachers' continous professional leanring needs

TALIS question 25 tapped teachers' professorial learning needs. The question stem was: *For each of the areas listed below, please indicate the extent to which you currently need professional learning activities.* The question was followed by a list of areas A to N and an open-ended response option O "other". Teachers responded to the listed aeras on a four point Likert-scale with response options: 1 = *no need at present*, 2 = *low level of need*, 3 = *moderate level of need*, 4 = *high level of need*.

Table 4: Descriptives CPL needs per country

		Australia						England						Japan						The Netherlands					
		freq. level of need				M	SD	freq. level of need				M	SD	freq. level of need				M	SD	freq. level of need				M	SD
area	Item 25: For each of the areas listed below, please indicate the extent to which you currently need professional learning activities	1	2	3	4			1	2	3	4			1	2	3	4			1	2	3	4		
A	Knowledge and understanding of my subject field(s)	733	1380	604	71	2	0.76	772	892	223	24	1.74	0.71	18	85	1381	1782	3.51	0.58	246	628	459	110	2.3	0.84
B	Pedagogical competencies in teaching my subject field(s)	702	1369	633	76	2.03	0.77	738	917	227	23	1.76	0.7	16	76	1182	1996	3.58	0.57	307	643	402	90	2.19	0.84
C	Knowledge of the curriculum	658	1407	634	87	2.05	0.77	779	906	198	25	1.72	0.7	34	454	1848	929	3.12	0.67	241	642	475	83	2.28	0.81
D	Student assessment practices	467	1161	991	164	2.31	0.82	580	879	396	48	1.95	0.78	23	187	1566	1482	3.38	0.63	319	642	400	80	2.17	0.83
E	ICT (information and communication technology) skills for teaching	261	855	1175	501	2.69	0.87	354	728	654	168	2.33	0.88	35	302	1685	1244	3.27	0.67	193	437	574	240	2.6	0.92
F	Student behaviour and classroom management	814	1213	593	164	2.04	0.86	867	828	184	25	1.67	0.7	28	224	1307	1699	3.44	0.66	208	521	538	174	2.47	0.88
G	School management and administration	963	1094	586	138	1.96	0.87	741	713	392	57	1.88	0.84	195	1061	1449	550	2.72	0.81	721	490	163	64	1.7	0.84
H	Approaches to individualised learning	466	1140	943	236	2.34	0.85	548	883	437	37	1.98	0.77	18	129	1344	1774	3.49	0.6	163	446	611	225	2.62	0.88
I	Teaching students with special needs	391	1009	1049	332	2.48	0.88	396	813	598	100	2.21	0.83	19	122	1163	1961	3.55	0.6	143	333	655	311	2.79	0.89
J	Teaching in a multicultural or multilingual setting	802	1039	735	203	2.12	0.91	682	710	427	81	1.95	0.87	89	755	1731	687	2.92	0.74	550	588	226	79	1.88	0.86
K	Teaching cross-curricular skills (e.g. creativity, critical thinking, problem solving)	390	1051	1085	260	2.44	0.84	563	826	455	58	2	0.81	35	350	1738	1135	3.22	0.67	144	366	647	285	2.74	0.89
L	Analysis and use of student assessments	486	1187	905	205	2.3	0.84	571	880	403	52	1.97	0.79	28	318	1729	1182	3.25	0.66	241	598	474	127	2.34	0.86
M	Teacher-parent/guardian co-operation	977	1305	418	80	1.86	0.77	915	774	188	25	1.64	0.71	40	384	1640	1195	3.22	0.7	375	668	314	85	2.08	0.84
N	Communicating with people from different cultures or countries	918	1215	526	123	1.95	0.83	789	760	296	56	1.8	0.81	76	690	1821	674	2.95	0.71	642	587	175	39	1.73	0.78

Correlations for teachers' continous professional leanring needs

Correlations were estimated per country. The correlatons matrix serves as input for the circumplex analyses. Multilevel multiple imputation was performed to impute missing values. Procedures followed descriptions as provided here.

Table 5: Australia Pearson correlations CPL needs

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.747	1												
C	Knowledge curriculum	0.659	0.645	1											
D	Student assessment	0.551	0.586	0.587	1										
E	ICT skills	0.296	0.3	0.329	0.337	1									
F	classroom management	0.469	0.504	0.452	0.48	0.251	1								
G	administration	0.247	0.311	0.291	0.284	0.178	0.342	1							
H	individualised learning	0.491	0.543	0.481	0.583	0.276	0.555	0.335	1						
I	teaching students with special needs	0.363	0.407	0.36	0.434	0.245	0.498	0.279	0.613	1					
J	teaching in multicultural settings	0.269	0.316	0.265	0.27	0.2	0.337	0.317	0.374	0.478	1				
K	teaching cross-curricular skills	0.381	0.438	0.41	0.476	0.327	0.383	0.308	0.548	0.452	0.401	1			
L	analysis of assessments	0.464	0.507	0.48	0.689	0.287	0.447	0.319	0.587	0.451	0.331	0.532	1		
M	teacher-parent cooperation	0.445	0.448	0.411	0.445	0.21	0.52	0.414	0.495	0.442	0.412	0.426	0.497	1	
N	Communicating other cultures	0.296	0.315	0.293	0.278	0.191	0.352	0.347	0.332	0.388	0.663	0.348	0.336	0.565	1

Table 6: England Pearson correlations CPL needs

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.741	1												
C	Knowledge curriculum	0.67	0.627	1											
D	Student assessment	0.498	0.527	0.589	1										
E	ICT skills	0.311	0.315	0.349	0.345	1									
F	classroom management	0.506	0.515	0.509	0.489	0.274	1								
G	administration	0.261	0.302	0.311	0.294	0.213	0.354	1							
H	individualised learning	0.477	0.523	0.493	0.54	0.284	0.54	0.335	1						
I	teaching students with special needs	0.401	0.433	0.402	0.419	0.266	0.475	0.297	0.63	1					
J	teaching in multicultural settings	0.269	0.313	0.264	0.298	0.198	0.338	0.292	0.394	0.458	1				
K	teaching cross-curricular skills	0.452	0.504	0.473	0.469	0.308	0.462	0.34	0.613	0.495	0.413	1			
L	analysis of assessments	0.428	0.47	0.486	0.643	0.304	0.465	0.362	0.519	0.434	0.348	0.535	1		
M	teacher-parent cooperation	0.419	0.455	0.45	0.411	0.232	0.527	0.392	0.488	0.41	0.411	0.505	0.49	1	
N	Communicating other cultures	0.283	0.324	0.305	0.3	0.207	0.349	0.333	0.371	0.374	0.664	0.397	0.359	0.563	1

Table 7: Japan Pearson correlations CPL needs

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.855	1												
C	Knowledge curriculum	0.424	0.397	1											
D	Student assessment	0.506	0.532	0.5	1										
E	ICT skills	0.312	0.329	0.36	0.401	1									
F	classroom management	0.508	0.531	0.375	0.518	0.419	1								
G	administration	0.227	0.204	0.48	0.342	0.345	0.331	1							
H	individualised learning	0.472	0.504	0.348	0.492	0.38	0.563	0.287	1						
I	teaching students with special needs	0.37	0.395	0.305	0.405	0.356	0.512	0.258	0.658	1					
J	teaching in multicultural settings	0.249	0.244	0.378	0.331	0.378	0.329	0.401	0.369	0.395	1				
K	teaching cross-curricular skills	0.397	0.397	0.458	0.45	0.397	0.454	0.399	0.442	0.395	0.515	1			
L	analysis of assessments	0.455	0.452	0.466	0.641	0.41	0.513	0.385	0.521	0.441	0.422	0.58	1		
M	teacher-parent cooperation	0.37	0.382	0.396	0.45	0.369	0.539	0.424	0.507	0.479	0.42	0.466	0.534	1	
N	Communicating other cultures	0.276	0.261	0.377	0.348	0.381	0.357	0.399	0.389	0.385	0.759	0.51	0.464	0.492	1

Table 8: The Netherlands Pearson correlations CPL needs

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.521	1												
C	Knowledge curriculum	0.481	0.456	1											
D	Student assessment	0.396	0.316	0.498	1										
E	ICT skills	0.182	0.167	0.217	0.26	1									
F	classroom management	0.382	0.567	0.326	0.291	0.215	1								
G	administration	0.164	0.107	0.192	0.246	0.1	0.114	1							
H	individualised learning	0.326	0.326	0.342	0.348	0.259	0.357	0.145	1						
I	teaching students with special needs	0.273	0.361	0.305	0.276	0.192	0.436	0.143	0.517	1					
J	teaching in multicultural settings	0.178	0.211	0.236	0.212	0.132	0.205	0.261	0.257	0.358	1				
K	teaching cross-curricular skills	0.293	0.276	0.294	0.3	0.24	0.315	0.176	0.426	0.39	0.267	1			
L	analysis of assessments	0.332	0.288	0.389	0.49	0.246	0.283	0.208	0.393	0.332	0.203	0.39	1		
M	teacher-parent cooperation	0.348	0.422	0.366	0.36	0.228	0.386	0.255	0.349	0.36	0.314	0.313	0.464	1	
N	Communicating other cultures	0.198	0.243	0.246	0.245	0.177	0.22	0.265	0.236	0.294	0.656	0.23	0.222	0.444	1

The use of latent measurement models to examine teachers' CPL needs

Latent measurement models intend to order persons along one continuum (or dimension) or along multiple continua. A continuum can follow any known mathematical logic to find ordering among persons in two-dimensional space (e.g., a linear line, a circular line, an ellipse, a hyperbolic). Social sciences is dominated by linear continua, however. The phrasing "*latent*" in latent measurement models, highlights that the object of measurement is unobservable. Most psychological constructs cannot be directly observed. Social scientists therefore apply logic. Any applied logic usually uses some combination of multiple observable characteristics to infer between-person differences on the object of measurement. According to Bond and Fox (2007), the estimation of latent measurement scales shares similarities with logic used by Physicians to order substances in terms of Density. The Density of substances is not directly observable. Instead, Physicians use an equation (logic) to derive an estimate of the Density of substances, which involves a division of Volume and Mass. Volume and Mass are both observable characteristics of substances. The estimation of modern latent measurement models starts with drawing an empty continuum. Linear continua typically express between-person differences using the unit: Logits (e.g., Figure 1). The only one model to estimate circular continua routinely expresses between-person differences using the unit: Degrees (Browne, 1992). It will returned on these circular models later.

-inf	...	-3	-2	-1	0	1	2	3	...	+inf
					unit = logits					

Figure 1: An example linear continuum with logits as unit.

Two vectors of estimated parameters typically accompany a continuum and give meaning to it. The first vector contains the item locations as well as other information about the items. Latent measurement models use some predefined logic, expressed in an equation, to order items along the continuum (see Figure 2).

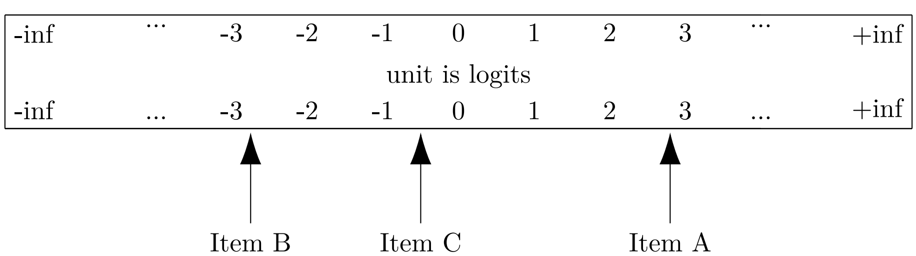


Figure 2: Further elaboration of the example linear continuum, now including a vector of three item locations.

The items assign meaning to different locations along continuum. To illustrate, suppose that some study inquires into persons' favorite color. Suppose that the resulting data set contains a set of dummy variables named by color and identifying whether the person responded that this color is his/her favorite color (1 = *Yes*) or not (0 = *No*). The latent measurement model will order and map the dummy variables to the continuum. Suppose the item locations are ordered as something like: red, orange, yellow, green, blue purple. The first part of the continuum is then the reddish part, meaning that if we estimate persons to be at that position, our estimates are that those persons have a color-preference for red (or perhaps orange). However, when moving further along the continuum, our interpretation of it changes. It no longer means a color-preference for reddish, but changes in into yellow and greenish and eventually into blueish. The central message here is that, although there is a single continuum, it must not have a single interpretation.

