

Through the eyes of the teacher

Mandy Klatt¹, Dr. Gregor Kachel^{1, 2}, Dr. Christin Lotz¹, & Prof. Dr. Anne Deiglmayr¹

¹ Leipzig University

² Max-Planck University for Evolutionary Anthropology

Author Note

The Ethics Advisory Board of Leipzig University has dealt with the research project and has come to the conclusion that there are no objections to the implementation of this research project. The Ethics Advisory Board points out that the scientific and ethical responsibility for the implementation of the project remains with the project director.

Correspondence concerning this article should be addressed to Mandy Klatt, Egelstraße 2a 04103 Leipzig. E-mail: mandy.klatt@uni-leipzig.de

Abstract

12

13 This document is a supplement to the paper and shows first graphs findings from the pilot
14 study.

15 *Keywords:* Professional Vision, Expert-Novice-Paradigm, Eye-Tracking

16 Word count: 1949

Through the eyes of the teacher

State of research

Teaching and classroom management are multidimensional settings in which teachers have to respond immediately to events as they develop (Barnes, 2004). The different interests and abilities of students must be managed in a way that maximizes the active learning time of students and minimizes disruptions whilst teaching. Learning to develop such classroom management skills and to teach effectively is a complicated and complex process (Wolff, Jarodzka, & Boshuizen, 2017).

During teaching, teachers must be able to select from a variety of visual and acoustic impressions to focus their attention on the essential and to distinguish between relevant and irrelevant events. This ability is called professional vision and is a key component of teacher expertise and successful teaching (Barth, 2017). Eye tracking technology has become a reliable means to study teachers' visual focus of attention (Bogert, 2016; Pouta, Lehtinen, & Palonen, 2020; Wolff, Jarodzka, & Boshuizen, 2017)

Educational research has repeatedly shown that there are differences between experienced and novice teachers in terms of perception and behavioral competencies (Barth, 2017; Bogert, 2016; Wolff, Jarodzka, & Boshuizen, 2017). For example, experts direct their attention more often and more evenly to all students, whereas novices only direct their attention to some students. The frequency and duration of fixations as eye movement are decisive (Stuermer, Seidel, Mueller, Häusler, & Cortina, 2017). Mobile eye-tracking technology has also shown that experienced teachers distribute their focus more efficiently to solve tasks (Jarodzka, Scheiter, Gerjets, & Van Gog, 2010). Furthermore, in contrast to novices, experts are able to focus their attention on the entire class and guide the class while giving feedback to individual students and answering questions (Cortina, Miller, McKenzie, & Epstein, 2015).

Research questions

The aim of the pilot study was to investigate whether there are differences in how expert and novice teachers manage scripted classroom disruptions. The disruptions were experimentally varied using a previously written script. Thus, our aim was to find out whether differences in the allocation of attention between expertise groups can be detected in this controlled context.

In order to answer this question, the hypothesis was formulated that teachers with more professional experience not only notice more disruptions but also notice them faster. In the hypothesis, therefore, it is necessary to check what has already been shown in the research literature: In complex teaching situations, experts have a more structured and elaborate professional knowledge than novices in order to perceive and interpret relevant events and to act appropriately (Berliner, 2001; Lachner, Jarodzka, & Nückles, 2016).

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

For the sample recruitment of the subjects ($N = 8$, experts $n = 2$, novices $n = 6$), schools in the city of Leipzig in Saxony were contacted. The institutions as well as the subjects were informed in detail about the aim and intention of the study in advance. Participation in the study was voluntary and only took place after written consent has been given.

The selection of the subjects was based on extreme groups, whereby professional experience is the crucial criterion for the selection of experts or novices. Novices were recruited as teachers who have been working in the teaching profession for no more than 3

Table 1

Demographic Information and Teaching Experience

group	N	Male	M age	Min age	Max age	SD age	M exp.	Min exp.	Max exp.	SD exp.
expert	2	1	47.50	44	51	4.95	20.00	15.00	25.00	7.07
novice	6	2	25.67	20	33	4.89	0.68	0.00	1.50	0.68

years, whereas experts were considered to have professional experience of 10 years or more (Messner & Reusser, 2000).

Procedure/ Data collection

Set up. For this study, scripted mini-lessons with $n = 2$ experts and $n = 6$ novices were recorded in the mobile Lab of the Empirical School and Classroom Research at the University of Leipzig. The subjects were divided into groups of four, so the study was conducted on two different sessions. All participants were asked to hold a 10-minute lesson. The duration of each appointment was approximately 2h: per group 10min briefing, 4 x 10min mini-lessons, 10min technical preparation and follow-up and 4x 10min transition points between the lessons and answering questionnaires.

One person from the group of 4 acted as a teacher, the other three subjects acted as the class. The subjects, who represented the class, were given behavioral instructions in a pre-written script to simulate typical events and disruptions in the classroom (e.g. putting their heads on the table, chatting, looking at their mobile phones, etc.).

The lesson disruptions were displayed as instructions during the lesson for all “students” but not the teacher. In order to avoid learning effects, the disruptions in each lesson were distributed pseudo-randomly over the short teaching phase. In addition, the order of the data collection was taken into account in the analyses and variance caused by

84 order was controlled.



Figure 1. Example for set up during a mini-lesson

85 **Questionnaire data.** After each mini-lesson, the students answered items on the
86 teaching quality using a validated questionnaire (Helmke et al., 2014) and scales on the
87 teacher's presence behavior (students $n = 24$). In addition, the teacher was asked to give a
88 self-assessment on his/her classroom management by completing the questionnaire after
89 each mini-lesson (teachers $n = 8$).

90 **Behavioral data.** The speech, sounds and voices were recorded with an audio
91 recorder installed in the middle of the Lab. Movements, facial expressions and gestures of
92 the subjects were recorded by four cameras from different angles. One camera was installed
93 to film the class from the side. Two more cameras were installed on the blackboard and at
94 the end of the Lab to film the teacher and class from the front and back. Furthermore, the
95 fourth camera was installed in such a way that only facial expressions and gestures of the

teacher were recorded, which enables a semi-automated analysis of the movement sequences.

Eyetracking data. A binocular Tobii Pro Glasses 2 eye-tracker consisting of a wearable head unit and a recording unit was used to record the eye movements of all 8 participants. The head unit is a measuring device with different sensitive sensors. A high-definition scene camera captures a full HD video and an integrated microphone records the surrounding sounds. Infrared light illuminators support the eye tracking sensors which record the eye orientation. The videos were recorded with a sampling rate of 50 Hz and a video resolution with 1920 x 1080 at 25 frames per second. The scene camera has a field of view of 90 deg. in 16:9 format (82 deg. horizontal and 52 deg. vertical) and has a frame dimension of 179 x 159 x 57mm (width x depth x height). The Tobii Pro Glasses Controller software was used to record and calibrate the eye movements.

Coding/ Data preparation/ Reliability

Questionnaire Data. The evaluation after each mini-lesson was conducted using paper questionnaires. Time needed to complete the questionnaire was about 5 minutes. The scales on the quality of teaching are a validated questionnaire (Helmke et al., 2014). Whereas the scales on the teacher's presence behavior were derived from the research literature (Brophy, 1986; Kiel, Frey, Weiß, & Weiss, 2013; Kounin, 2006; Marzano, 2007; Nolting, 2012) and were used in the pilot for the first time. The questionnaire is 4-point Likert scale (1 = Strongly Disagree; 2 = Disagree; 3 = Agree; 4 = Strongly Agree). Data was obtained from N = 32 subjects (students n = 24, teachers n = 8).

The following scales were assessed:

- (1) Classroom management
- (2) Positive climate and motivation
- (3) Clarity and structuredness

- (4) Activation and support
- (5) Presence: posture/gaze
- (6) Presence: voice
- (7) Presence: verbal and non-verbal intervention
- (8) Natural behaviour

Table 2 provides an overview over the mean, the standard deviation, the range, Cronbach's Alpha and the Skewness & Kurtosis of all scales for the teachers' self-assessment.

```
##
## Reliability analysis
## Call: alpha(x = self.as.wide[, -1])
##
##      raw_alpha std.alpha G6(smc) average_r S/N      ase mean      sd median_r
##      0.85      0.87      0.85      0.69 6.8 0.095   2.5 0.97      0.62
##
## lower alpha upper      95% confidence boundaries
## 0.67 0.85 1.04
##
## Reliability if an item is dropped:
##
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
## af_beiträge      0.75      0.76      0.62      0.62 3.2      0.169      NA
## af_nachdenken      0.71      0.74      0.59      0.59 2.9      0.185      NA
## af_wechselseitig      0.93      0.93      0.87      0.87 13.7      0.049      NA
##
## med.r
## af_beiträge      0.62
## af_nachdenken      0.59
```



```

147 ## af_wechselseitig 0.87
148 ##
149 ## Item statistics
150 ##
151 ##          n raw.r std.r r.cor r.drop mean  sd
152 ## af_beiträge      8 0.90 0.92 0.90 0.80 3.0 0.93
153 ## af_nachdenken    8 0.92 0.93 0.92 0.81 2.4 1.06
154 ## af_wechselseitig 8 0.86 0.82 0.65 0.62 2.0 1.31
155 ##
156 ## Non missing response frequency for each item
157 ##          1    2    3    4 miss
158 ## af_beiträge    0.00 0.38 0.25 0.38    0
159 ## af_nachdenken  0.25 0.25 0.38 0.12    0
160 ## af_wechselseitig 0.50 0.25 0.00 0.25    0
161 ##
162 ## Warning in alpha(self.cs.wide[, -1]): Some items were negatively correlated with the
163 ## should be reversed.
164 ## To do this, run the function again with the 'check.keys=TRUE' option
165 ##
166 ## Some items ( ks_sichtbar ) were negatively correlated with the total scale and
167 ## probably should be reversed.
168 ## To do this, run the function again with the 'check.keys=TRUE' option
169 ##
170 ## Warning in sqrt(Vtc): NaNs wurden erzeugt
171 ##
172 ## Reliability analysis
173 ## Call: alpha(x = self.cs.wide[, -1])
174 ##

```

```

171 ##   raw_alpha std.alpha G6(smc) average_r   S/N   ase mean   sd median_r
172 ##      -0.76      -1.1   -0.36      -0.36 -0.53 0.94   2.8 0.53      -0.36
173 ##
174 ##   lower alpha upper      95% confidence boundaries
175 ## -2.6 -0.76 1.08
176 ##
177 ##   Reliability if an item is dropped:
178 ##               raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r
179 ## ks_sichtbar      -0.75      -0.36    0.13      -0.36 -0.26      NA    0
180 ## ks_verständlich  -0.17      -0.36    0.13      -0.36 -0.26      NA    0
181 ##               med.r
182 ## ks_sichtbar      -0.36
183 ## ks_verständlich -0.36
184 ##
185 ##   Item statistics
186 ##               n raw.r std.r r.cor r.drop mean   sd
187 ## ks_sichtbar     8 0.88 0.57  NaN  -0.36  3.1 1.13
188 ## ks_verständlich 8 0.13 0.57  NaN  -0.36  2.5 0.53
189 ##
190 ##   Non missing response frequency for each item
191 ##               1    2    3    4 miss
192 ## ks_sichtbar    0.12 0.12 0.25 0.5    0
193 ## ks_verständlich 0.00 0.50 0.50 0.0    0
194 ##
195 ##   Reliability analysis
196 ##   Call: alpha(x = self.cm.wide[, -1])
197 ##