The second vector contains the ordering of persons and could possibly include other information about persons. Latent measurement models use some predefined logic, expressed in an equation, to order persons (in our case teachers) along the continuum. The position of the person can match with certain location of an item which match provides information about the person (in our case the teacher) (see Figure 3).

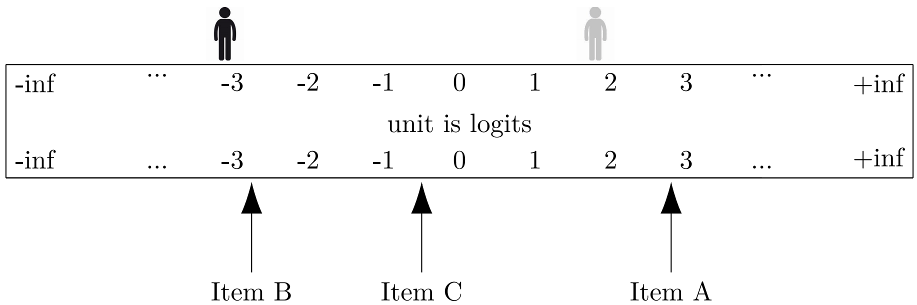


Figure 3: Yet further elaboration of the example linear continuum, now including a vector of two person positions.

When persons are ordered the difference between persons is expressed in by score-difference. In general, the size of the score difference signals how much two persons differ. However, the exact meaning of a score-difference varies between latent measurement models. Guttman (1954) introduced a broad distinction between "differences in degree models" and "differences in kind models". According to Guttman, people can describe differences between one another using phrases that one is "better" or has "more of something" than another person. This way to express between-person differences he referred to as following the logic of "differences in degree". Here it is also referred to this logic by dominance models. In dominance models persons either dominate the item or are dominated by the item. In dominance "differences in degree" type of measurement models, the person's position indicates that this person is able to perform all items located left to his/her position. In Figure 4, the black icon teacher is estimated as not yet able to perform well on whatever the content is of item B, C, and A (note that the model estimates that this teacher is almost able to perform well on item B (whatever its content is).

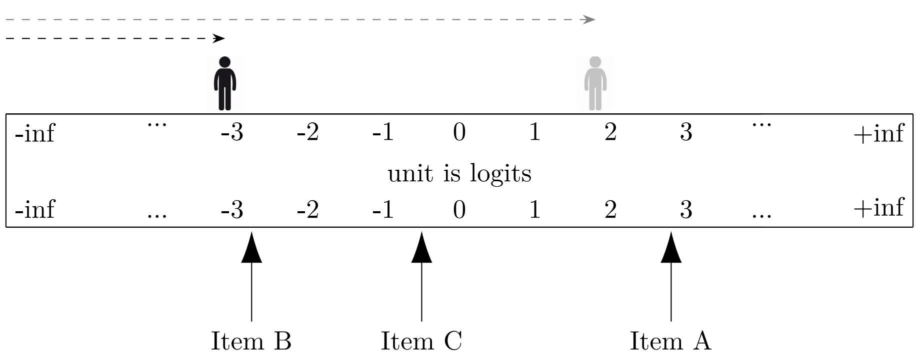


Figure 4: The dominance "difference in degree" models interpretation: person dominates all items up to his/her estimated position.

The teacher positioned at the gray icon position, instead, is estimated as able to perform well on whatever is the content of item B and C, but as not yet able to perform on whatever the content is of item C. It is clear that persons positioned near the gray icon are "better", or have "more of something", than persons positioned near the black icon, because persons positioned near the gray icon likely perform well on all content that the persons positioned near the black icon can do *plus* some additional content. Examples of latent measurement models that apply the “differences in degree” type of logic are: Item Response Theory, Mokken analysis and to a lesser extend Factor Analysis. Guttman (1954) further proposed that people also can decribe difference between one another using phrasings like on being “distinct” or “unalike” another person. This way to express between-person differences he referred to as following the logic of “differences in kind”. Others have referred to this type of logic as “unfolding” (Roberts, Donoghue, Service, and Laughlin, 1996). In unfolding “differences in kind” models, the person position indicates some area along the continuum with of item content that typify the person well or which best describe that person (e.g., Figure 5). There is, thus, no claim that the person dominates or is able to perform on items left of the person’s position. Examples of unfolding latent measurement models are the generalized graded unfolding models (GGUM) and circumplex models.

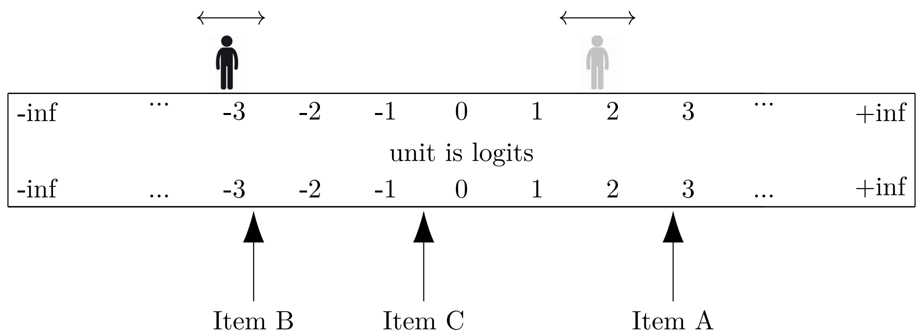


Figure 5: The unfolding "difference in kind" models interpretation: person most likely endorses the item content nearest to his/her estimated position.

Dependent on the latent measurement model chosen, score-differences can express both dominance “differences in degree” (how some persons are better or have more of something than other persons) or they can express unfolding “differences in kind” (how some persons are distinct or unalike) type of individual differences. Importantly, scientific constructs are not necessarily bounded to one type of logic. For some very well known constructs, the dominance “differences in degree” type of logic and the unfolding “differences in kind” type of logic are both applied and can be both considered to be informative. Take for example the Physicians measurement of the construct *Time*. *Time* can be measured using the linear dominance “difference in degree” type of logic - in which running 15 minutes is *longer* or *more* than running 5 minutes - and by using the unfolding “difference in kind” logic - in which 5PM is a distinctive moment of the day than 2AM.

It follows from the above that we may examine the same item responses (item 25) using both types of logic. The results for each serves distinct purposes. Kyriakides, Creemers, and Panayiotou (2018) have used the terminology of “quantitative” versus “qualitative” to refer to these distinct measurement purposes. With “quantitative” measures Kyriakides et al. (2018) refers to measures that describe individual differences between teachers in terms of frequencies such as whether some teachers have *more* or *higher* CPL needs than other teachers. With “qualitative” measures, as we understand, Kyriakides et al. (2018) refers to descriptive measures of individual differences other than frequencies such as whether some teachers have *distinct types* of CPL needs.

Looking at the current evidence, TALIS 2008, 2013, 2018 mostly describe individual differences in terms of their quantity. Central to prior studies was the exploration whether teachers in different countries had different levels of CPL needs (Economic Co-operation and (OECD), 2019). To answer this question, the TALIS questionnaire inquired teachers’ level of need with respect to various areas for professional learning. It clustered areas using factor analysis logical criterion of parallelism (i.e. if higher scores on area A typically coincide with higher scores on area B, than A and B form a cluster). Using this criterion, TALIS 2018 extracted two factors, namely: “*Need for professional development in subject matter and pedagogy*” and “*Need for professional development for teaching for diversity*”. Finally, TALIS published results report on the teachers’ score-differences describing how some teachers have an overall *higher* need for professional learning on these factors than other teachers.

A simple non-latent dominance model: Frequencies in CPL needs

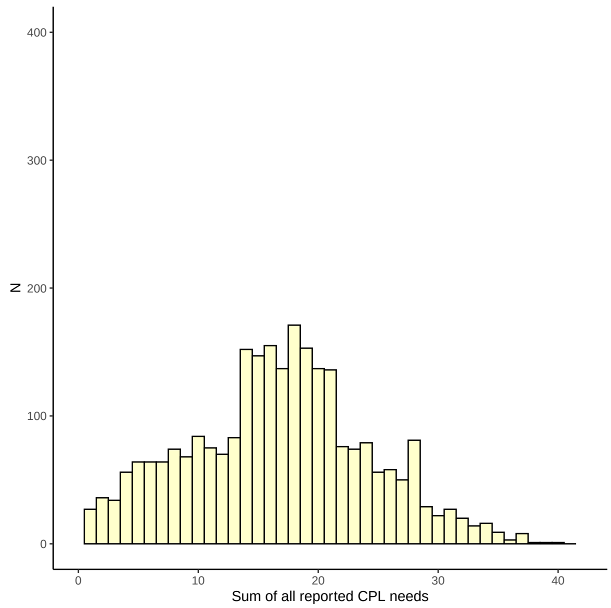
The most simplistic manner to describe individual differences in terms of dominance “differences in degree” is by summing or averaging over all responses. This gives an estimate of the average need to learn on any area. Below we plot the sum scores to describe variation between Australian, Dutch, English, and Japanese teachers in terms of the quantity of their learning needs teachers (see Figure 6 on next page).

The TALIS data shows that teachers vary in the quantity of CPL needs (see Figure 2). Some teachers expressing no high need for none of the areas of CPL, others expressing high need only to some specific areas of CPL, and yet others express need to almost all areas of CPL. Moreover, the Australian, Dutch, and English teacher population have fairly similar distributions, characterized by a normal curve suggesting that most teachers have a overall moderate level of CPL need. The Japanese teacher populations is characterized by a left skewed distribution, suggesting that most Japanese teachers have a overall high level of learning need. In summary, the use of the simple frequency statistic would result in us assessing that the Japanese teachers typically have a high level of CPL need, whereas the teacher populations in the Western countries have mostly have a moderate level of CPL need.

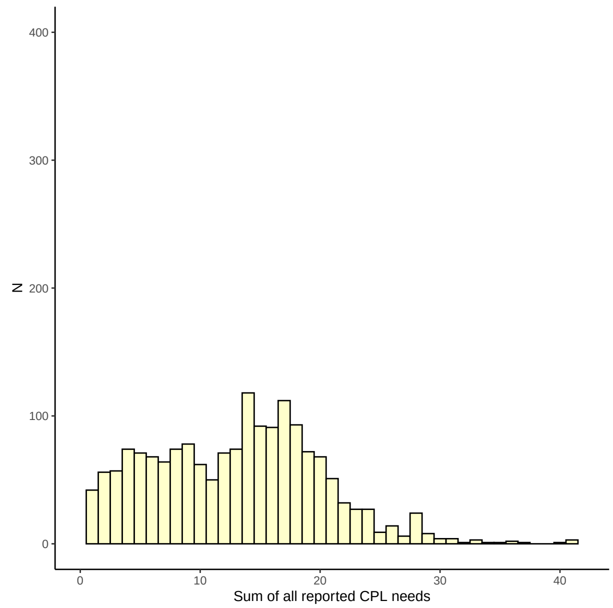
An application of three Latent measurement models: Overview and model fit criteria

This section assesses the value of (1) dominance “differences in degree” models, (2) linear unfolding “differences in kind” models, and (3) circular unfolding “differences in kind” models to describe individual differences in teachers’ CPL needs. The utility of dominance “differences in degree” models is assessed by fitting a Generalized Partial Credit Model (GPCM) (Muraki, 1992) to the data. The utility of linear unfolding “differences in kind” models is assessed by fitten a Generalized Graded Unfolding Model (GGUM) (Roberts et al., 1996) to the data. Finally, the utility of circular unfolding “differences in kind” models is assessed by fitting a circumplex model (Guttman, 1954) to the data. An important consideration for the model choice is model fit. Model fit coefficients describe to extend to which model predicted item responses overlaps with the observed item responses in the data set. The following coefficients for model fit were used:

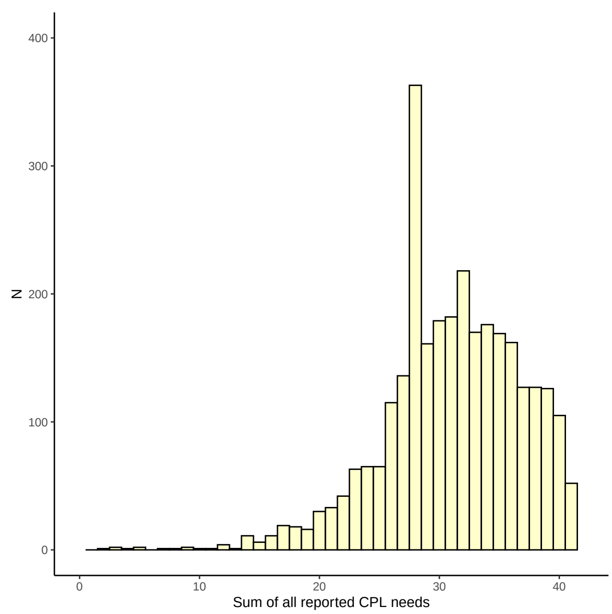
- **Absolute model fit coefficients:**
  - Root Mean Square Error of Approximation (RMSEA), wherein:
    - $RMSEA \leq 0.05$ = good fit
    - $0.05 < RMSEA \leq 0.08$ = sufficient fit
    - $RMSEA \geq 0.08$ = insufficient fit
  - Comparative Fit Index (CFI), wherein:
    - $CFI \geq 0.95$ = good fit
    - $CFI \geq 0.90$ = sufficient fit
    - $CFI < 0.90$ = insufficient fit
- **Comparative model fit coefficients:**
  - $\frac{\chi^2}{df}$ ratio, wherein:
    - **Rule:** $< \frac{\chi^2}{df}$ means better model fit compared to other models
  - The Bayesian Information Criterion, wherein:
    - **Rule:** $< BIC$ means better model fit compared to other models



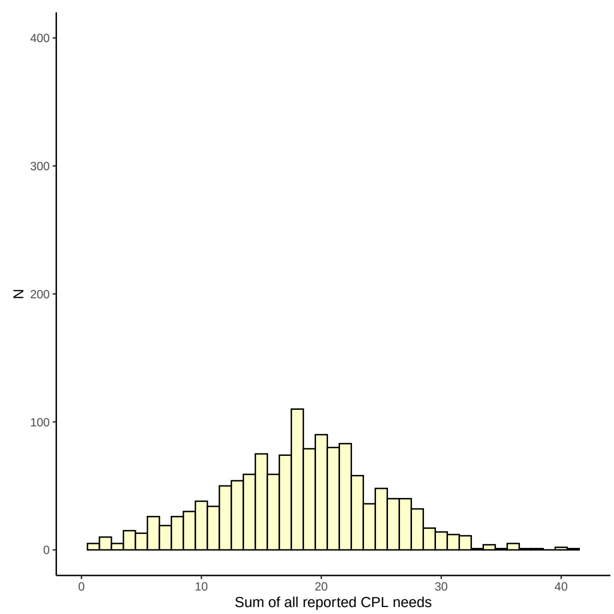
(a) Sumscore CPL needs (AUS)



(b) Sumscore CPL needs (ENG)



(c) Sumscore CPL needs (JPN)



(d) Sumscore CPL needs (NLD)

Figure 6: Differences in quantity of CPL needs

An application of the Generalized Partial Credit Model

Dominance "difference in degree" latent measurement models have advantages over an approach using the mean or frequency scores. Frequency and mean statistics are insensitive to between-person differences on the items. Take for example the above histograms. These histograms suggest huge similarities between the Australian, the Dutch and the English teacher populations. However, when we would look at the exact item responses, than it can be observed that Dutch teachers are much more likely to assign a high level of need to one of the items, while assigning a low level of need to other items. In specific, 739 or 54% of the Dutch teachers indicates a high level of needs on one of the items, whereas responses by the English teachers indicate that only 348 or 20% of the English teachers has a high level of needs on one of the items. In summary, the frequencies above hide the considerable difference between Dutch and English teacher populations. Dutch teachers discriminate more precisely between what they feel they need to learn and what they feel they don't need to learn whereas the Australian and English teachers do not discriminate that much but generally assign the same level of need (moderate) to all items. It are these types of differences that can become visible in latent measurement models. We will now turn to discussing the models and their results.

To explore the applicability of dominance "difference in degree" measurement, the GPCM was fitted to the data (Muraki, 1992). The GPCM is a polytomous two-parameter item response theory (IRT) model. IRT models predict that the probability on a positive (higher) item response increases when persons are *better* or *have more* of the object of measurement (see Figure 7). IRT models order items along one continuum using the prediction that the probability on a positive response on an item located more to the right of the continuum increases conditionally on the probability on a positive response on items located more to left of the continuum. Bond and Fox (2007) refer to this conditional probability as following a linear-additive logic. It predicts that teachers' level of CPL need increases additive such that teachers' with a low level of CPL need endorsed only some few items (e.g., item A, E, and G), teachers with an overall moderate level of CPL need are predicted to endorse those same few items plus some additional items (e.g., item A, E, and G + item F and B), and teacher having a high level of CPL need endorse those same few items, those same additional items plus even more items (e.g., item A, E, G, F and B plus item C and D). Conceptually, an application of dominance "difference in degree" models thus implies an underlying theoretical logic explaining how a need for certain CPL activities must precede a need for other CPL activities.

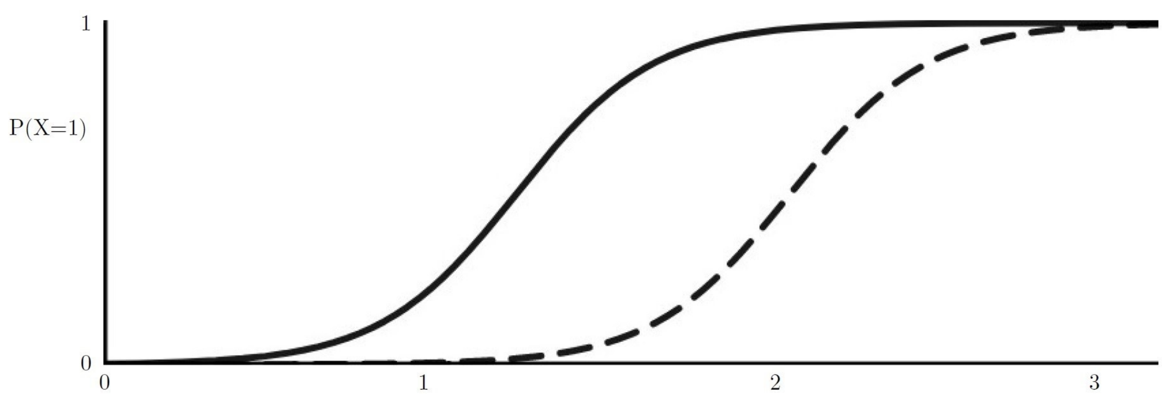


Figure 7: A visualization of the linear dominance GPCM probability curve

Polytomous IRT models introduce an additional complexity because they subdivide item responses into item steps. An item step conceptualizes the stepping of a lower response category k to the higher response category k+1. The Figure 8 below illustrates the three item steps found in this study's four-point Likert scale. Item step are valued in logits. The item step value is our best estimate of the required increase in the level of teachers' CPL needs to reach the point where the teacher has an probability >50% to mark the one-step higher response category (k+1).Thus, suppose the teacher has responded the Likert scale value 1 = "*no need at present*" and the item step value is 0.79, then the polytomous IRT model predicts that an increase of 0.79 points in the teachers' level of CPL needs will make that the teacher is more likely to mark the higher response option value 2 = "*low level of need*" than the response option currently marked.

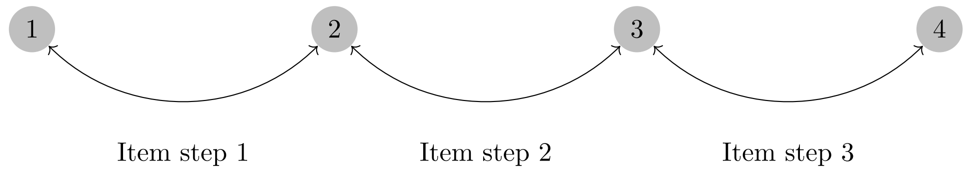


Figure 8: A visualization of the concept item step

```
PCM.AUS <- mirt::mirt(na.omit(CPL.needs.AUS[, 3:16])-1, 1, itemtype = 'gpcm',
                      IRT.param = TRUE, technical = list(NCYCLES = 2000))
PCM.ENG <- mirt::mirt(na.omit(CPL.needs.ENG[, 3:16])-1, 1, itemtype = 'gpcm',
                      IRT.param = TRUE, technical = list(NCYCLES = 2000))
PCM.JPN <- mirt::mirt(na.omit(CPL.needs.JPN[, 3:16])-1, 1, itemtype = 'gpcm',
                      IRT.param = TRUE, technical = list(NCYCLES = 2000))
PCM.NLD <- mirt::mirt(na.omit(CPL.needs.NLD[, 3:16])-1, 1, itemtype = 'gpcm',
                      IRT.param = TRUE, technical = list(NCYCLES = 2000))
# fit coefficients GGUM
fit.PCM.AUS <- mirt::M2(PCM.AUS)
fit.PCM.ENG <- mirt::M2(PCM.ENG)
fit.PCM.JPN <- mirt::M2(PCM.JPN)
fit.PCM.NLD <- mirt::M2(PCM.NLD)
```

GPCM: Model fit

The GPCM was estimated using the "mirt" package (Chalmers, 2012). Syntax ran in R is printed above. Model fit coefficients were insufficient in absolute terms. RMSEA for the four countries in alphabetical order were: 0.141, 0.105, 0.105, and 0.099. CFI for the four countries in alphabetical ordere were: 0.86, 0.93, 0.88, and 0.86. Because this is the first measurement model explored, conclusions on the relative fit are found in subsequent sections. BIC values were 74597.9, 45611.7, 68209.9, 42731. $\frac{\chi^2}{df}$ values ranged between: 53.77, 20.63, 35.86, and 14.61. An overview of fit coefficients is given in the Table below.

Table 9: Fit coefficients for PCM model

	CFI	RMSEA	RMSEA 5%	RMSEA 95%	Chi	df	Chi/df	BIC
Austrlia	0.86	0.141	0.136	0.146	2634.718	49	53.77	74597.9
England	0.93	0.105	0.100	0.111	1010.920	49	20.63	45611.7
Japan	0.88	0.105	0.101	0.109	1757.006	49	35.86	68209.9
Netherlands	0.86	0.099	0.093	0.106	715.978	49	14.61	42731.0

GPCM: Interpretation output

It might be possible to keep a subset of items that fit the GPCM. Item selection is a time consuming process however. Whether such effort might be worthwhile depends on the utility that the GPCM has for identifying individual teachers' CPL needs. In what follows we briefly explore the utility of the results estimated by the current GPCM as serving as a thought experiment ("what if this model would have fitted?"). This thought experiment is performed by analyzing a WrightMap plot. Central to WrightMap plots is the continuum which is vertically plotted. The continuum starts at the top and runs to the bottom. Left side of the vertical continuum are the person positions plotted. Right of the vertical continuum are the item locations. Following the logic of the GPCM, teachers located at the top of the continuum have high probability to mark the item response category "no need at present". Hence, these teachers are estimated to have no or very low levels of CPL need on all or most of the areas. Teachers at the bottom of the continuum have high probability to mark the item response category "high level of need". Hence, these teachers are estimate to have a high level of CPL need on all or most areas. Most of interest is whether areas vary over the continuum. Variation would indicate that, although teachers have low or moderate levels of CPL need overall, the probability that they mark item responses is not uniformly distributed over the areas.

```
person.scores.AUS <- mirt::fscores(PCM.AUS)

## Make data.frame object to be inserted in the function wrightMap()
item.parameters.AUS <- as.data.frame(mirt::coef(PCM.AUS,IRTpars = T,
                                simplify = T)[[1]][, 2:4])
item.parameters.AUS <- item.parameters.AUS - rowMeans(mirt::coef(
                        PCM.AUS,IRTpars = T, simplify = T)[[1]][, 2:4])
rownames(item.parameters.AUS) <- c("item A","item B","item C","item D","item E",
                                   "item F","item G","item H","item I","item J",
                                   "Item K","item L","item M","item N")
```

```
# Coloring items in functioning wrightMap()
itemcolors.wrightmap <- c("red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4","red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4","red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4")

# syntax for plotting Wrightmap
WrightMap.AUS <- WrightMap::wrightMap(person.scores.AUS, item.parameters.AUS,
          person.side = personDens, show.thr.lab = FALSE,
          main.title = "Wrightmap GPCM Teachers' CPL needs Australia", min.l = -6,
          max.l = 4, item.prop = 0.8, thr.sym.col.fg = itemcolors.wrightmap)
```

Below we print the syntax to extract the teacher positions and the item locations from the mirt-object "PCM.AUS". Under that code chunk follows the code to generate a Wrightmap plot using the r package "WirghtMap". The wrightmaps for the GPCM are printed below (see Figure 9 and 10). On the x-axis are the 14 items "A" to "N". The three vertically plotted dots above the items are our estimates of the item-step locations. The highest dot is the first step. Teacher positioned at the location of the first item step have a probability of 50% to choose the higher response category of the competing categories "no need at present" and a "low level of need", the further teachers position is beyond the item location (thus the lower teachers are positioned at the continuum), the higher the probability is that teachers endorse the higher response category out of these two categories. It can be observed that the GPCM estimates equal distances between the first the second and the third item steps. This result suggests that taking an item step (meaning that the teacher likely will endorse the higher response category $k + 1$) requires a similar increase in teachers' level of CPL need.