```

```

198 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
199 ##       0.83       0.82       0.93       0.54 4.6 0.084   2.4 0.72       0.55
200 ##
201 ##   lower alpha upper       95% confidence boundaries
202 ## 0.67 0.83 1
203 ##
204 ## Reliability if an item is dropped:
205 ##               raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
206 ## km_aktiv           0.84       0.83       0.83       0.62 5.0       0.091 0.049 0.57
207 ## km_klar            0.71       0.68       0.86       0.42 2.2       0.172 0.167 0.44
208 ## km_mitbekommen     0.89       0.89       0.94       0.74 8.4       0.070 0.033 0.82
209 ## km_ungestört       0.65       0.63       0.70       0.37 1.7       0.182 0.101 0.53
210 ##
211 ## Item statistics
212 ##               n raw.r std.r r.cor r.drop mean   sd
213 ## km_aktiv       8 0.73 0.73 0.71 0.56 2.0 0.76
214 ## km_klar        8 0.93 0.92 0.90 0.84 2.5 1.07
215 ## km_mitbekommen 8 0.60 0.62 0.51 0.40 2.8 0.71
216 ## km_ungestört   8 0.97 0.97 0.99 0.93 2.5 0.93
217 ##
218 ## Non missing response frequency for each item
219 ##               1    2    3    4 miss
220 ## km_aktiv      0.25 0.50 0.25 0.00 0
221 ## km_klar       0.12 0.50 0.12 0.25 0
222 ## km_mitbekommen 0.00 0.38 0.50 0.12 0
223 ## km_ungestört  0.12 0.38 0.38 0.12 0
224 ##

```

```

225 ## Reliability analysis
226 ## Call: alpha(x = self.nb.wide[, -1])
227 ##
228 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
229 ##       0.9       0.9   0.87       0.76 9.4 0.061   2.8 0.71       0.76
230 ##
231 ##   lower alpha upper      95% confidence boundaries
232 ## 0.78 0.9 1.02
233 ##
234 ## Reliability if an item is dropped:
235 ##           raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r med.r
236 ## m_fiktiv       0.91       0.91   0.83       0.83 10.1   0.065   NA 0.83
237 ## m_natürlich    0.81       0.81   0.68       0.68  4.2   0.135   NA 0.68
238 ## m_verhalten    0.86       0.86   0.76       0.76  6.4   0.096   NA 0.76
239 ##
240 ## Item statistics
241 ##           n raw.r std.r r.cor r.drop mean   sd
242 ## m_fiktiv   8 0.88 0.89 0.79 0.75 3.0 0.76
243 ## m_natürlich 8 0.94 0.94 0.92 0.87 2.6 0.74
244 ## m_verhalten 8 0.92 0.91 0.86 0.81 2.9 0.83
245 ##
246 ## Non missing response frequency for each item
247 ##           2    3    4 miss
248 ## m_fiktiv   0.25 0.50 0.25    0
249 ## m_natürlich 0.50 0.38 0.12    0
250 ## m_verhalten 0.38 0.38 0.25    0
251 ##

```

```

252 ## Reliability analysis
253 ## Call: alpha(x = self.pcm.wide[, -1])
254 ##
255 ##   raw_alpha std.alpha G6(smc) average_r S/N ase mean   sd median_r
256 ##      0.78      0.78    0.97      0.37 3.5 0.11   2.9 0.61      0.41
257 ##
258 ## lower alpha upper      95% confidence boundaries
259 ## 0.57 0.78 1
260 ##
261 ## Reliability if an item is dropped:
262 ##               raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
263 ## lkm_ausreden      0.81      0.81    0.88      0.47 4.4      0.10 0.083
264 ## lkm_freundlich    0.80      0.78    0.91      0.41 3.5      0.09 0.120
265 ## lkm_interesse     0.73      0.74    0.98      0.36 2.8      0.15 0.114
266 ## lkm_kritik        0.69      0.66    0.84      0.28 2.0      0.16 0.122
267 ## lkm_rückmeldungen 0.67      0.68    0.79      0.30 2.1      0.18 0.088
268 ## lkm_überlegen     0.77      0.78    0.88      0.41 3.5      0.12 0.091
269 ##               med.r
270 ## lkm_ausreden      0.48
271 ## lkm_freundlich    0.46
272 ## lkm_interesse     0.41
273 ## lkm_kritik        0.24
274 ## lkm_rückmeldungen 0.41
275 ## lkm_überlegen     0.43
276 ##
277 ## Item statistics
278 ##               n raw.r std.r r.cor r.drop mean   sd

```

```

279 ## lkm_ausreden      8  0.32  0.46  0.46   0.20  3.2 0.46
280 ## lkm_freundlich    8  0.51  0.59  0.59   0.31  3.1 0.83
281 ## lkm_interesse     8  0.80  0.72  0.66   0.63  2.6 1.19
282 ## lkm_kritik        8  0.88  0.90  0.91   0.82  3.0 0.76
283 ## lkm_rückmeldungen 8  0.91  0.86  0.87   0.85  2.5 0.93
284 ## lkm_überlegen     8  0.66  0.60  0.60   0.48  2.6 0.92
285 ##
286 ## Non missing response frequency for each item
287 ##              1    2    3    4 miss
288 ## lkm_ausreden    0.00 0.00 0.75 0.25    0
289 ## lkm_freundlich  0.00 0.25 0.38 0.38    0
290 ## lkm_interesse   0.25 0.12 0.38 0.25    0
291 ## lkm_kritik      0.00 0.25 0.50 0.25    0
292 ## lkm_rückmeldungen 0.12 0.38 0.38 0.12    0
293 ## lkm_überlegen   0.12 0.25 0.50 0.12    0

294 ## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was done

295 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
296
297 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
298
299 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done

300 ##
301 ## Reliability analysis
302 ## Call: alpha(x = self.ppg.wide[, -1])
303 ##

```

```

304 ## raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
305 ## 0.85 0.86 0.94 0.47 6.3 0.077 2.7 0.69 0.45
306 ##
307 ## lower alpha upper 95% confidence boundaries
308 ## 0.7 0.85 1.01
309 ##
310 ## Reliability if an item is dropped:
311 ## raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
312 ## phb_alleangesehen 0.87 0.88 0.97 0.55 7.3 0.072 0.099
313 ## phb_augen 0.81 0.82 0.98 0.44 4.7 0.104 0.113
314 ## phb_blick 0.79 0.80 0.99 0.40 3.9 0.114 0.102
315 ## phb_gestik 0.88 0.88 0.96 0.56 7.5 0.066 0.081
316 ## phb_raum 0.88 0.89 0.97 0.57 8.1 0.068 0.078
317 ## phb_stand 0.78 0.80 0.86 0.40 4.0 0.118 0.107
318 ## phb_vorsichgeht 0.80 0.80 0.94 0.40 3.9 0.108 0.110
319 ## med.r
320 ## phb_alleangesehen 0.49
321 ## phb_augen 0.45
322 ## phb_blick 0.42
323 ## phb_gestik 0.52
324 ## phb_raum 0.62
325 ## phb_stand 0.42
326 ## phb_vorsichgeht 0.42
327 ##
328 ## Item statistics
329 ## n raw.r std.r r.cor r.drop mean sd
330 ## phb_alleangesehen 8 0.51 0.52 0.46 0.35 3.0 0.93

```

```

331 ## phb_augen          8  0.85  0.84  0.83   0.76  3.4 1.06
332 ## phb_blick          8  0.96  0.96  0.98   0.94  2.8 0.89
333 ## phb_gestik         8  0.52  0.50  0.48   0.33  2.6 1.06
334 ## phb_raum           8  0.43  0.45  0.36   0.26  1.8 0.89
335 ## phb_stand          8  0.95  0.95  0.97   0.93  2.9 0.99
336 ## phb_vorsichgeht    8  0.96  0.96  0.96   0.94  2.5 0.76
337 ##
338 ## Non missing response frequency for each item
339 ##                  1    2    3    4 miss
340 ## phb_alleangesehen 0.12 0.00 0.62 0.25    0
341 ## phb_augen         0.12 0.00 0.25 0.62    0
342 ## phb_blick         0.12 0.12 0.62 0.12    0
343 ## phb_gestik        0.12 0.38 0.25 0.25    0
344 ## phb_raum          0.50 0.25 0.25 0.00    0
345 ## phb_stand         0.12 0.12 0.50 0.25    0
346 ## phb_vorsichgeht   0.12 0.25 0.62 0.00    0

347 ## Warning in alpha(self.pvni.wide[, -1]): Some items were negatively correlated with th
348 ## should be reversed.
349 ## To do this, run the function again with the 'check.keys=TRUE' option

350 ## Some items ( pi_nonverbal ) were negatively correlated with the total scale and
351 ## probably should be reversed.
352 ## To do this, run the function again with the 'check.keys=TRUE' option

353 ## Warning in sqrt(Vtc): NaNs wurden erzeugt

354 ##

```



```

355 ## Reliability analysis
356 ## Call: alpha(x = self.pvni.wide[, -1])
357 ##
358 ##   raw_alpha std.alpha G6(smc) average_r   S/N ase mean   sd median_r
359 ##      -0.75      -0.81    -0.42      -0.18 -0.45    1  2.8 0.35      -0.17
360 ##
361 ##   lower alpha upper      95% confidence boundaries
362 ##  -2.76 -0.75 1.26
363 ##
364 ## Reliability if an item is dropped:
365 ##           raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r   med.r
366 ## pi_direkt      -0.58      -0.72  -0.264      -0.264 -0.42      0.95    NA -0.264
367 ## pi_nonverbal    -0.21      -0.22  -0.098      -0.098 -0.18      0.81    NA -0.098
368 ## pi_zubewegen   -0.38      -0.40  -0.165      -0.165 -0.28      0.94    NA -0.165
369 ##
370 ## Item statistics
371 ##           n raw.r std.r r.cor r.drop mean   sd
372 ## pi_direkt   8  0.49  0.53   NaN  -0.22  3.2 0.71
373 ## pi_nonverbal 8  0.64  0.41   NaN  -0.30  2.4 0.92
374 ## pi_zubewegen 8  0.20  0.46   NaN  -0.29  2.6 0.52
375 ##
376 ## Non missing response frequency for each item
377 ##           1    2    3    4 miss
378 ## pi_direkt  0.00 0.12 0.50 0.38    0
379 ## pi_nonverbal 0.12 0.50 0.25 0.12    0
380 ## pi_zubewegen 0.00 0.38 0.62 0.00    0
381 ##

```

```

382 ## Reliability analysis
383 ## Call: alpha(x = self.pv.wide[, -1])
384 ##
385 ##   raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
386 ##      0.77      0.76      0.82      0.51 3.2 0.12  2.8 0.62      0.54
387 ##
388 ## lower alpha upper      95% confidence boundaries
389 ## 0.53 0.77 1.01
390 ##
391 ## Reliability if an item is dropped:
392 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
393 ## ps_deutlich      0.65      0.70      0.54      0.54 2.32      0.213      NA 0.54
394 ## ps_impulse      0.89      0.89      0.79      0.79 7.75      0.081      NA 0.79
395 ## ps_klar      0.31      0.34      0.21      0.21 0.52      0.434      NA 0.21
396 ##
397 ## Item statistics
398 ##           n raw.r std.r r.cor r.drop mean sd
399 ## ps_deutlich 8 0.87 0.81 0.76 0.64 2.9 0.83
400 ## ps_impulse 8 0.62 0.71 0.52 0.39 2.6 0.52
401 ## ps_klar 8 0.96 0.95 0.95 0.88 2.9 0.83
402 ##
403 ## Non missing response frequency for each item
404 ##           2 3 4 miss
405 ## ps_deutlich 0.38 0.38 0.25 0
406 ## ps_impulse 0.38 0.62 0.00 0
407 ## ps_klar 0.38 0.38 0.25 0
408 ## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was done

```

```
409 ## Warning in alpha(self.p.wide[, -1]): Some items were negatively correlated with the t
410 ## should be reversed.
411 ## To do this, run the function again with the 'check.keys=TRUE' option

412 ## Some items ( pi_nonverbal ) were negatively correlated with the total scale and
413 ## probably should be reversed.
414 ## To do this, run the function again with the 'check.keys=TRUE' option

415 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done

416 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
417
418 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
419
420 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
421
422 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
423
424 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
425
426 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
427
428 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
429
430 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
431
432 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
433
```

```

434 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
435
436 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
437
438 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
439
440 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
441
442 ## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done
443 ##
444 ## Reliability analysis
445 ## Call: alpha(x = self.p.wide[, -1])
446 ##
447 ##   raw_alpha std.alpha G6(smc) average_r S/N ase mean   sd median_r
448 ##      0.74      0.74   0.73      0.18 2.8 0.1   2.7 0.42      0.18
449 ##
450 ##   lower alpha upper      95% confidence boundaries
451 ## 0.54 0.74 0.95
452 ##
453 ## Reliability if an item is dropped:
454 ##
455 ##           raw_alpha std.alpha G6(smc) average_r S/N var.r med.r
456 ## phb_alleangesehen    0.71    0.70    0.74      0.16 2.3  0.20 0.159
457 ## phb_augen            0.70    0.70    0.72      0.16 2.4  0.18 0.159
458 ## phb_blick            0.66    0.66    0.67      0.14 1.9  0.17 0.145
459 ## phb_gestik           0.74    0.73    0.75      0.19 2.8  0.20 0.178
460 ## phb_raum             0.74    0.72    0.72      0.18 2.6  0.20 0.178
461 ## phb_stand            0.65    0.65    0.67      0.14 1.9  0.17 0.129

```

```

461 ## phb_vorsichgeht      0.68      0.66      0.68      0.14 1.9  0.18 0.096
462 ## pi_direkt            0.76      0.76      0.75      0.21 3.1  0.21 0.273
463 ## pi_nonverbal         0.84      0.83      0.81      0.29 4.9  0.12 0.291
464 ## pi_zubewegen         0.76      0.76      0.78      0.21 3.1  0.21 0.214
465 ## ps_deutlich          0.71      0.70      0.77      0.17 2.4  0.20 0.145
466 ## ps_impulse           0.74      0.72      0.74      0.18 2.6  0.21 0.202
467 ## ps_klar              0.68      0.66      0.69      0.14 2.0  0.19 0.130
468 ##
469 ## Item statistics
470 ##          n raw.r std.r  r.cor r.drop mean  sd
471 ## phb_alleangesehen 8  0.617  0.65  0.644  0.495  3.0 0.93
472 ## phb_augen          8  0.673  0.62  0.587  0.545  3.4 1.06
473 ## phb_blick          8  0.908  0.88  0.845  0.872  2.8 0.89
474 ## phb_gestik         8  0.453  0.38  0.257  0.280  2.6 1.06
475 ## phb_raum           8  0.439  0.49  0.389  0.296  1.8 0.89
476 ## phb_stand          8  0.956  0.92  0.891  0.935  2.9 0.99
477 ## phb_vorsichgeht    8  0.893  0.87  0.834  0.859  2.5 0.76
478 ## pi_direkt          8  0.147  0.18  0.023  0.019  3.2 0.71
479 ## pi_nonverbal       8 -0.723 -0.71 -1.014 -0.789  2.4 0.92
480 ## pi_zubewegen       8  0.075  0.17  0.072 -0.019  2.6 0.52
481 ## ps_deutlich        8  0.638  0.60  0.622  0.534  2.9 0.83
482 ## ps_impulse         8  0.376  0.47  0.452  0.291  2.6 0.52
483 ## ps_klar            8  0.855  0.86  0.819  0.805  2.9 0.83
484 ##
485 ## Non missing response frequency for each item
486 ##          1    2    3    4 miss
487 ## phb_alleangesehen 0.12 0.00 0.62 0.25    0

```

488	## phb_augen	0.12	0.00	0.25	0.62	0
489	## phb_blick	0.12	0.12	0.62	0.12	0
490	## phb_gestik	0.12	0.38	0.25	0.25	0
491	## phb_raum	0.50	0.25	0.25	0.00	0
492	## phb_stand	0.12	0.12	0.50	0.25	0
493	## phb_vorsichgeht	0.12	0.25	0.62	0.00	0
494	## pi_direkt	0.00	0.12	0.50	0.38	0
495	## pi_nonverbal	0.12	0.50	0.25	0.12	0
496	## pi_zubewegen	0.00	0.38	0.62	0.00	0
497	## ps_deutlich	0.00	0.38	0.38	0.25	0
498	## ps_impulse	0.00	0.38	0.62	0.00	0
499	## ps_klar	0.00	0.38	0.38	0.25	0

Table 2

Scale analysis for teachers' self-assessment

scale	M	SD	Min	Max	Skewness	Kurtosis
Activation and support	2.46	1.14	1.00	4.00	0.10	1.65
Clarity and structuredness	2.81	0.91	1.00	4.00	-0.17	2.17
Classroom management	2.44	0.88	1.00	4.00	0.19	2.38
Natural behaviour	2.83	0.76	2.00	4.00	0.28	1.84
Positive climate and motivation	2.85	0.87	1.00	4.00	-0.49	2.65
Presence: posture/gaze	2.70	1.01	1.00	4.00	-0.44	2.15
Presence: verbal and non-verbal intervention	2.75	0.79	1.00	4.00	-0.07	2.48
Presence: voice	2.79	0.72	2.00	4.00	0.32	2.03

Table 3 provides an overview over the mean, the standard deviation, the range, Cronbach's Alpha and the Skewness & Kurtosis of all scales for the students' perception of

502 the teacher's behaviour in class.

503 ##

504 ## Reliability analysis

505 ## Call: alpha(x = stud.as.wide[, -1])

506 ##

507 ## raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r

508 ## 0.48 0.5 0.41 0.25 0.99 0.17 2.6 0.71 0.22

509 ##

510 ## lower alpha upper 95% confidence boundaries

511 ## 0.14 0.48 0.82

512 ##

513 ## Reliability if an item is dropped:

514 ## raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r

515 ## af_beiträge 0.34 0.35 0.22 0.22 0.55 0.26 NA

516 ## af_nachdenken 0.49 0.54 0.37 0.37 1.18 0.18 NA

517 ## af_wechselseitig 0.26 0.27 0.16 0.16 0.37 0.29 NA

518 ## med.r

519 ## af_beiträge 0.22

520 ## af_nachdenken 0.37

521 ## af_wechselseitig 0.16

522 ##

523 ## Item statistics

524 ## n raw.r std.r r.cor r.drop mean sd

525 ## af_beiträge 24 0.64 0.72 0.49 0.35 3.3 0.75

526 ## af_nachdenken 24 0.64 0.65 0.31 0.23 2.6 0.97

527 ## af_wechselseitig 24 0.82 0.75 0.55 0.37 2.0 1.25

528 ##

```

529 ## Non missing response frequency for each item
530 ##           1     2     3     4 miss
531 ## af_beträge      0.04 0.04 0.50 0.42    0
532 ## af_nachdenken   0.12 0.33 0.33 0.21    0
533 ## af_wechselseitig 0.58 0.00 0.25 0.17    0

534 ##

535 ## Reliability analysis
536 ## Call: alpha(x = stud.cs.wide[, -1])
537 ##

538 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
539 ##       0.62      0.64    0.47      0.47 1.8 0.15   3.6 0.58     0.47

540 ##

541 ##   lower alpha upper      95% confidence boundaries
542 ## 0.34 0.62 0.91
543 ##

544 ## Reliability if an item is dropped:
545 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
546 ## ks_sichtbar      0.64    0.47    0.22      0.47 0.9     NA    0 0.47
547 ## ks_verständlich   0.35    0.47    0.22      0.47 0.9     NA    0 0.47

548 ##

549 ## Item statistics
550 ##           n raw.r std.r r.cor r.drop mean   sd
551 ## ks_sichtbar  24 0.90 0.86 0.59 0.47 3.5 0.78
552 ## ks_verständlich 24 0.81 0.86 0.59 0.47 3.6 0.58

553 ##

554 ## Non missing response frequency for each item
555 ##           1     2     3     4 miss

```



```

556 ## ks_sichtbar      0.04 0.04 0.25 0.67      0
557 ## ks_verständlich 0.00 0.04 0.29 0.67      0

558 ##

559 ## Reliability analysis

560 ## Call: alpha(x = stud.cm.wide[, -1])

561 ##

562 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
563 ##      0.75      0.75    0.73      0.42 2.9 0.085   3.2 0.69      0.37

564 ##

565 ##   lower alpha upper      95% confidence boundaries

566 ## 0.58 0.75 0.91

567 ##

568 ## Reliability if an item is dropped:

569 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se   var.r med.r
570 ## km_aktiv           0.59      0.59    0.50      0.33 1.5   0.143 0.0038 0.35
571 ## km_klar            0.68      0.68    0.59      0.41 2.1   0.114 0.0063 0.37
572 ## km_mitbekommen     0.76      0.76    0.73      0.52 3.2   0.085 0.0319 0.50
573 ## km_ungestört       0.70      0.70    0.67      0.44 2.4   0.109 0.0555 0.36

574 ##

575 ## Item statistics

576 ##           n raw.r std.r r.cor r.drop mean   sd
577 ## km_aktiv   24 0.85 0.85 0.83 0.70 3.0 0.93
578 ## km_klar    24 0.76 0.77 0.70 0.56 3.3 0.86
579 ## km_mitbekommen 24 0.66 0.66 0.45 0.40 3.3 0.92
580 ## km_ungestört 24 0.74 0.74 0.59 0.52 3.1 0.93

581 ##

582 ## Non missing response frequency for each item