What is further apparent from the Wrightmaps is that the GPCM results are ambiguous, such that they do not match unique teacher positions with unique areas in need for CPL. When moving top-down along the continuum then there are three points at which the model predicts that the teacher will take an item step. However, the model predicts that the teacher takes the item step for all items simultaneously. This is observed in all countries.

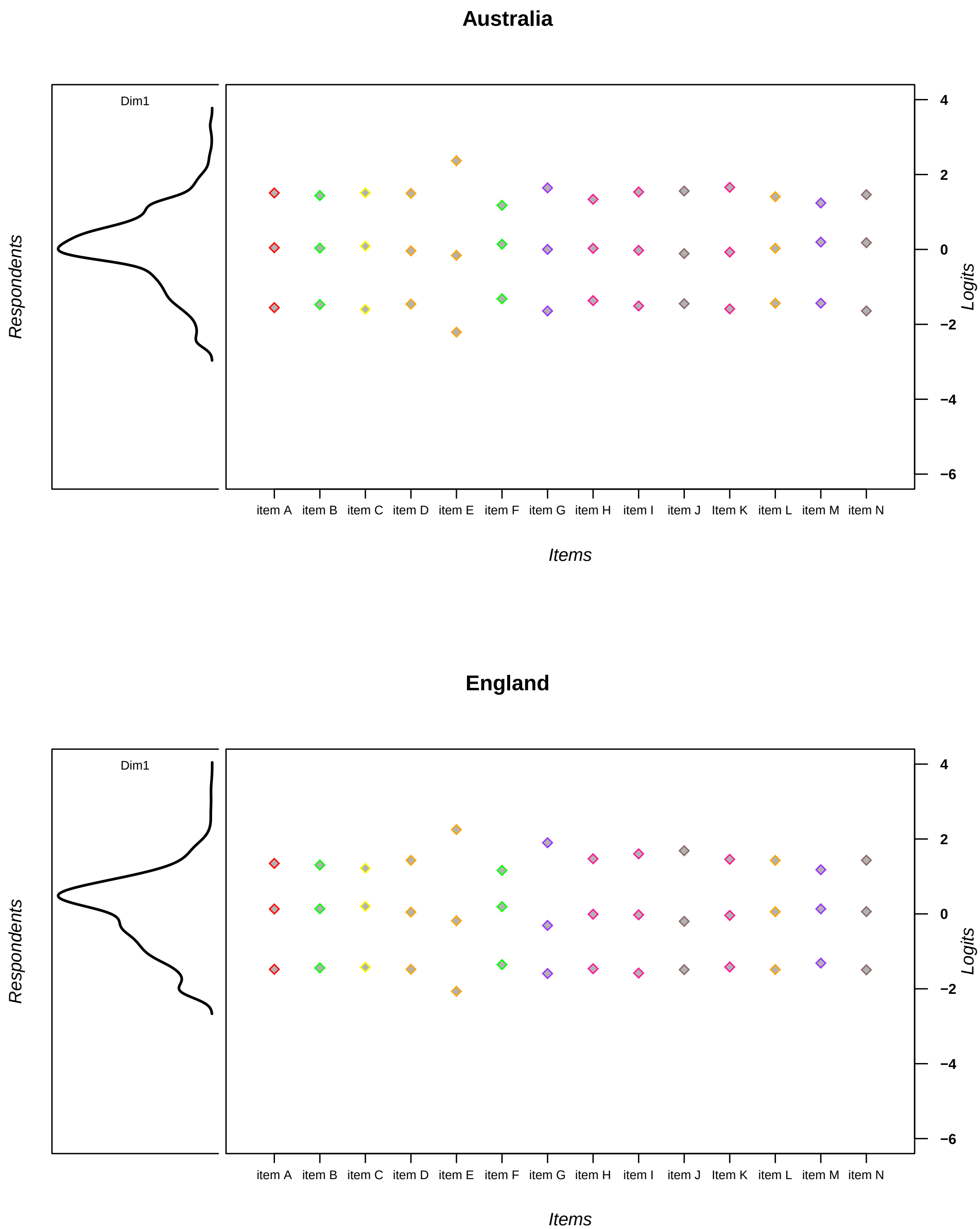


Figure 9: wrightmap GPCM (Australia left - England right plot)

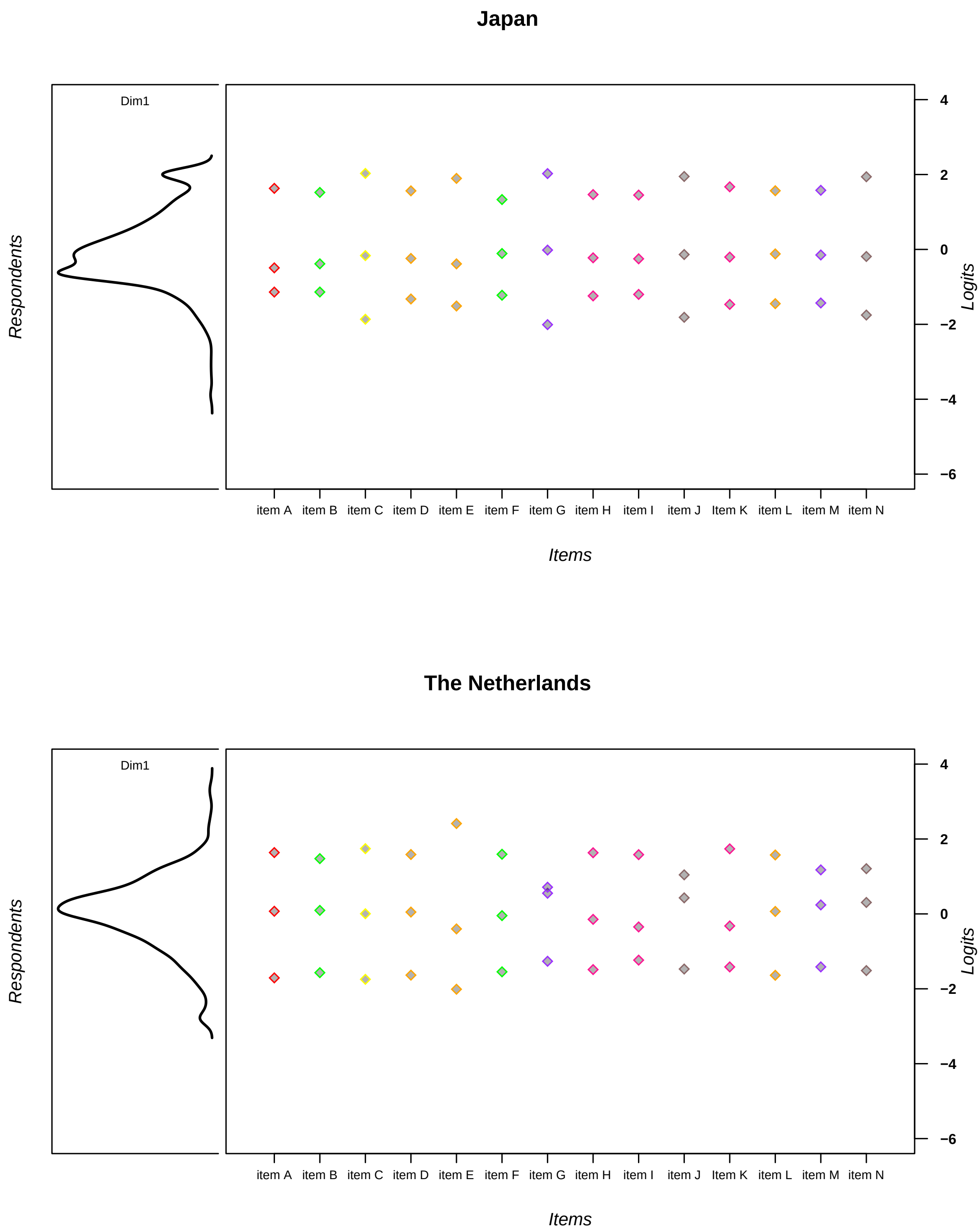


Figure 10: wrightmap GPCM (Japan above - the Netherlands below plot)

The Generalized Graded Unfolding Model (GGUM)

The Generalized Graded Unfolding Model (GGUM) was fitted to the data to test the applicability of the linear unfolding measurement. Unfolding models have the purpose to characterize or typify persons. Conceptually it believes that the probability that teachers endorse a high level of CPL needs peaks for some specific areas (items) but is low for all others. In the Figure 11, the probability peaks at the position of the black triangle. Here, the black triangle represents the location of an item. According to the model estimations, the probability is high that the teacher marks the black triangle item with *a high level of need* while the probability on such mark is lower for the other items. GGUM is a polytomous unfolding model, meaning that like the GPCM it estimates item steps.

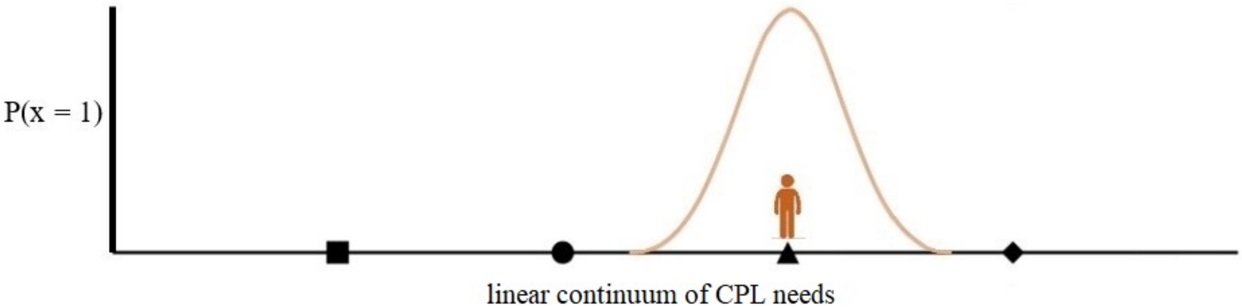


Figure 11: An illustration of the linear unfolding peak probiblity distribution

```
GGUM.AUS <- mirt::mirt(na.omit(CPL.needs.AUS[, 3:16])-1, 1, itemtype = 'ggum',
                       technical = list(NCYCLES = 2000))
GGUM.ENG <- mirt::mirt(na.omit(CPL.needs.ENG[, 3:16])-1, 1, itemtype = 'ggum',
                       technical = list(NCYCLES = 2000))
GGUM.JPN <- mirt::mirt(na.omit(CPL.needs.JPN[, 3:16])-1, 1, itemtype = 'ggum',
                       technical = list(NCYCLES = 2000))
GGUM.NLD <- mirt::mirt(na.omit(CPL.needs.NLD[, 3:16])-1, 1, itemtype = 'ggum',
                       technical = list(NCYCLES = 2000))

# fit coefficients GGUM
fit.GGUM.AUS <- mirt::M2(GGUM.AUS)
fit.GGUM.ENG <- mirt::M2(GGUM.ENG)
fit.GGUM.JPN <- mirt::M2(GGUM.JPN)
fit.GGUM.NLD <- mirt::M2(GGUM.NLD)
```

GGUM: Model fit

The GPCM was estimated using the “mirt” package (Chalmers, 2012).Syntax ran in R is printed. Fit the of the GGUM ranged between sufficient to insufficient in absolute terms. RMSEA for the four countries in alphabetical ordere were: 0.084, 0.096, 0.106, and 0.067. CFI for the four countries in alphabetical ordere were: 0.69, 0.76, 0.96, and 0.93. Because this is the first measurement model explored, conclusions on the relative fit are found in subsequent sections. BIC values were 74715.9, 45715.5, 68280, 42864.9. $\frac{\chi^2}{df}$ values ranged between: 19.6, 17.48, 36.92, and 7.15. An overview of fit coefficients is provided in the Table below.