```

```

583 ##           1      2      3      4 miss
584 ## km_aktiv      0.04 0.29 0.29 0.38      0
585 ## km_klar       0.04 0.12 0.33 0.50      0
586 ## km_mitbekommen 0.08 0.04 0.33 0.54      0
587 ## km_ungestört  0.08 0.12 0.42 0.38      0

588 ##
589 ## Reliability analysis
590 ## Call: alpha(x = stud.nb.wide[, -1])
591 ##
592 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
593 ##       0.85      0.85      0.82      0.66 5.8 0.056   3.4 0.59      0.61
594 ##
595 ##   lower alpha upper      95% confidence boundaries
596 ## 0.74 0.85 0.96
597 ##
598 ## Reliability if an item is dropped:
599 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
600 ## m_fiktiv      0.71      0.71      0.56      0.56 2.5   0.120   NA 0.56
601 ## m_natürlich    0.89      0.90      0.82      0.82 8.9   0.044   NA 0.82
602 ## m_verhalten    0.75      0.75      0.61      0.61 3.1   0.103   NA 0.61
603 ##
604 ## Item statistics
605 ##           n raw.r std.r r.cor r.drop mean   sd
606 ## m_fiktiv   23  0.92  0.92  0.89   0.79  3.4 0.72
607 ## m_natürlich 23  0.83  0.82  0.64   0.61  3.3 0.71
608 ## m_verhalten 23  0.89  0.90  0.86   0.77  3.5 0.59
609 ##

```

```

610 ## Non missing response frequency for each item
611 ##           2    3    4 miss
612 ## m_fiktiv    0.13 0.35 0.52    0
613 ## m_natürlich 0.13 0.39 0.48    0
614 ## m_verhalten 0.04 0.43 0.52    0

615 ##

616 ## Reliability analysis
617 ## Call: alpha(x = stud.pcm.wide[, -1])
618 ##

619 ##   raw_alpha std.alpha G6(smc) average_r S/N  ase mean  sd median_r
620 ##       0.82       0.81    0.86       0.41 4.2 0.05   3.4 0.6    0.44
621 ##

622 ##   lower alpha upper    95% confidence boundaries
623 ## 0.72 0.82 0.92
624 ##

625 ## Reliability if an item is dropped:
626 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
627 ## lkm_ausreden      0.78    0.77    0.78    0.40 3.3   0.063 0.043
628 ## lkm_freundlich    0.85    0.86    0.86    0.54 5.9   0.048 0.014
629 ## lkm_interesse     0.79    0.76    0.85    0.39 3.2   0.060 0.079
630 ## lkm_kritik        0.81    0.80    0.85    0.44 3.9   0.052 0.062
631 ## lkm_rückmeldungen 0.76    0.74    0.81    0.36 2.8   0.067 0.075
632 ## lkm_überlegen     0.75    0.72    0.76    0.34 2.6   0.070 0.057
633 ##           med.r
634 ## lkm_ausreden    0.41
635 ## lkm_freundlich  0.54
636 ## lkm_interesse   0.42

```

```

637 ## lkm_kritik          0.51
638 ## lkm_rückmeldungen  0.41
639 ## lkm_überlegen      0.43
640 ##
641 ## Item statistics
642 ##
643 ##          n raw.r std.r r.cor r.drop mean  sd
644 ## lkm_ausreden      24  0.77  0.75  0.75  0.66  3.6 0.77
645 ## lkm_freundlich     24  0.32  0.41  0.31  0.20  3.7 0.46
646 ## lkm_interesse     24  0.78  0.76  0.68  0.64  3.2 0.96
647 ## lkm_kritik        24  0.68  0.65  0.54  0.51  3.2 0.88
648 ## lkm_rückmeldungen 24  0.84  0.84  0.80  0.73  3.3 0.92
649 ## lkm_überlegen     24  0.87  0.88  0.90  0.79  3.4 0.83
650 ##
651 ## Non missing response frequency for each item
652 ##          1    2    3    4 miss
653 ## lkm_ausreden  0.04 0.04 0.17 0.75  0
654 ## lkm_freundlich 0.00 0.00 0.29 0.71  0
655 ## lkm_interesse  0.04 0.25 0.21 0.50  0
656 ## lkm_kritik    0.08 0.04 0.46 0.42  0
657 ## lkm_rückmeldungen 0.08 0.04 0.33 0.54  0
658 ## lkm_überlegen  0.04 0.08 0.29 0.58  0
659 ##
660 ## Warning in alpha(stud.ppg.wide[, -1]): Some items were negatively correlated with the
661 ## should be reversed.
662 ## To do this, run the function again with the 'check.keys=TRUE' option
663 ##
664 ## Some items ( phb_stand ) were negatively correlated with the total scale and
665 ## probably should be reversed.

```

```

663 ## To do this, run the function again with the 'check.keys=TRUE' option

664 ##

665 ## Reliability analysis

666 ## Call: alpha(x = stud.ppg.wide[, -1])

667 ##

668 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
669 ##      0.59      0.54    0.74      0.14 1.2 0.11   3.2 0.39      0.17

670 ##

671 ##   lower alpha upper      95% confidence boundaries

672 ## 0.37 0.59 0.81

673 ##

674 ## Reliability if an item is dropped:

675 ##               raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r
676 ## phb_alleangesehen      0.58      0.54    0.74      0.163 1.17    0.119 0.120
677 ## phb_augen              0.59      0.52    0.74      0.156 1.10    0.113 0.138
678 ## phb_blick              0.34      0.26    0.47      0.055 0.35    0.187 0.074
679 ## phb_gestik             0.43      0.36    0.66      0.085 0.56    0.160 0.106
680 ## phb_raum               0.67      0.61    0.76      0.203 1.53    0.086 0.109
681 ## phb_stand              0.68      0.68    0.76      0.259 2.09    0.096 0.073
682 ## phb_vorsichgeht        0.39      0.31    0.54      0.070 0.45    0.171 0.100

683 ##               med.r
684 ## phb_alleangesehen 0.182
685 ## phb_augen         0.135
686 ## phb_blick         0.062
687 ## phb_gestik        0.135
688 ## phb_raum          0.215
689 ## phb_stand         0.215

```

```

690 ## phb_vorsichgeht    0.062
691 ##
692 ## Item statistics
693 ##
694 ##           n raw.r std.r  r.cor r.drop mean  sd
695 ## phb_alleangesehen 24  0.45 0.425  0.287  0.231  3.5 0.66
696 ## phb_augen          24  0.33 0.455  0.291  0.158  3.6 0.49
697 ## phb_blick          24  0.90 0.876  0.972  0.827  3.5 0.72
698 ## phb_gestik         24  0.77 0.748  0.706  0.577  3.0 0.86
699 ## phb_raum           24  0.40 0.256  0.085  0.048  2.1 0.97
700 ## phb_stand          24 -0.14 0.025 -0.130 -0.295  3.7 0.46
701 ## phb_vorsichgeht    24  0.82 0.813  0.880  0.671  3.3 0.82
702 ##
703 ## Non missing response frequency for each item
704 ##           1    2    3    4 miss
705 ## phb_alleangesehen 0.00 0.08 0.29 0.62    0
706 ## phb_augen          0.00 0.00 0.38 0.62    0
707 ## phb_blick          0.04 0.00 0.42 0.54    0
708 ## phb_gestik         0.04 0.25 0.42 0.29    0
709 ## phb_raum           0.33 0.33 0.25 0.08    0
710 ## phb_stand          0.00 0.00 0.29 0.71    0
711 ## phb_vorsichgeht    0.04 0.08 0.38 0.50    0
712 ##
713 ## Warning in alpha(stud.pvni.wide[, -1]): Some items were negatively correlated with th
714 ## should be reversed.
715 ## To do this, run the function again with the 'check.keys=TRUE' option
716 ##
717 ## Some items ( pi_direkt ) were negatively correlated with the total scale and
718 ## probably should be reversed.

```

```

716 ## To do this, run the function again with the 'check.keys=TRUE' option

717 ##

718 ## Reliability analysis

719 ## Call: alpha(x = stud.pvni.wide[, -1])

720 ##

721 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
722 ##      0.2      0.16   0.23      0.062 0.2 0.27   3.2 0.46      0.18
723 ##

724 ##   lower alpha upper      95% confidence boundaries
725 ## -0.34 0.2 0.73
726 ##

727 ## Reliability if an item is dropped:

728 ##           raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r med.r
729 ## pi_direkt      0.40      0.40   0.25      0.25 0.68      0.24   NA 0.25
730 ## pi_nonverbal    0.30      0.30   0.18      0.18 0.44      0.28   NA 0.18
731 ## pi_zubewegen   -0.64     -0.66  -0.25     -0.25 -0.40      0.66   NA -0.25
732 ##

733 ## Item statistics

734 ##           n raw.r std.r r.cor r.drop mean   sd
735 ## pi_direkt  24 0.43 0.51 0.063 -0.040  3.6 0.65
736 ## pi_nonverbal 24 0.59 0.55 0.181  0.038  2.8 0.78
737 ## pi_zubewegen 24 0.80 0.78 0.636  0.356  3.1 0.80
738 ##

739 ## Non missing response frequency for each item

740 ##           1    2    3    4 miss
741 ## pi_direkt  0.00 0.08 0.21 0.71    0
742 ## pi_nonverbal 0.04 0.29 0.50 0.17    0

```

```

743 ## pi_zubewegen 0.04 0.12 0.50 0.33    0

744 ##

745 ## Reliability analysis

746 ## Call: alpha(x = stud.pv.wide[, -1])

747 ##

748 ##   raw_alpha std.alpha G6(smc) average_r S/N ase mean   sd median_r
749 ##      0.71      0.72      0.65      0.46 2.6 0.1   3.6 0.48      0.44

750 ##

751 ##   lower alpha upper      95% confidence boundaries

752 ## 0.51 0.71 0.92

753 ##

754 ## Reliability if an item is dropped:

755 ##           raw_alpha std.alpha G6(smc) average_r   S/N alpha se var.r med.r
756 ## ps_deutlich      0.61      0.62      0.44      0.44 1.60      0.16   NA  0.44
757 ## ps_impulse      0.76      0.76      0.61      0.61 3.08      0.10   NA  0.61
758 ## ps_klar        0.49      0.50      0.33      0.33 0.98      0.20   NA  0.33

759 ##

760 ## Item statistics

761 ##           n raw.r std.r r.cor r.drop mean   sd
762 ## ps_deutlich 24  0.79  0.81  0.67   0.54  3.6 0.58
763 ## ps_impulse  24  0.76  0.74  0.50   0.43  3.5 0.66
764 ## ps_klar     24  0.84  0.85  0.77   0.64  3.6 0.58

765 ##

766 ## Non missing response frequency for each item

767 ##           2    3    4 miss
768 ## ps_deutlich 0.04 0.29 0.67    0
769 ## ps_impulse  0.08 0.29 0.62    0

```



```

770 ## ps_klar      0.04 0.29 0.67    0

771 ## Warning in alpha(stud.p.wide[, -1]): Some items were negatively correlated with the t
772 ## should be reversed.
773 ## To do this, run the function again with the 'check.keys=TRUE' option

774 ## Some items ( phb_stand ) were negatively correlated with the total scale and
775 ## probably should be reversed.
776 ## To do this, run the function again with the 'check.keys=TRUE' option

777 ##
778 ## Reliability analysis
779 ## Call: alpha(x = stud.p.wide[, -1])
780 ##
781 ##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
782 ##      0.71      0.67    0.89      0.13   2 0.081   3.3 0.33      0.11
783 ##
784 ##   lower alpha upper      95% confidence boundaries
785 ## 0.55 0.71 0.87
786 ##
787 ## Reliability if an item is dropped:
788 ##           raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r
789 ## phb_alleangesehen    0.68    0.63    0.86    0.125 1.7   0.088 0.079
790 ## phb_augen            0.72    0.68    0.90    0.150 2.1   0.080 0.083
791 ## phb_blick            0.63    0.57    0.83    0.099 1.3   0.106 0.068
792 ## phb_gestik           0.65    0.60    0.86    0.110 1.5   0.100 0.072
793 ## phb_raum             0.70    0.65    0.87    0.135 1.9   0.081 0.078
794 ## phb_stand            0.74    0.73    0.90    0.182 2.7   0.073 0.066

```

```

795 ## phb_vorsichgeht      0.65      0.60      0.85      0.110 1.5      0.099 0.072
796 ## pi_direkt            0.72      0.69      0.90      0.155 2.2      0.078 0.078
797 ## pi_nonverbal         0.67      0.62      0.86      0.122 1.7      0.093 0.070
798 ## pi_zubewegen         0.70      0.66      0.87      0.141 2.0      0.082 0.079
799 ## ps_deutlich          0.72      0.67      0.89      0.147 2.1      0.077 0.078
800 ## ps_impulse           0.67      0.61      0.87      0.115 1.6      0.092 0.079
801 ## ps_klar              0.70      0.65      0.85      0.133 1.8      0.081 0.076
802 ##                      med.r
803 ## phb_alleangesehen 0.097
804 ## phb_augen            0.137
805 ## phb_blick            0.062
806 ## phb_gestik           0.101
807 ## phb_raum             0.101
808 ## phb_stand            0.181
809 ## phb_vorsichgeht     0.093
810 ## pi_direkt            0.152
811 ## pi_nonverbal         0.101
812 ## pi_zubewegen         0.093
813 ## ps_deutlich          0.152
814 ## ps_impulse           0.084
815 ## ps_klar              0.114
816 ##
817 ## Item statistics
818 ##                      n raw.r std.r r.cor r.drop mean   sd
819 ## phb_alleangesehen 24  0.52  0.53  0.53  0.401  3.5 0.66
820 ## phb_augen          24  0.19  0.25  0.16  0.077  3.6 0.49
821 ## phb_blick           24  0.84  0.83  0.85  0.781  3.5 0.72

```

```

822 ## phb_gestik          24  0.72  0.70  0.69  0.606  3.0 0.86
823 ## phb_raum            24  0.49  0.42  0.40  0.295  2.1 0.97
824 ## phb_stand           24 -0.21 -0.11 -0.18 -0.308  3.7 0.46
825 ## phb_vorsichgeht     24  0.72  0.70  0.72  0.608  3.3 0.82
826 ## pi_direkt           24  0.22  0.19  0.12  0.075  3.6 0.65
827 ## pi_nonverbal        24  0.62  0.57  0.57  0.490  2.8 0.78
828 ## pi_zubewegen        24  0.41  0.35  0.33  0.244  3.1 0.80
829 ## ps_deutlich         24  0.20  0.28  0.23  0.066  3.6 0.58
830 ## ps_impulse          24  0.63  0.65  0.62  0.526  3.5 0.66
831 ## ps_klar             24  0.37  0.44  0.44  0.250  3.6 0.58
832 ##
833 ## Non missing response frequency for each item
834 ##                    1    2    3    4 miss
835 ## phb_alleangesehen 0.00 0.08 0.29 0.62    0
836 ## phb_augen         0.00 0.00 0.38 0.62    0
837 ## phb_blick         0.04 0.00 0.42 0.54    0
838 ## phb_gestik        0.04 0.25 0.42 0.29    0
839 ## phb_raum          0.33 0.33 0.25 0.08    0
840 ## phb_stand         0.00 0.00 0.29 0.71    0
841 ## phb_vorsichgeht   0.04 0.08 0.38 0.50    0
842 ## pi_direkt         0.00 0.08 0.21 0.71    0
843 ## pi_nonverbal      0.04 0.29 0.50 0.17    0
844 ## pi_zubewegen      0.04 0.12 0.50 0.33    0
845 ## ps_deutlich       0.00 0.04 0.29 0.67    0
846 ## ps_impulse        0.00 0.08 0.29 0.62    0
847 ## ps_klar           0.00 0.04 0.29 0.67    0

```

Table 3

Scale analysis for students' perspective

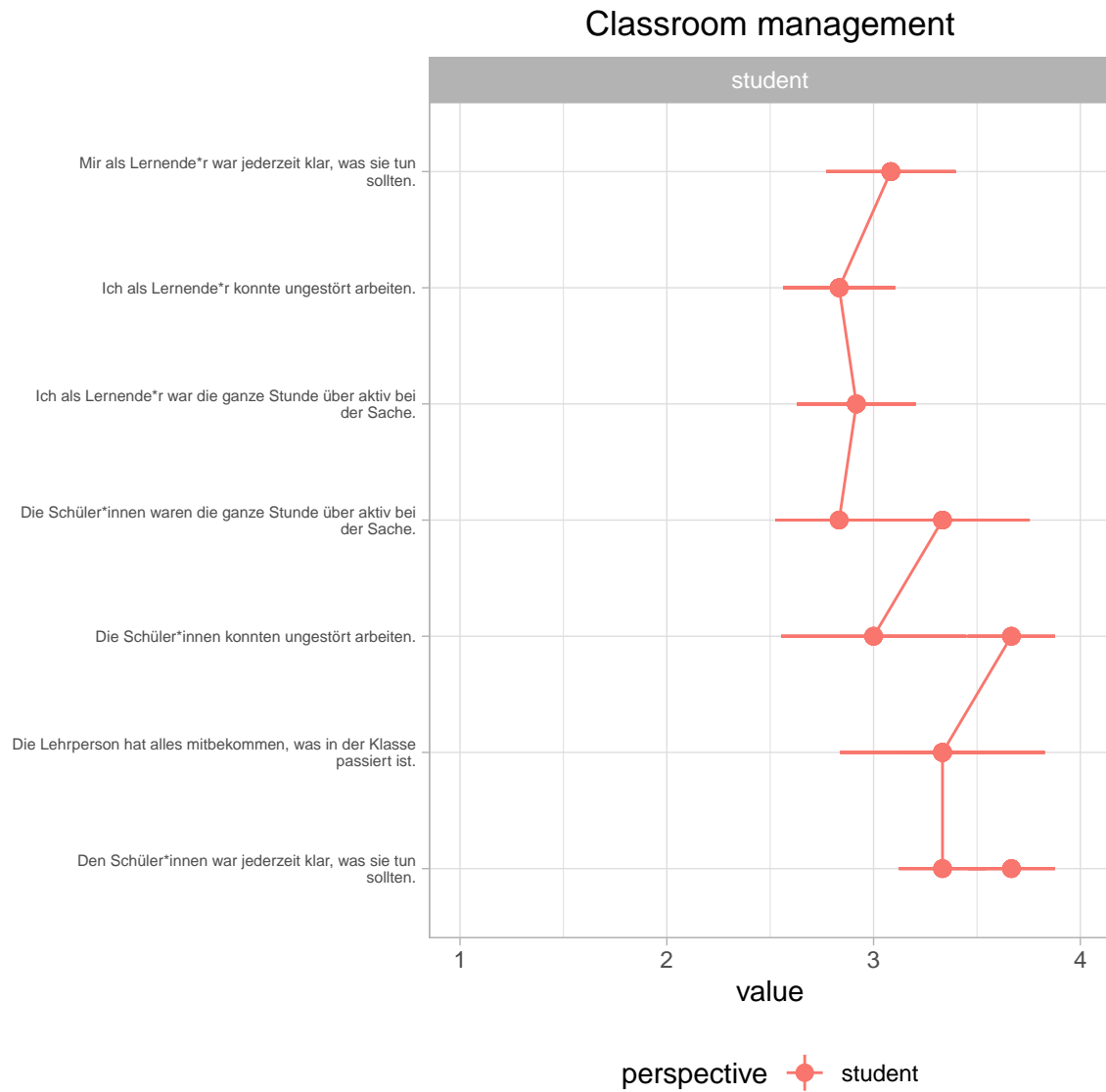
scale	M	SD	Min	Max	Skewness	Kurtosis
Activation and support	2.64	1.13	1.00	4.00	-0.32	1.73
Clarity and structuredness	3.58	0.68	1.00	4.00	-1.75	6.09
Classroom management	3.18	0.91	1.00	4.00	-0.87	2.86
Natural behaviour	3.41	0.67	2.00	4.00	-0.68	2.38
Positive climate and motivation	3.41	0.83	1.00	4.00	-1.40	4.29
Presence: posture/gaze	3.24	0.89	1.00	4.00	-1.05	3.31
Presence: verbal and non-verbal intervention	3.18	0.81	1.00	4.00	-0.66	2.70
Presence: voice	3.60	0.60	2.00	4.00	-1.18	3.36

848

The individual items of a scale are further represented in graphs.