Table 10: Fit coefficients for GGUM model

	CFI	RMSEA	RMSEA 5%	RMSEA 95%	Chi	df	Chi/df	BIC
Austrlia	0.69	0.084	0.078	0.089	686.051	35	19.60	74715.9
England	0.76	0.096	0.090	0.103	611.691	35	17.48	45715.5
Japan	0.96	0.106	0.101	0.111	1292.044	35	36.92	68280.0
Netherlands	0.93	0.067	0.059	0.075	250.399	35	7.15	42864.9

Overall, model fit coefficients favored the GGUM model over the GPCM, however fit coefficients for the GGUM were inconsistent within countries suggesting no sufficient absolute fit in most countries. In Japan, for example, the RMSEA was with 0.106 highest and indicating 'bad' absolute fit, whereas CFI was 0.69 which is also highest and indicating 'good' absolute fit to the data. Only for the Netherlands did the GGUM give predictions of sufficient fit on both CFI and RMSEA.

GGUM: Interpretation output

Though the GGUMN does not fit the data, it might be possible to keep a subset of items that fit the GGUM. Item selection is a time consuming process however. Whether such effort might be worthwhile depends on the utility that the GGUM has for identifying individual teachers' CPL needs. In what follows we briefly examine the utility of the results estimated by the current GGUM as being a though experiment ("what if the GGUM would have fitted?"). This though experiment is performed using a WrightMapplot. The Wirghtmap uses a visualization analogous to Figure 1a to 1f above). Central to these figures is the dimension with on the upper side the person positions and on the lower side the item locations. In a Wrightmap, the person positions are at the left side of the dimension and the item locations are at the right side of it. Following the logic of the GGUM, teachers located near an item (or an item step) are most likely to endorse that item while not endorsing items located at other locations on the continuum.

```r
person.scores.AUS <- mirt::fscores(GGUM.AUS)

## Make data.frame object to be inserted in the function wrightMap()
item.parameters.AUS <- as.data.frame(rbind(mirt::coef(GGUM.AUS)[[1]][3:5],
        mirt::coef(GGUM.AUS)[[2]][3:5], mirt::coef(GGUM.AUS)[[3]][3:5],
        mirt::coef(GGUM.AUS)[[4]][3:5], mirt::coef(GGUM.AUS)[[5]][3:5],
        mirt::coef(GGUM.AUS)[[6]][3:5], mirt::coef(GGUM.AUS)[[7]][3:5],
        mirt::coef(GGUM.AUS)[[8]][3:5], mirt::coef(GGUM.AUS)[[9]][3:5],
        mirt::coef(GGUM.AUS)[[10]][3:5], mirt::coef(GGUM.AUS)[[11]][3:5],
        mirt::coef(GGUM.AUS)[[12]][3:5], mirt::coef(GGUM.AUS)[[13]][3:5],
        mirt::coef(GGUM.AUS)[[14]][3:5]))
rownames(item.parameters.AUS) <- c("item A","item B","item C","item D","item E",
                                   "item F","item G","item H","item I","item J",
                                   "Item K","item L","item M","item N")

# Coloring items in functioning wrightMap()
itemcolors.wrightmap <- c("red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4","red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4","red","green","yellow","orange","orange","green",
      "Purple1", "DeepPink1","DeepPink1","RosyBrown4","DeepPink1","orange",
      "Purple1","RosyBrown4")
```

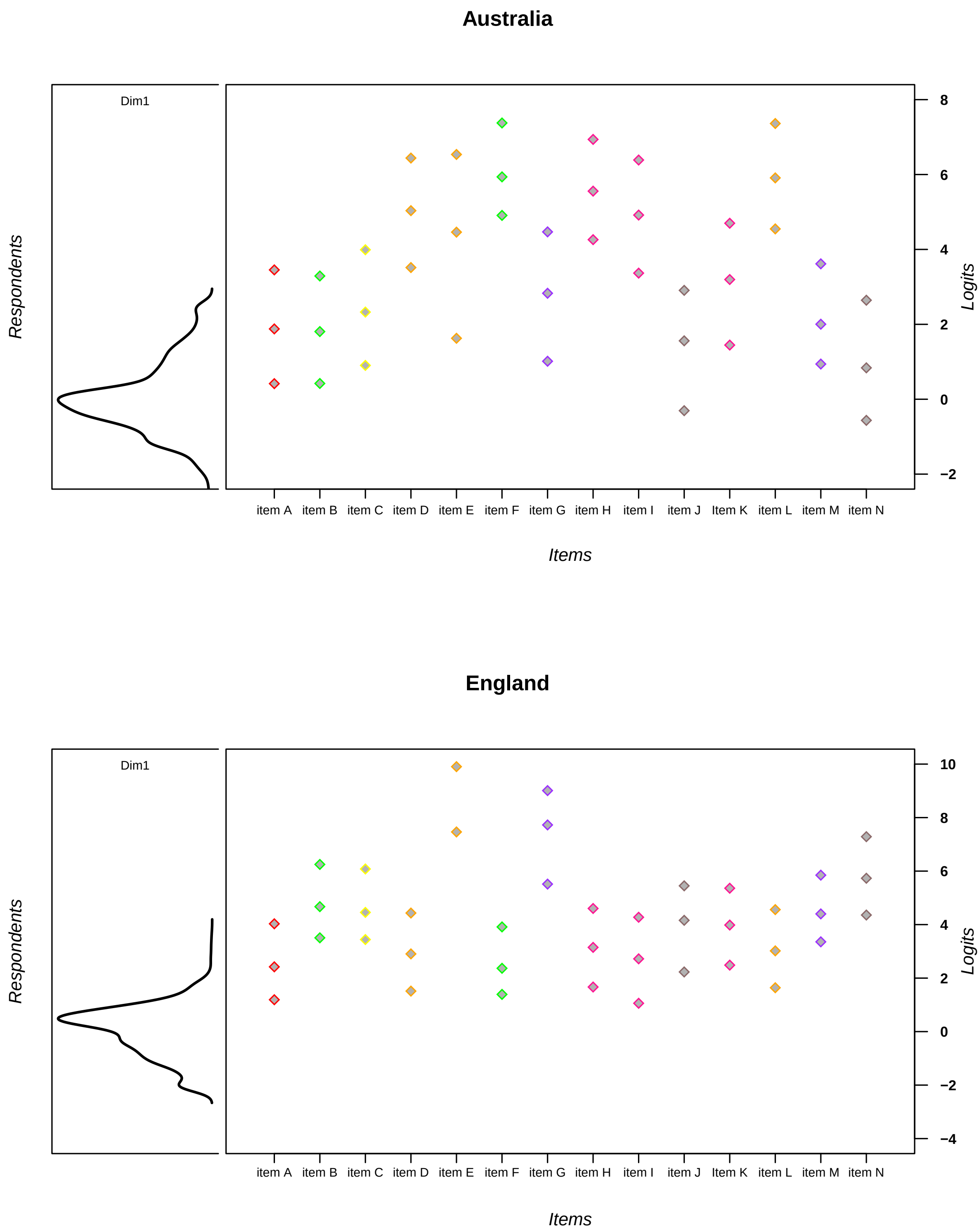
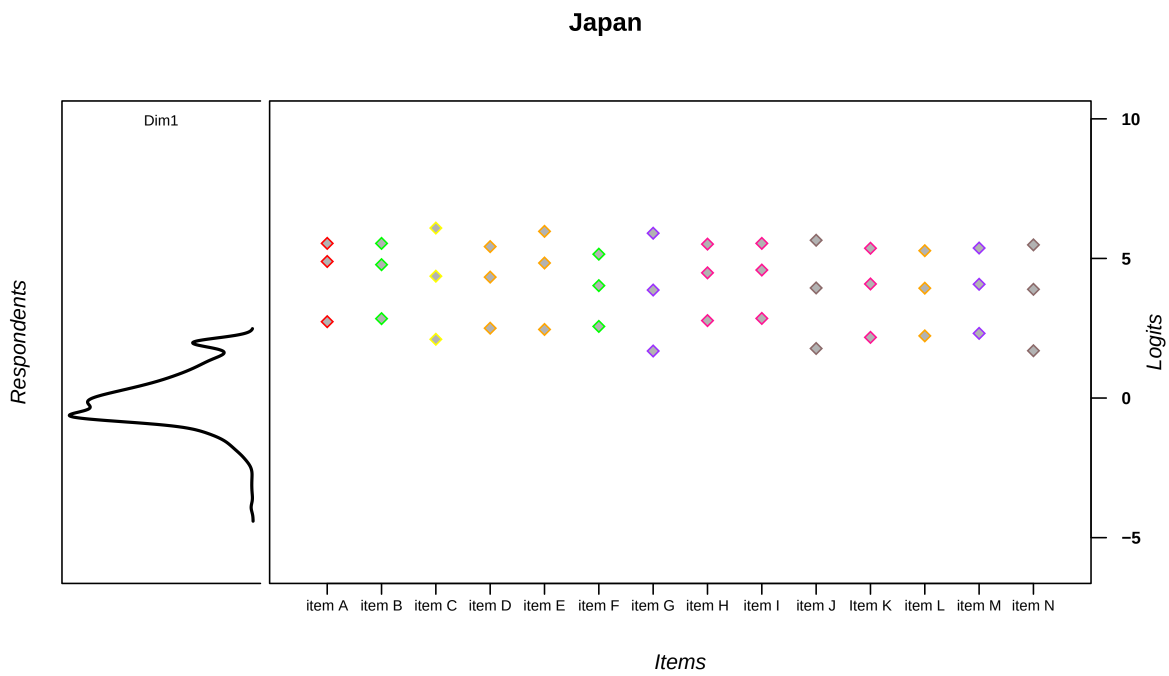


Figure 12: wrightmap GGUM (Australia left - England right plot)



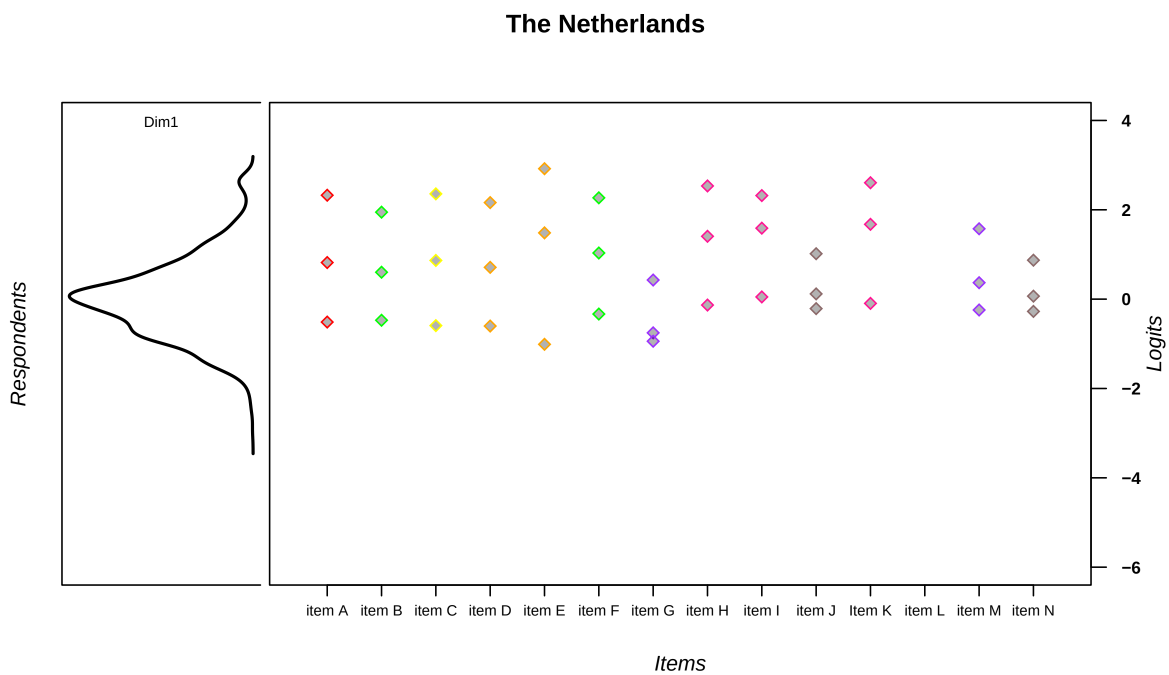


Figure 13: wrightmap GGUM (Japan left - the Netherlands right plot)

Below we print the syntax to extract the teacher positions and the item locations from the mirt-object "GGUM.AUS". Under that code chunk follows the code to generate a Wrightmap plot using the r package "WirghtMap". The wrightmaps for the GGUM are presented in Figure 11 and 12. On the x-axis are the 14 items "A" to "N". The three vertically plotted dots above each other represent the item-step locations. The highest dot is the first step. Teachers positioned at the location of the first item step have a low probability to endorse the higher response category of the competing categories $k =$ "no need at present" and $k + 1 =$ "low level of need". The further teachers position is below the item step location, the higher the probability is that teachers marks the higher response category $k + 1$.