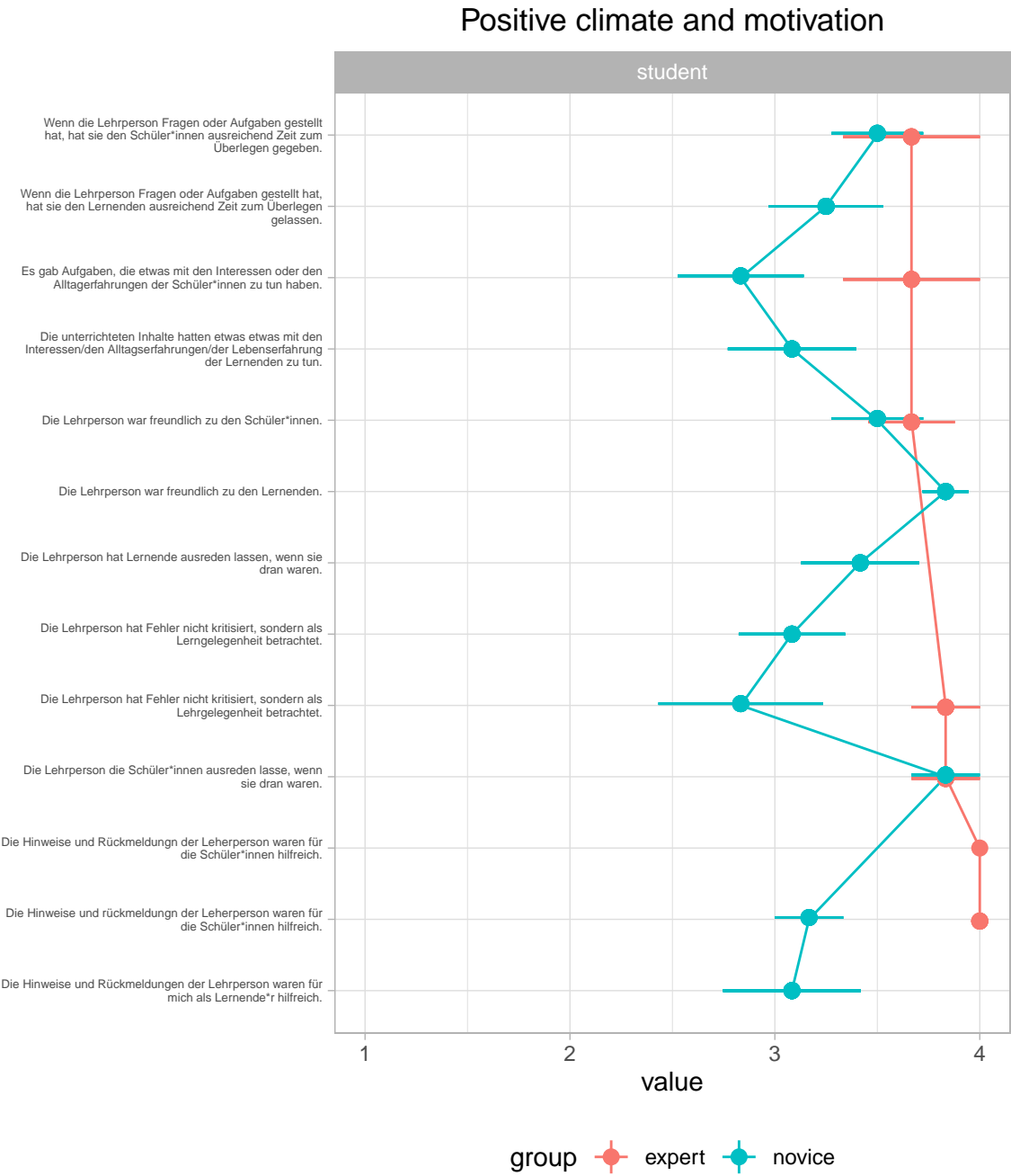
849

(1) Classroom management

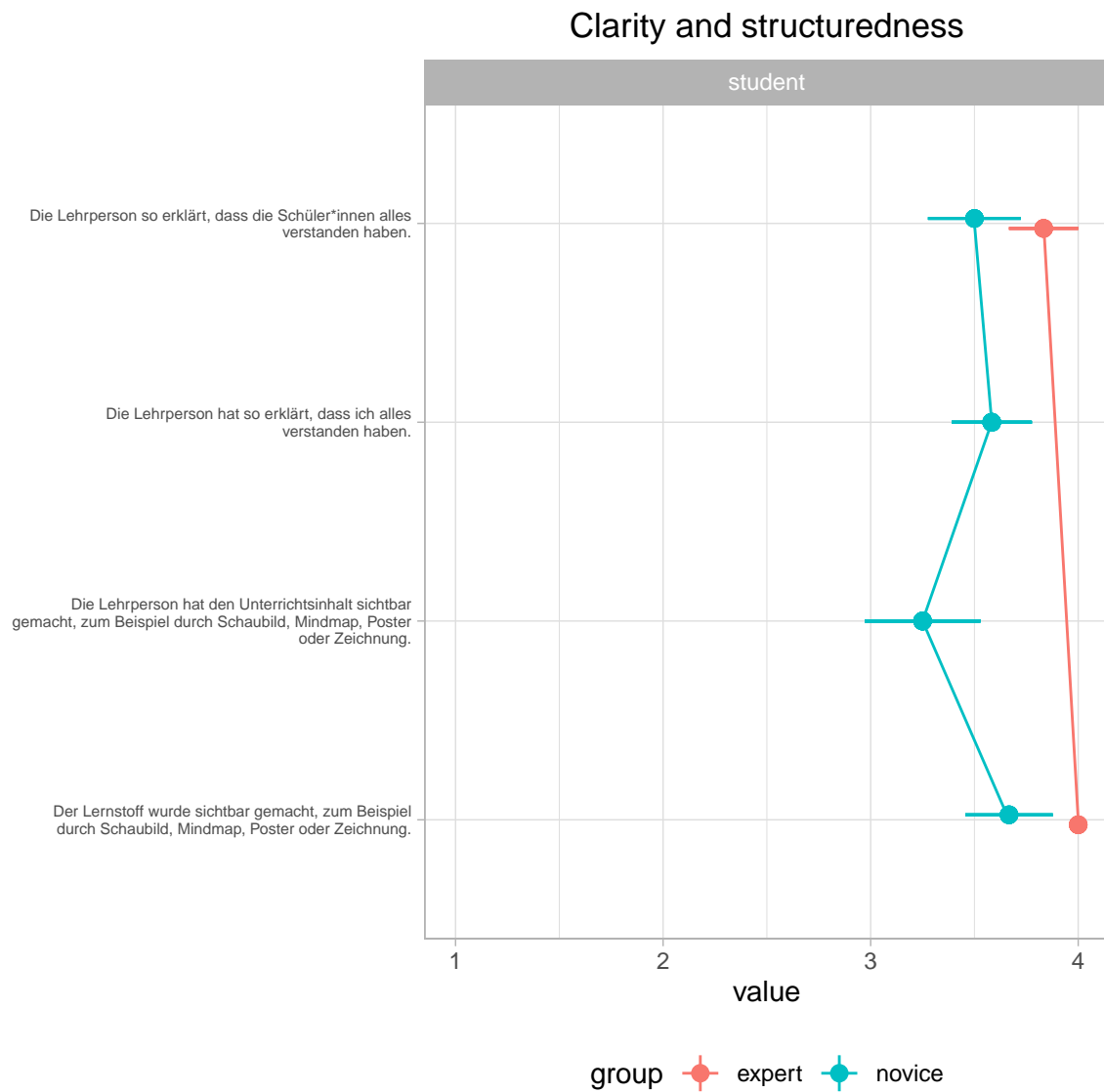


851

(2) Positive climate and motivation



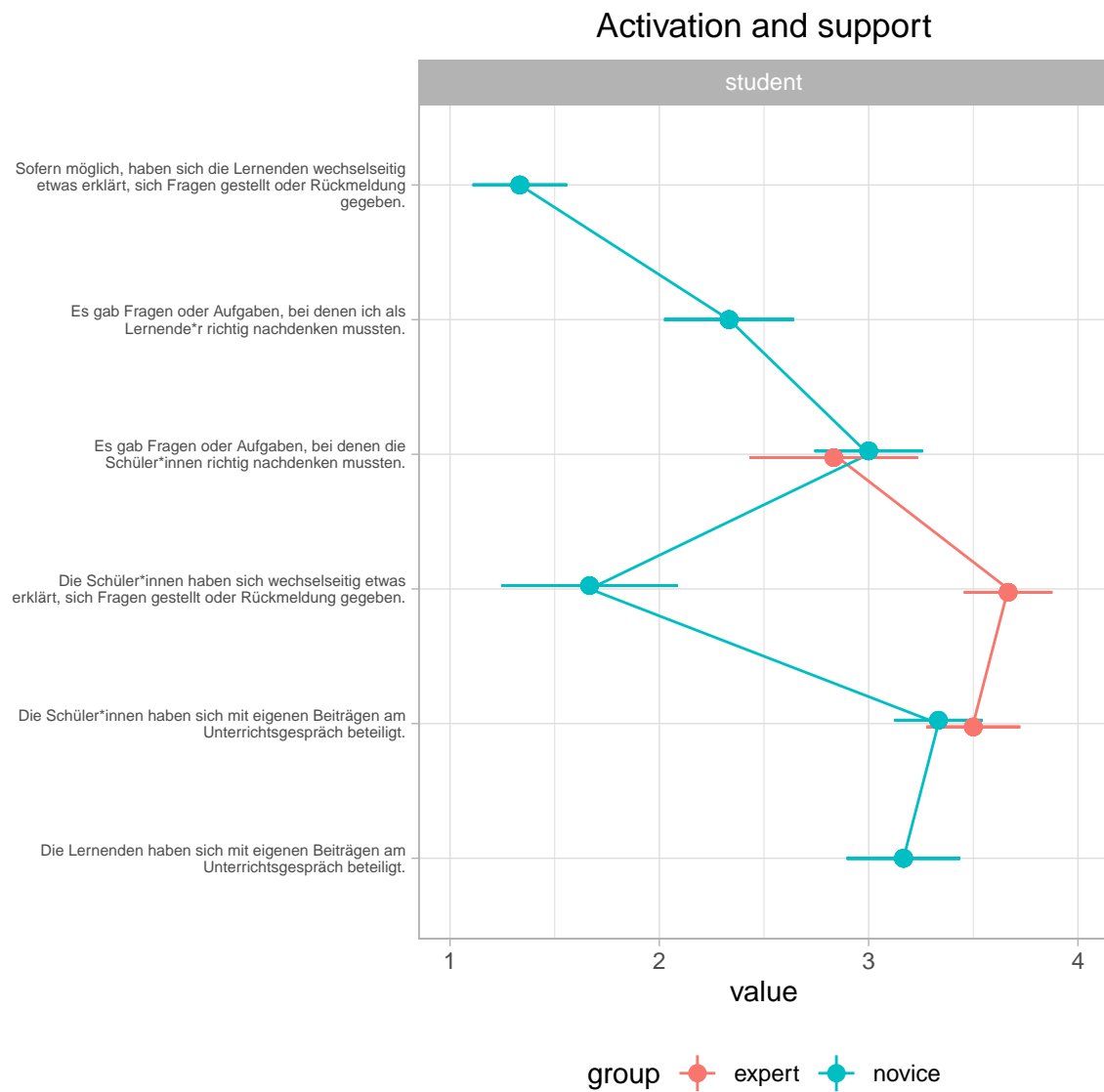
853 (3) Clarity and structuredness



854

855

(4) Activation and support



856

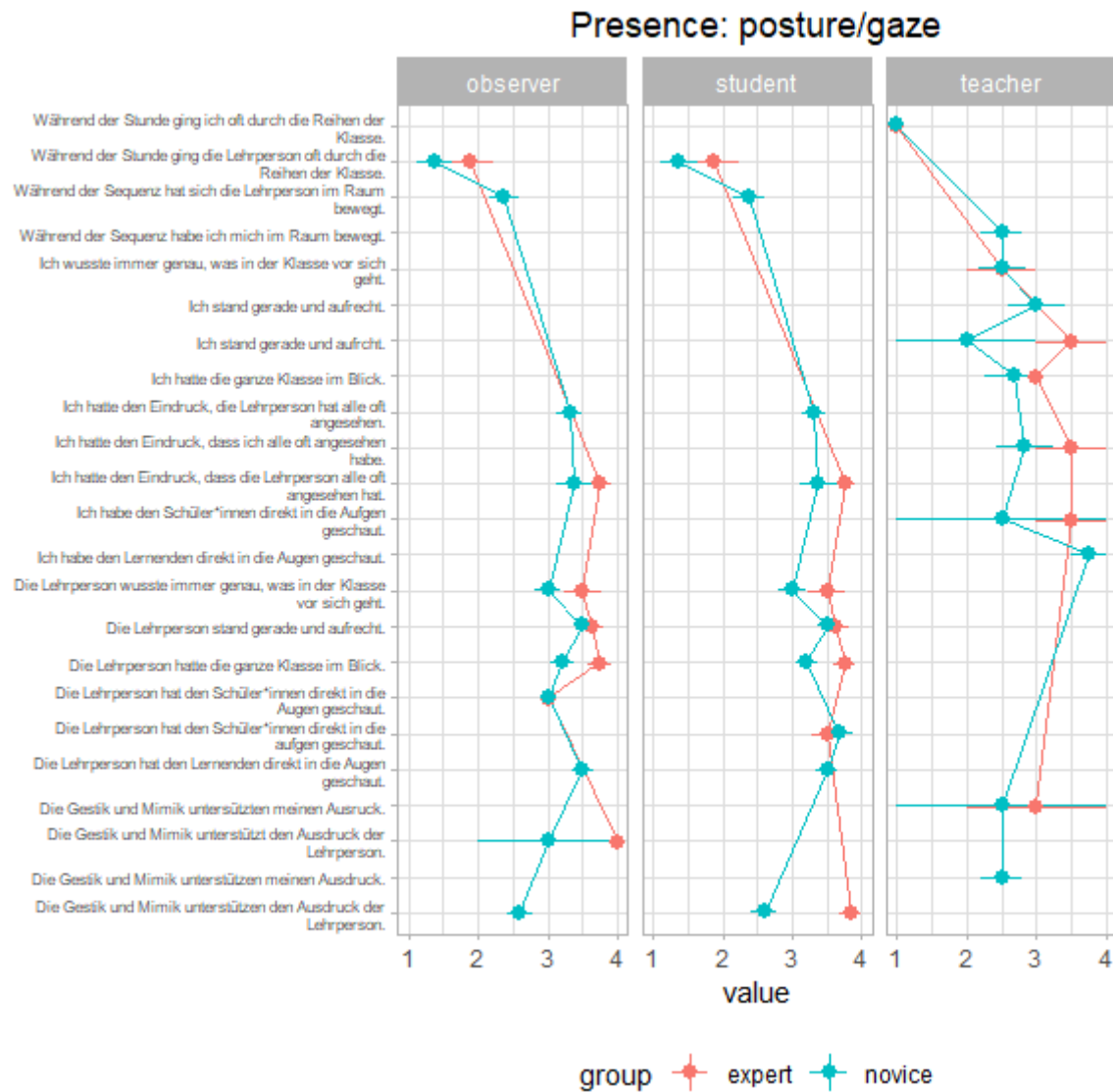


Figure 2. (5) Presence: posture/gaze

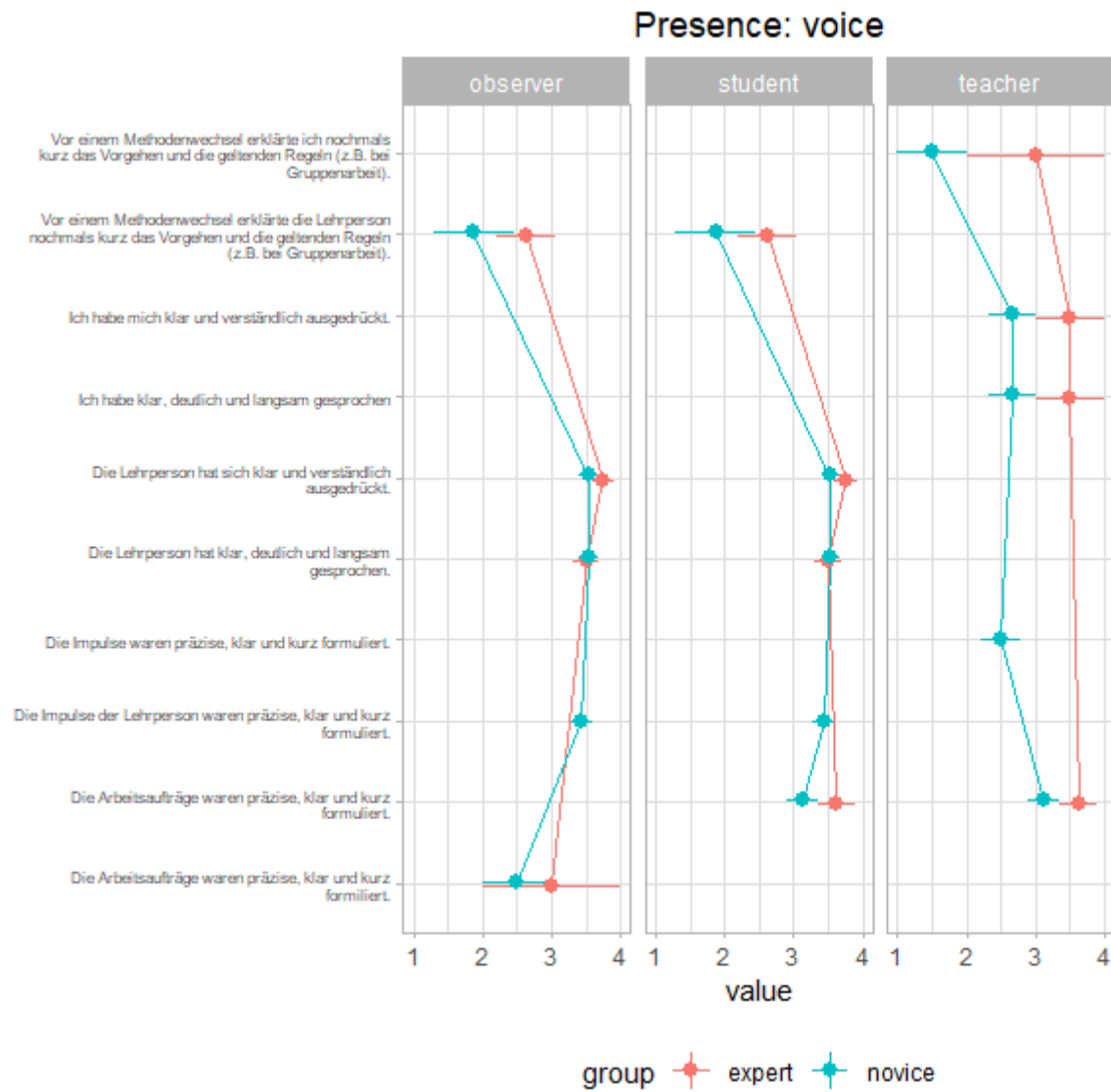


Figure 3. (6) Presence: voice

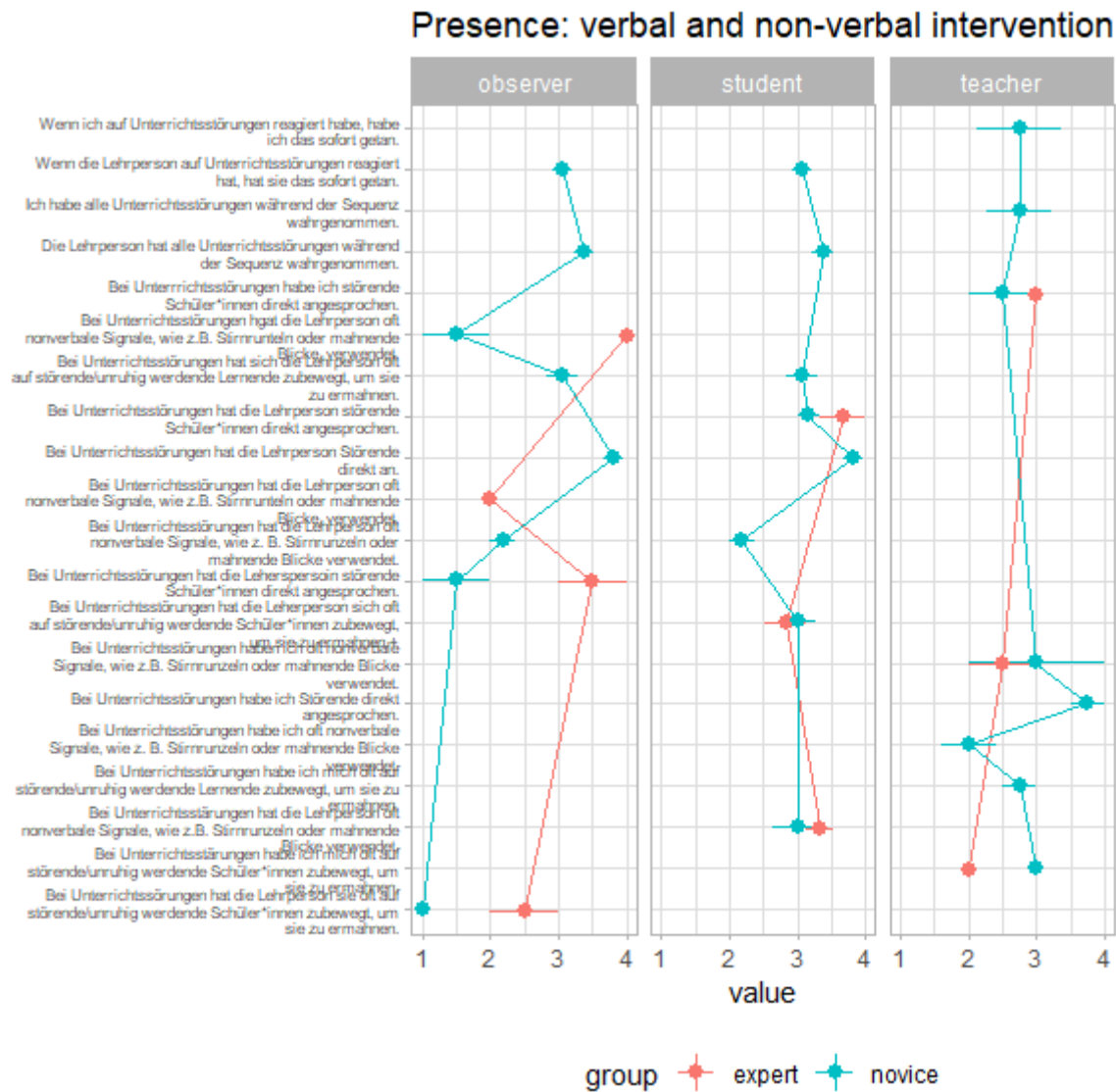
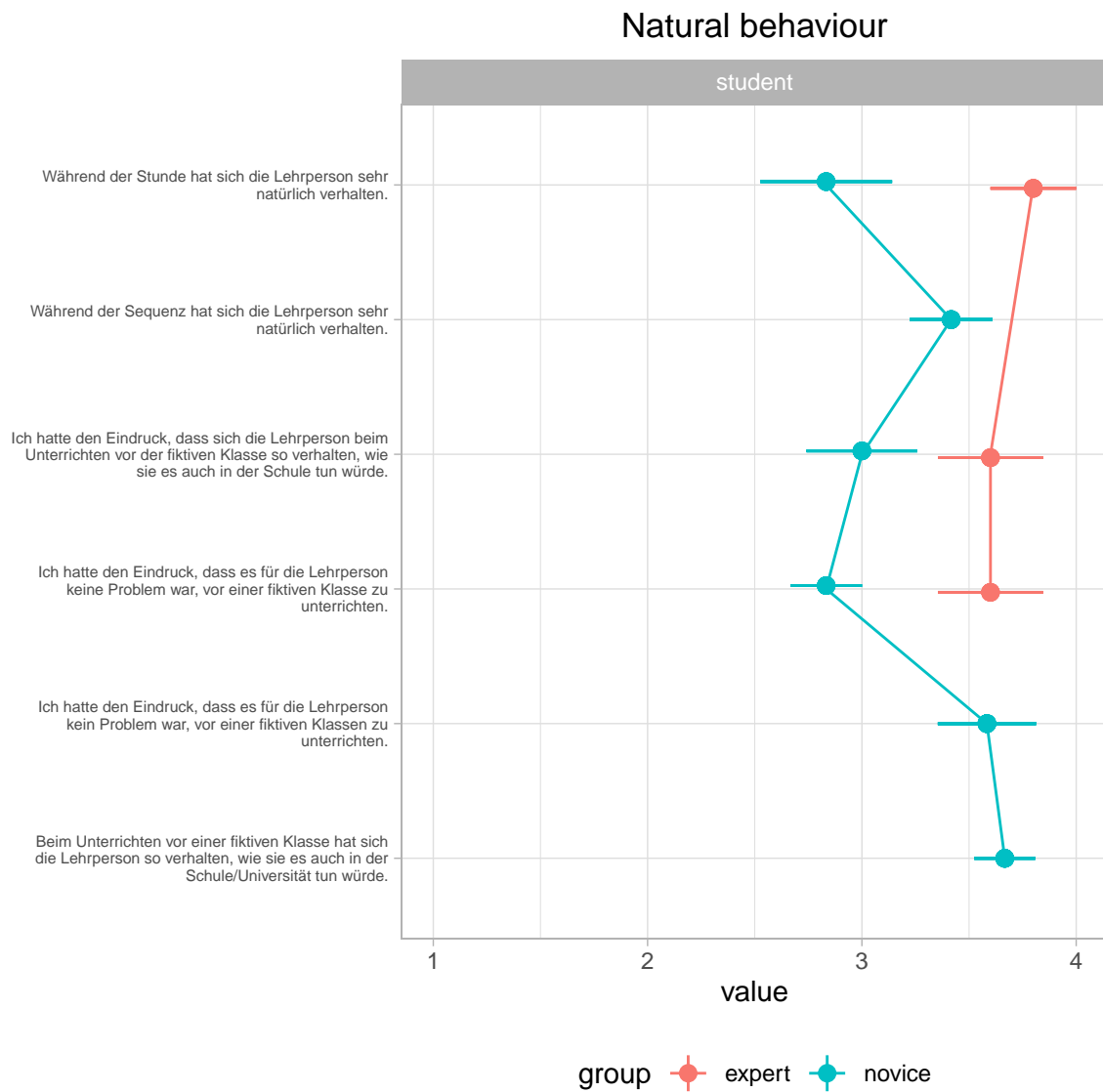