Alike the GPCM, the GGUM also estimates equal distances between the first the second and the third item steps. This result suggests that if CPL needs increase, the probability that the teachers marks the response category $k + 1$ increases uniformly for all the items. This model behavior is not useful for the purpose to identify unique CPL needs. Furthermore, and again similar to the GPCM, the GGUM results are ambiguous. It does not match unique teacher positions with unique areas in need for CPL. When moving top-down along the continuum then there are three points at which the model predicts that the teacher will take an item step. However, the model predicts that the teacher takes the item step for all items simultaneously. This is observed in all countries.

The circumplex model

In the article, the circumplex is illustrated in 2D space, due to which it appears flat. The actual circumplex, however, must be viewed as a cylinder having the height equal to the number of response categories on the Likert-scale. Predicted Likert-scale responses increase if item content is more associated with the teachers' attitudes or in our case needs. The red dashed line in the Figure 14 reflects how the probability that the teacher marks the response category *high level of need* is highest for items 7 and 8 and lowest for items 3 and 4, whatever the content of these items may be. Note that the circumplex uses the correlation metric and not the probability metric. Though technically different, we argue that conceptually the logic just stated can be applied.

Some further explanation about the circumplex model can be given using this cylindrical Figure. First, the Figure illustrates an ideal in which the probability increase uniformly when moving from items 3 and 4 to items 7 and 8. This ideal is almost never present. Browne (1992) introduced the m parameter, to model steeper and flatter parts of the increase in predicted Likert-scale responses. By default m is set 3. This study also applied this default. The number 3 is related to the three factor model which parameters are used as starting input to estimate the circumplex model. The rationale is that items belonging to the same factor must have been responded to more similar and, therefore, will cluster along the circumference. Among items within the same cluster the increase in predicted Likert-scale responses likely is flatter.

In the Figure, the upward distances from response category "*no need at present*" to *"low level of need"* to *"moderate level of need"* and finally *"high level of need"* are all equal. This illustrates that the circumplex assumes that the Likert-scale responses are continuous scales with equal step sizes. This assumption is related to the standard input Pearson correlation matrix. Pearson correlations assume two continuous scales. The assumption that Likert scales are continous is uncommon in other psychometric models, like factor analysis, the unfolding IRT models and the Partial Credit Model (PCM), however. These latter models typically define ordinal item thresholds or ordinal item step functions. Moreover, the ordinal treatment of Likert scale data generally improves model fit. Hence, a truly fair comparison of model fit between the circumplex model and other psychometric models, would require an estimate of circumplex model fit defining the item Likert-scales as ordinal. As far as we know, this has never been attempted before. The model fit indices presented in the article depend on a new approach and estimation strategy that defines the Likert-scales as ordinal. The script for this strategy is found below the Figure.

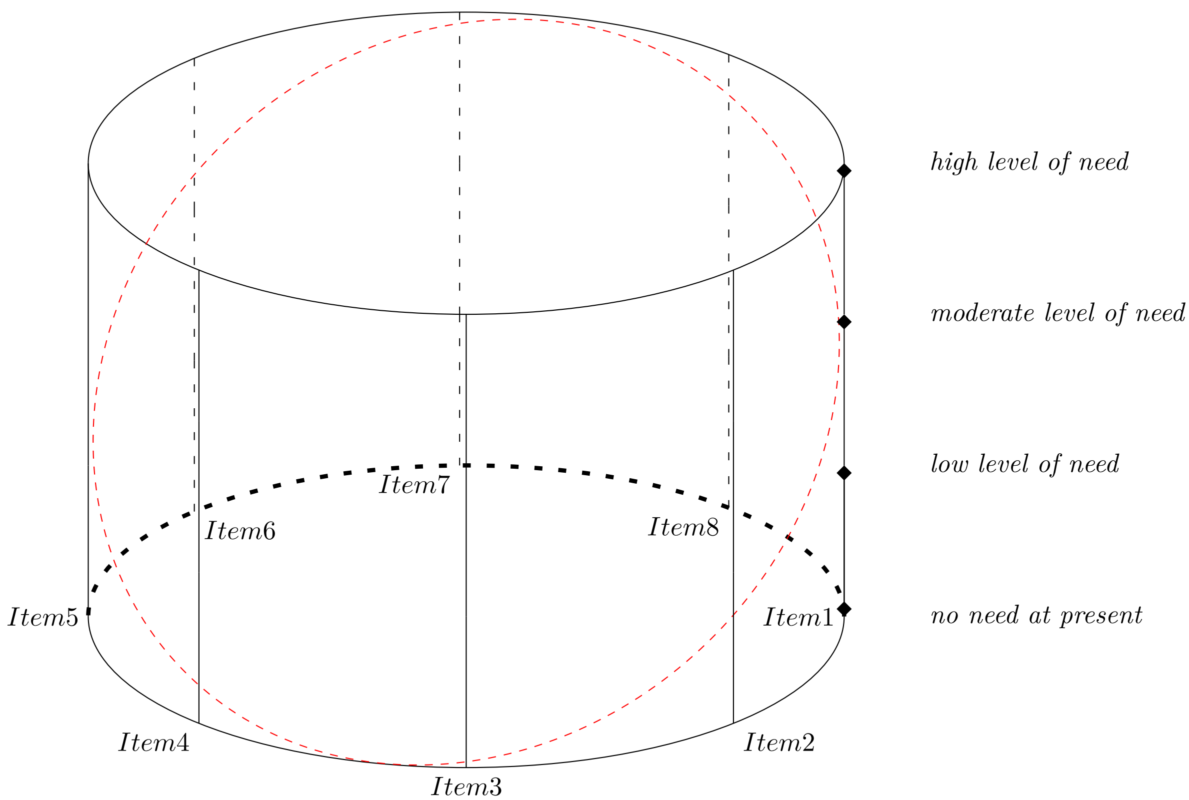


Figure 14: 3D visualization of the cirumplex model. The Likert scale with four levels places four circumplexes above each other together making a cylinder

Item steps coding One of the key concerns with the circumplex is that is assumes Likert scales to be continous interval level scales. The coding below was developed to code item responses in line with item step coding. Item step coding is different from item step estimation. Also, the item step coding is not performed analogously to Sijtsma and Molenaar (2002), because the standard Mokkken item step coding was developed for the dominance model type of logic. Instead, the item step coding was developed such that it would 'cut' slices from the above cylinder for each level of the Likert scale. An example of the R syntax used to code item steps is presented below.

```
# Three empty vectors for item A step codes
CPL.needs.A.step1 <- rep(NA, nrow(CPL.needs.AUS))
CPL.needs.A.step2 <- rep(NA, nrow(CPL.needs.AUS))
CPL.needs.A.step3 <- rep(NA, nrow(CPL.needs.AUS))

# item A step code 1
for(i in 1:nrow(CPL.needs.AUS)) {
  if((CPL.needs.AUS[i,3]-1) == 0 & !is.na(CPL.needs.AUS[i,3]-1)) {
         CPL.needs.A.step1[i] <- 0 }
  else if((CPL.needs.AUS[i,3]-1) == 1 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step1[i] <- 1  }
  else if((CPL.needs.AUS[i,3]-1) == 2 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step1[i] <- 0  }
  else if((CPL.needs.AUS[i,3]-1) == 3 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step1[i] <- 0  }
}
```

```
# item A step code 2
for(i in 1:nrow(CPL.needs.AUS)) {
  if((CPL.needs.AUS[i,3]-1) == 0 & !is.na(CPL.needs.AUS[i,3]-1)) {
         CPL.needs.A.step2[i] <- 0 }
  else if((CPL.needs.AUS[i,3]-1) == 1 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step2[i] <- 0  }
  else if((CPL.needs.AUS[i,3]-1) == 2 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step2[i] <- 1  }
  else if((CPL.needs.AUS[i,3]-1) == 3 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step2[i] <- 0  }
}
# item A step code 3
for(i in 1:nrow(CPL.needs.AUS)) {
  if((CPL.needs.AUS[i,3]-1) == 0 & !is.na(CPL.needs.AUS[i,3]-1)) {
         CPL.needs.A.step3[i] <- 0 }
  else if((CPL.needs.AUS[i,3]-1) == 1 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step3[i] <- 0  }
  else if((CPL.needs.AUS[i,3]-1) == 2 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step3[i] <- 0  }
  else if((CPL.needs.AUS[i,3]-1) == 3 & !is.na(CPL.needs.AUS[i,3]-1))  {
         CPL.needs.A.step3[i] <- 1  }
}

# Three empty vectors for item B step codes
CPL.needs.B.step1 <- rep(NA, nrow(CPL.needs.AUS))
CPL.needs.B.step2 <- rep(NA, nrow(CPL.needs.AUS))
CPL.needs.B.step3 <- rep(NA, nrow(CPL.needs.AUS))

# item B step code 1
for(i in 1:nrow(CPL.needs.AUS)) {
  if((CPL.needs.AUS[i,4]-1) == 0 & !is.na(CPL.needs.AUS[i,4]-1)) {
         CPL.needs.B.step1[i] <- 0 }
  else if((CPL.needs.AUS[i,4]-1) == 1 & !is.na(CPL.needs.AUS[i,4]-1))  {
         CPL.needs.B.step1[i] <- 1  }
  else if((CPL.needs.AUS[i,4]-1) == 2 & !is.na(CPL.needs.AUS[i,4]-1))  {
         CPL.needs.B.step1[i] <- 0  }
  else if((CPL.needs.AUS[i,4]-1) == 3 & !is.na(CPL.needs.AUS[i,4]-1))  {
         CPL.needs.B.step1[i] <- 0  }
}

# code will repeat for every item B till N.
```

pooling correlation matrices After computation of the item steps, three correlation matrices were estimated one for every step code matrix. Then, all three correlation matrices were pooled into one correlation matrix using Fisher's Z-transformation. An example code follows below.

```
# define the variable ID school as factor
# the object Item.step.1 is a data.frame object with item step codes for
# item step 1
Item.step.1.AUS$IDSCHOOL <- as.factor(Item.step.1.AUS$IDSCHOOL)

# Impute missing data in item step 1 Australian data as example
# this step is repeated three times for every Item step
```

```
mice_imp <- mice(select(Item.step.1.AUS,TT3G27A,TT3G27B,
                  TT3G27C,TT3G27D,TT3G27E,TT3G27F,TT3G27G,TT3G27H,
                  TT3G27I,TT3G27J,TT3G27K,TT3G27L,TT3G27M,TT3G27N),
                 m = 40, seed = 987)

# generate the 40 different data sets with impute missing values
# this step is repeated three times for every step
Item.step.1.AUS_imp <- NULL
for(i in 1:40) Item.step.1.AUS_imp[[i]] <- cbind(as.factor(
  Item.step.1.AUS$IDSCHOOL),complete(mice_imp, action=i, inc=FALSE))
for(i in 1:40) names(Item.step.1.AUS_imp[[i]]) <-
  make.names(names(Item.step.1.AUS[,c(2:16)]))

# Perform a Fisher Z-transformation on the estimated correlations
# this step is repeated three times for every step
Zcor.Item.step.1.AUS <- DescTools::FisherZ(cbind(
  correlation(Item.step.1.AUS_imp[[1]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[2]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[3]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[4]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[5]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[6]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[7]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[8]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[9]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[10]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[11]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[12]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[13]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[14]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[15]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[16]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[17]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[18]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[19]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[20]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[21]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[22]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[23]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[24]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[25]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[26]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[27]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[28]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[29]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[30]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[31]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[32]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[33]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[34]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[35]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[36]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[37]], multilevel = FALSE)[, 3],
```

```r
      correlation(Item.step.1.AUS_imp[[38]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[39]], multilevel = FALSE)[, 3],
      correlation(Item.step.1.AUS_imp[[40]], multilevel = FALSE)[, 3]))

# Compute mean Fisher Z-transformed across the 40 imputed datasets
# Inverse Fisher's Z to retrieve Pearson correlations
# performed three times
FisherInv.Item.step.1_imp.AUS <- FisherZInv(rowMeans(Zcor.Item.step.1.AUS))

# organize the Pearson correlations in a diagonal matrix for printing
# performed three times
n_elements <- 14
m <- diag(n_elements)
m_upper <- m_lower <- matrix(0, n_elements, n_elements)
m[lower.tri(m_lower, diag = FALSE)] <- FisherInv.Item.step.1_imp.AUS
m[upper.tri(m)] <- NA
Cor.Item.step.1.AUS_imp <- round(m, 3)
Cor.Item.step.1.AUS_imp[is.na(Cor.Item.step.1.AUS_imp)] <- 0
Cor.Item.step.1.AUS_imp <- Cor.Item.step.1.AUS_imp + (t(
  Cor.Item.step.1.AUS_imp) - diag(diag(Cor.Item.step.1.AUS_imp)))

# Z-Transform correlations -- three times performed
Zcor.needs1 <- DescTools::FisherZ(Cor.Item.step.1.AUS_imp)

# Replace "inf" values on the diagonal for the value 1
Zcor.needs1 <- as.data.frame(rbind(as.numeric(gsub("Inf", "1",
                            (as.character(Zcor.needs1[1,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[2,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[3,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[4,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[5,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[6,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[7,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[8,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[9,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[10,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[11,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[12,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[13,])))),
        as.numeric(gsub("Inf", "1", (as.character(Zcor.needs1[14,]))))))

# after performing three times the above, there are three Zcor.needs objects.
# these contain Z-transformations that can be average using the following r code.
n <- nrow(Zcor.needs1)
pooled.Fisher.need.AUS <- matrix(NA, nrow=14, ncol = 14)
for(j in 1:n) {
  for(i in 1:n) {
        pooled.Fisher.need.AUS[i,j] <- ifelse((as.numeric(!is.na(
          Zcor.needs1[i,j]))==1),
             mean(c(Zcor.needs1[i,j], Zcor.needs2[i,j], Zcor.needs3[i,j]),
                  na.rm = TRUE), NA)
  }
}
```

```r
# the inverse Z-transformation was applied to obtain Pearson correlations
# these Pearson correlations are the mean correlations between items
# over the item steps
pooled.FisherInv.need.AUS <- FisherZInv(pooled.Fisher.need.AUS)

# please note that the object pooled.FisherInv.need.AUS contains class <chr>
# these were reverted to numeric. This part of the syntax is not provided.
```

```
## Warning: Number of logged events: 2800

## Warning: Number of logged events: 2800

## Warning: Number of logged events: 2800
```

Input correlations matrices

The concept of item step is novel to circumplex model analysis. Item steps reflect the process of stepping from the lower response category k to the higher response category *k+1*. Contrary to GPCM and GGUM, the circumplex does not standard include item step estimation for polytomous scored items. Within GPCM and GGUM, the inclusion of item steps usually improves model fit. The unavailabilty of item steps gave the circumplex model a disadvantage. To include item steps, we applied a strategy that codes item steps. Coding of item steps is slightly different from estimating item steps. Item step coding is routine practice in non-parametric IRT model (Sijtsma and Molenaar (2002)). Item step codes recoded the original four-point Likert scale into three dummies variables. When teachers endorsed the response category $k = 2$ (i.e. *"moderate level of need"*), then the item steps $k = 1$ and $k = 3$ were coded 0 and the item step $k = 1$ was coded 1. The chunk with Rcode is printed above for further explanation. This method of item step coding corresponds with expected item response patterns in unfolding models (Mokken, Schuur W. H., and J. (2001)). Three item correlations matrices (D) were estimated, namely: D_{step1}, D_{step2} , D_{step3}, where D represents the correlation matrix and the subscript identifies the 14 item step dummy variables. For example, D_{step2} is the correlation matrix of the 14 items coding for step 2. The three correlation matrices were then pooled ($D - pooled$) resulting in one correlations matrix representing the the combination of the three steps. $D - pooled$ was used as input for the circumplex analysis to estimate uniform item locations that apply to all steps. After item step coding was performed - as outlined above - input correlation matrices were estimated. These input correlation matrices were inserted in the CIRCUM program to estimate circumplex model fit (see Table 8, 9, 10, and 11).

Table 11: Australia correlations CPL needs item step coding

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.611	1												
C	Knowledge curriculum	0.486	0.486	1											
D	Student assessment	0.309	0.343	0.358	1										
E	ICT skills	0.102	0.114	0.12	0.157	1									
F	classroom management	0.275	0.282	0.264	0.261	0.119	1								
G	administration	0.154	0.185	0.174	0.131	0.101	0.208	1							
H	individualised learning	0.237	0.3	0.256	0.337	0.136	0.304	0.16	1						
I	teaching students with special needs	0.155	0.174	0.166	0.203	0.133	0.254	0.109	0.449	1					
J	teaching in multicultural settings	0.125	0.149	0.131	0.125	0.111	0.175	0.159	0.226	0.33	1				
K	teaching cross-curricular skills	0.188	0.216	0.22	0.237	0.169	0.184	0.136	0.322	0.239	0.232	1			
L	analysis of assessments	0.235	0.256	0.287	0.496	0.152	0.248	0.147	0.366	0.242	0.182	0.324	1		
M	teacher-parent cooperation	0.277	0.273	0.248	0.211	0.089	0.299	0.232	0.224	0.174	0.212	0.226	0.272	1	
N	Communicating other cultures	0.153	0.144	0.16	0.116	0.073	0.186	0.191	0.145	0.183	0.455	0.189	0.189	0.394	1

Table 12: England correlations CPL needs item step coding

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.608	1												
C	Knowledge curriculum	0.528	0.43	1											
D	Student assessment	0.282	0.31	0.346	1										
E	ICT skills	0.088	0.08	0.099	0.151	1									
F	classroom management	0.343	0.31	0.323	0.262	0.114	1								
G	administration	0.142	0.133	0.158	0.15	0.11	0.188	1							
H	individualised learning	0.282	0.275	0.254	0.321	0.115	0.326	0.163	1						
I	teaching students with special needs	0.164	0.159	0.16	0.213	0.125	0.173	0.123	0.431	1					
J	teaching in multicultural settings	0.136	0.129	0.133	0.164	0.12	0.145	0.143	0.221	0.309	1				
K	teaching cross-curricular skills	0.245	0.263	0.227	0.282	0.157	0.24	0.149	0.363	0.26	0.218	1			
L	analysis of assessments	0.243	0.297	0.266	0.464	0.115	0.255	0.202	0.294	0.195	0.174	0.33	1		
M	teacher-parent cooperation	0.251	0.254	0.265	0.207	0.086	0.314	0.2	0.262	0.155	0.203	0.249	0.283	1	
N	Communicating other cultures	0.176	0.169	0.171	0.154	0.095	0.172	0.139	0.188	0.175	0.458	0.206	0.188	0.355	1

Table 13: Japan correlations CPL needs item step coding

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.799	1												
C	Knowledge curriculum	0.305	0.274	1											
D	Student assessment	0.413	0.426	0.384	1										
E	ICT skills	0.234	0.246	0.292	0.304	1									
F	classroom management	0.414	0.423	0.278	0.407	0.311	1								
G	administration	0.131	0.122	0.335	0.179	0.203	0.153	1							
H	individualised learning	0.382	0.4	0.233	0.381	0.278	0.455	0.123	1						
I	teaching students with special needs	0.284	0.293	0.202	0.314	0.237	0.38	0.117	0.556	1					
J	teaching in multicultural settings	0.169	0.162	0.278	0.212	0.262	0.18	0.3	0.217	0.212	1				
K	teaching cross-curricular skills	0.299	0.291	0.353	0.346	0.303	0.325	0.243	0.326	0.273	0.361	1			
L	analysis of assessments	0.359	0.351	0.364	0.527	0.312	0.402	0.214	0.411	0.332	0.284	0.478	1		
M	teacher-parent cooperation	0.284	0.291	0.304	0.34	0.246	0.389	0.256	0.354	0.323	0.285	0.341	0.426	1	
N	Communicating other cultures	0.188	0.175	0.277	0.236	0.275	0.224	0.291	0.235	0.215	0.659	0.365	0.335	0.365	1

Table 14: The Netherlands correlations CPL needs item step coding

1	2	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	Subject Knowledge	1													
B	Pedagogical competency	0.375	1												
C	Knowledge curriculum	0.289	0.225	1											
D	Student assessment	0.231	0.175	0.32	1										
E	ICT skills	0.059	0.058	0.072	0.092	1									
F	classroom management	0.213	0.33	0.151	0.157	0.089	1								
G	administration	0.053	0.038	0.067	0.135	0.015	0.041	1							
H	individualised learning	0.143	0.142	0.175	0.158	0.119	0.182	0.02	1						
I	teaching students with special needs	0.121	0.129	0.108	0.113	0.074	0.22	-0.01	0.297	1					
J	teaching in multicultural settings	0.062	0.082	0.134	0.131	0.044	0.06	0.118	0.09	0.146	1				
K	teaching cross-curricular skills	0.087	0.103	0.109	0.152	0.092	0.132	0.052	0.227	0.206	0.073	1			
L	analysis of assessments	0.152	0.146	0.194	0.27	0.131	0.163	0.085	0.171	0.135	0.088	0.164	1		
M	teacher-parent cooperation	0.169	0.216	0.163	0.155	0.096	0.177	0.106	0.11	0.112	0.151	0.088	0.26	1	
N	Communicating other cultures	0.095	0.112	0.144	0.133	0.063	0.076	0.119	0.047	0.095	0.451	0.077	0.122	0.261	1

Formulas for estimation of teacher positions with item step coding

The teacher positions estimates the areas that describe the professional learning needs that the teacher most likely has highest need for. The GPCM and GGUM have standard implemented algorithms to estimate teacher positions. For the circumplex, teacher position $\angle_{(p)}$ was estimated using the following formula:

If:

$$\cos(x_i * \angle_{(i)} * \frac{\pi}{180}) \leq 0 \tag{1}$$

then:

$$\angle_{(p)} = 360 + \left(\frac{180}{\pi}\right) \tan^{-1}\left(\frac{\sum_{i=1}^{k}(h_i * x_{ik} * \sin(x_{ik} * \angle_i * \frac{\pi}{180}))}{\sum_{i=1}^{k}(h_i * x_{ik} * \cos(x_{ik} * \angle_i * \frac{\pi}{180}))}\right) \tag{2}$$

And if:

$$\cos(x_i * \angle_{(i)} * \frac{\pi}{180} > 0 \tag{3}$$

then:

$$\angle_{(p)} = \left(\frac{180}{\pi}\right) \tan^{-1}\left(\frac{\sum_{i=1}^{k}(h_i * x_{ik} * \sin(x_{ik} * \angle_i * \frac{\pi}{180}))}{\sum_{i=1}^{k}(h_i * x_{ik} * \cos(x_{ik} * \angle_i * \frac{\pi}{180}))}\right) \tag{4}$$

Wherein $\angle(p)$ defines the angular teacher position; $\angle(i)$ defines the angular item position; h_i is the model estimated item communality, and x_{ik} is the dummy item step code value.

Teacher locations were subsequently estimated using the item step code variables as indicated in the formulas 1-4. Teachers' position was estimated per item step code, meaning that we had up to three teachers positions. One position estimating the area(s) for which the teacher most likely has a *high level of need*, one position estimating the area(s) for which the teacher most likely has a *moderate level of need*, and one position estimating the area(s) for which the teacher most likely has a *low level of need*. Teachers failing to respond with a certain level (e.g., none of the items is responded to with *low level of need*) or teachers with full response patterns (i.e. meaning that teachers respond to all items with the same level of need) had fewer than three estimated positions, however. Positions for teachers with full response patterns of 0 = *no need at present* could not be estimated. These teacher had NA.

Model fit estimates: comparison GPCM, GGUM and Circumplex

Table 15: Fit coefficients for PCM GGUM and Circumplex models

model	CFI	RMSEA	RMSEA 5%	RMSEA 95%	M2 / Chi	df
Austrlia						
GPCM	0.86	0.141	0.136	0.146	2634.718	49
GPCM	0.69	0.084	0.078	0.089	686.051	35
Circumplex	NA	0.071	0.067	0.075	901	61
England						
GPCM	0.93	0.105	0.1	0.111	1010.92	49
GPCM	0.76	0.096	0.09	0.103	611.691	35
Circumplex	NA	0.076	0.071	0.081	710	61
Japan						
GPCM	0.88	0.105	0.101	0.109	1757.006	49
GPCM	0.96	0.106	0.101	0.111	1292.044	35
Circumplex	NA	0.075	0.072	0.079	1183	61
Netherlands						
GPCM	0.86	0.099	0.093	0.106	715.978	49
GPCM	0.93	0.067	0.059	0.075	250.399	35
Circumplex	NA	0.068	0.062	0.062	452	61

Application circumplex: Visual output circumplex model

CIRCUM output files used as input for the circumplex visualizations:

- TALIS PD needs AUS (par steps).out
- TALIS PD needs ENG (par steps).out
- TALIS PD needs JPN (par steps).out
- TALIS PD needs NLD (par steps).out

Most of the output of the circumplex is extensively discussed in the article. It will not be further discussed here. Just the figures are printed.

Circumplex: item angular locations

CPL needs circlar unfolding Australia: item locations

```
Circumplex.CPL.Needs.AUS.step <-

## cirkel schattingen --> item 27 PD needs.

Item.angles.AUS <- c(0,120,234,212,201,208,210,175,176,190,195,203,211,202)

item.position.AUS.1 <- circular(Item.angles.AUS[[1]], type = "angles",
                                units = "degrees", names = Item.angles.names)
item.position.AUS.2 <- circular(Item.angles.AUS[[2]], type = "angles",
```

```
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.3 <- circular(Item.angles.AUS[[3]], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.4 <- circular(Item.angles.AUS[c(8,9,11)], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.5 <- circular(Item.angles.AUS[c(10,14)], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.6 <- circular(Item.angles.AUS[c(5,12)], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.7 <- circular(Item.angles.AUS[c(7,13)], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.8 <- circular(Item.angles.AUS[[4]], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
item.position.AUS.9 <- circular(Item.angles.AUS[[6]], type = "angles",
                                                          units = "degrees", names = Item.angles.names)
Table.item.position.AUS <- data.frame(cbind(c("Subject Knowledge",
    "Pedagogical competency","individualised learning",
    "teaching students with special needs", "teaching in multicultural settings",
   "teaching cross-curricular skills","ICT skills",
   "communicating with people of other cultures","analysis of assessments",
   "classroom management","teacher-parent cooperation", "administration",
   "student assessment practices","Knowledge of the curriculum"),
   as.character(c(0,120,175,176,190,195,201,202,203,208,210,211,212,234))))
rownames(Table.item.position.AUS) <- c("A","B","H","I","J","K","E","N","L","F",
                                       "G","M","D","C")
colnames(Table.item.position.AUS) <- c("Professional learning area", "angle")
color.rownames <- c("red", "green","DeepPink1","DeepPink1","RosyBrown4",
                    "DeepPink1","orange","RosyBrown4","orange","green",
                    "Purple1","Purple1","orange","yellow")
```

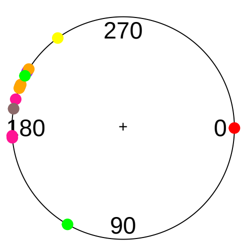


Figure 15: Circular Unfolding Australian teachers' CPL needs

Table 16: Item locations Australia circular unfolding

	Professional learning area	angle
A	Subject Knowledge	0
B	Pedagogical competency	120
H	individualised learning	175
I	teaching students with special needs	176
J	teaching in multicultural settings	190
K	teaching cross-curricular skills	195
E	ICT skills	201
N	communicating with people of other cultures	202
L	analysis of assessments	203
F	classroom management	208
G	teacher-parent cooperation	210
M	administration	211
D	student assessment practices	212
C	Knowledge of the curriculum	234

CPL needs circlar unfolding England: item locations

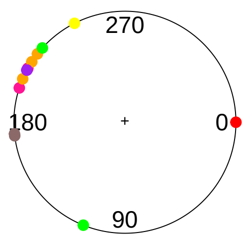


Figure 16: Circular unfolding Englisch teachers' CPL needs

Table 17: Item locations England circular unfolding

	Professional learning area	angle
A	Subject Knowledge	0
B	Pedagogical competency	112
N	communicating with people of other cultures	173
J	teaching in multicultural settings	174
I	teaching students with special needs	198
E	ICT skills	203
G	administration	208
M	teacher-parent cooperation	208
K	teaching cross-curricular skills	209
H	individualised learning	209
L	analysis of assessments	213
D	student assessment practices	218
F	classroom management	222
C	Knowledge of the curriculum	243

CPL needs circlar unfolding Japan: item locations

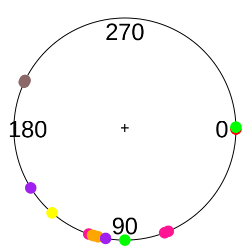


Figure 17: Circular unfolding Japanese teachers' CPL needs

Table 18: Item locations Japan circular unfolding

	Professional learning area	angle
A	Subject Knowledge	0
H	individualised learning	67
I	teaching students with special needs	69
F	classroom management	90
M	teacher-parent cooperation	100
L	analysis of assessments	104
D	student assessment practices	106
E	ICT skills	107
K	teaching cross-curricular skills	109
C	Knowledge of the curriculum	131
G	administration	148
J	teaching in multicultural settings	205
N	communicating with people of other cultures	206
B	Pedagogical competency	359

CPL needs circlar unfolding the Netherlands: item locations

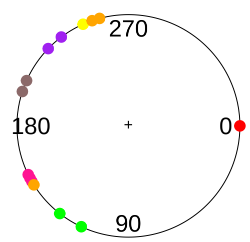


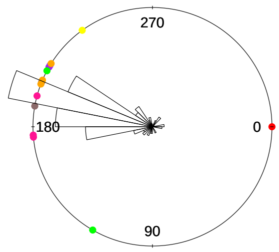
Figure 18: Circular unfolding Dutch teachers' CPL needs

Table 19: Item locations The Netherlands circular unfolding

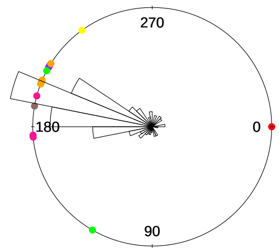
	Professional learning area	angle
A	Subject Knowledge	0
B	Pedagogical competency	115
F	classroom management	128
E	ICT skills	148
H	individualised learning	150
K	teaching cross-curricular skills	152
I	teaching students with special needs	154
J	teaching in multicultural settings	198
N	communicating with people of other cultures	204
G	administration	224
M	teacher-parent cooperation	233
C	Knowledge of the curriculum	246
D	student assessment practices	251
L	analysis of assessments	255

Teacher positions with item step coding

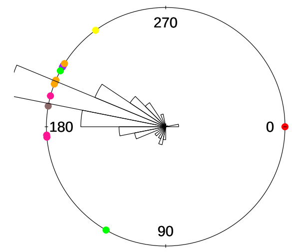
CPL needs country distributions Figure 19 and Figure 20 present the empirical distributions of the teachers' CPL needs in Australia, England, Japan and the Netherlands. Three plots are generated: at the right are teachers who have expressed a high level of need for learning on one or more of the areas. In the middle are teachers who have no high level of need for any of the areas but who did express a moderate level of need on one or more of the areas. At the left are the teachers who have neither expressed a high level of need, nor a moderate level of need for learning on one of the areas, but who did express a low level of need.



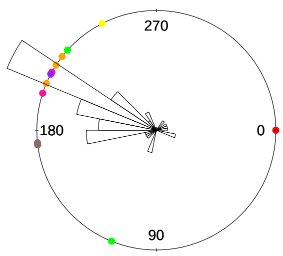
(a) Australian teachers with high CPL needs (n = 1012)



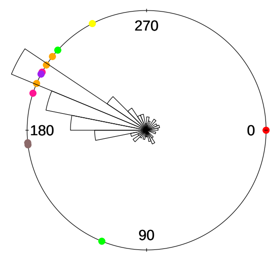
(b) Australian teachers with only moderate CPL needs (n = 1411)



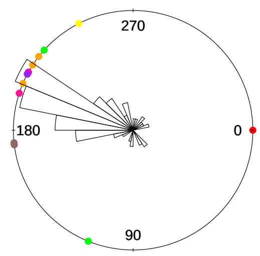
(c) Australian teachers with only low CPL needs (n = 275)



(d) Englisn teachers with high CPL needs (n = 358)

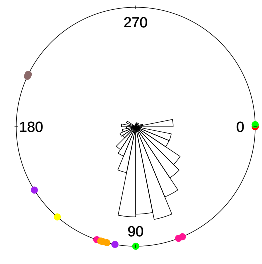


(e) English teachers with only moderate CPL needs (n = 1046)

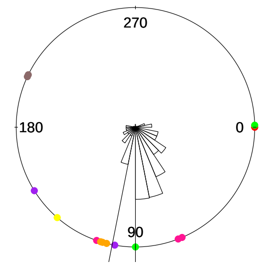


(f) English teachers with only low CPL needs (n = 351)

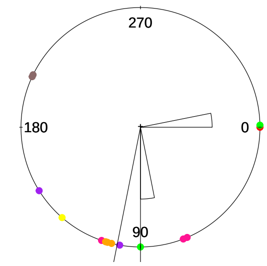
Figure 19



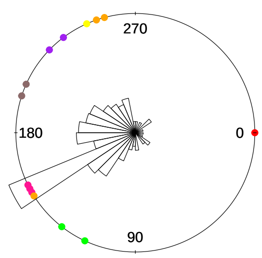
(a) Japanese teachers with high CPL needs (n = 2710)



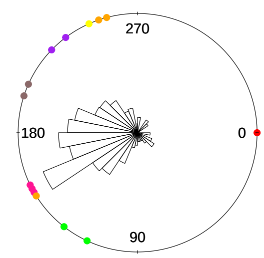
(b) Japanese teachers with only moderate CPL needs (n = 518)



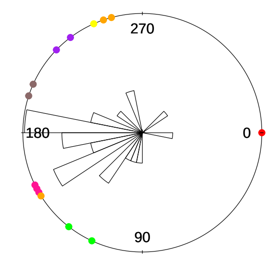
(c) Japanese teachers with only low CPL needs (n = 9)



(d) Dutch teachers with high CPL needs (n = 745)



(e) Dutch teachers with only moderate CPL needs (n = 585)



(f) Dutch teachers with only low CPL needs (n = 50)

Figure 20

Bond, T. G., and Fox, C. M. (2007). *Applying the rasch model: Fundamental measurement in the human sciences* (2nd Edition). Taylor & Francis Group.

Browne, M. W. (1992). Circumplex model for correlation matrices. *Psychometrika*, *57*, 469–497. https://doi.org/10.1007/BF02294416

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the r environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Economic Co-operation, O. for, and (OECD), D. S. (2019). *TALIS 2018 results (volume i)*. OECD. https://doi.org/10.1787/1d0bc92a-en

Guttman, L. (1954). Guttman (1954). A new approach to factor analysis. The radex. In P. F. Lazarsfeld (Ed.), *Mathematical Thinking in the Social Sciences* (pp. 258–348). Glencoe, Illinois: Free Press.

Kyriakides, L., Creemers, B. P. M., and Panayiotou, A. (2018). Using educational effectiveness research to promote quality of teaching: The contribution of the dynamic model. *ZDM - Mathematics Education*, *50*, 381–393. https://doi.org/10.1007/s11858-018-0919-3

Mokken, R. J., Schuur W. H., van, and J., L. A. (2001). The circles of our minds: A nonparametric IRT model for the circumplex. In A. Boomsma, M. A. J. van Duijn, and T. B. Snijders (Eds.), *Essays on item response theory* (pp. 339–356). New York, USA: Springer.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. In *ETS Research Report Series* (Vol. 1992, pp. 1–30). Educational Testing Service; Wiley. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x

Roberts, J. S., Donoghue, J. R., Service, T., and Laughlin, J. E. (1996). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, *24*, 3–32.

Sijtsma, K., and Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. London, England: Sage Publications Ltd.