Figure 4. (7) Presence: verbal and non-verbal intervention

857 (8) Natural behaviour



859 In addition, we plotted all scales. Graph provides boxplots and individual data for
860 experts and novices.

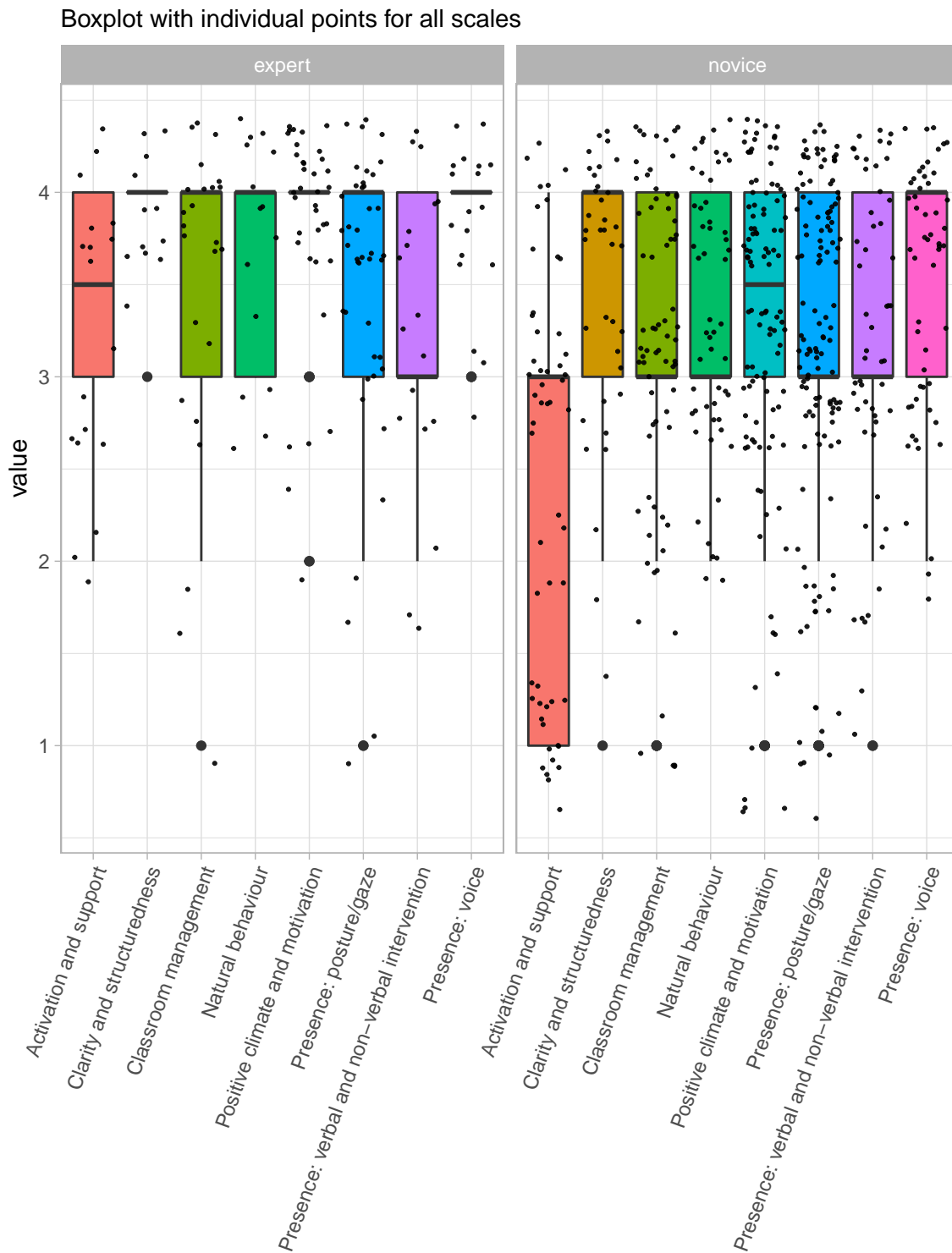


Figure 5. (#fig:boxplot scales) Boxplots and individual data for experts and novices

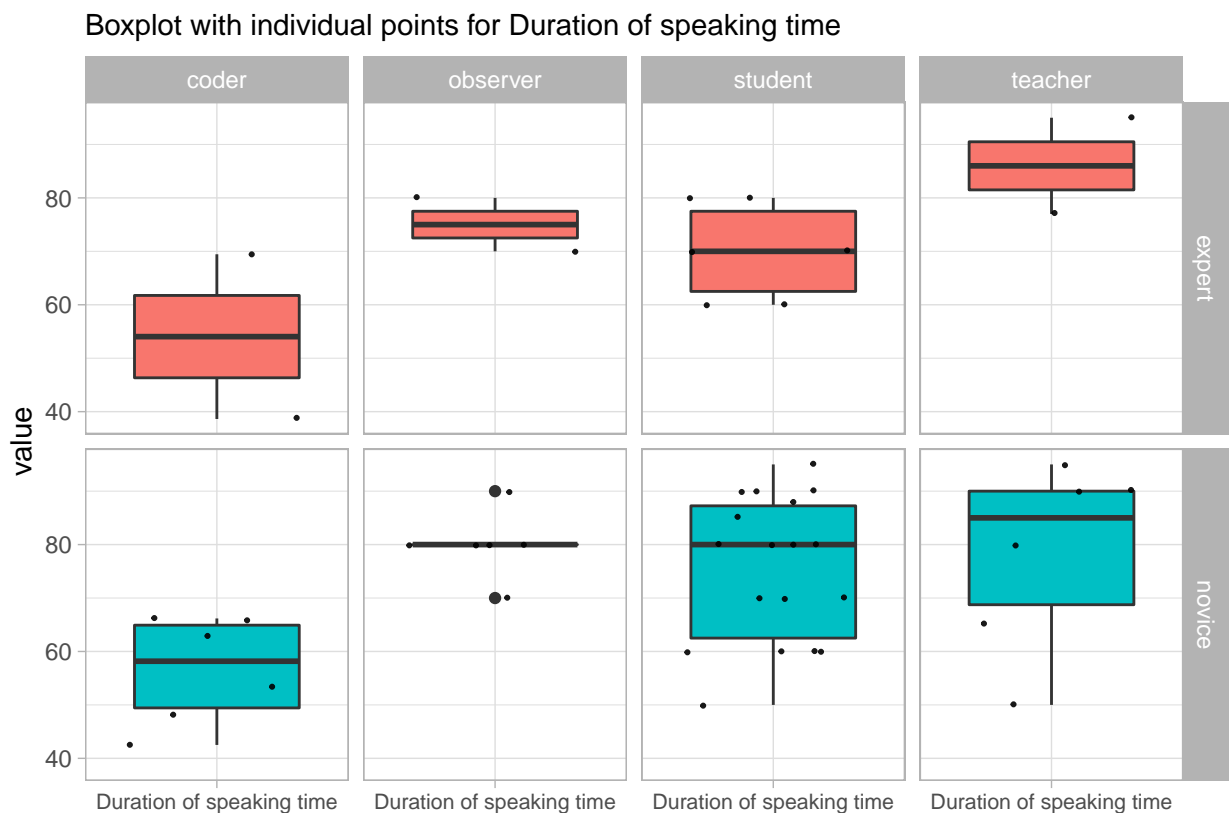
Behavioral Data. The recorded lessons were coded in a post-hoc procedure with the coding software MAXQDA by previously trained raters (Kuckartz & Rädiker, 2019). The following coding scheme was developed:

- phase - lesson begin, state event: teacher starts the lesson with a noise, talk, taking a position in class
- phase - lesson end, state event: teacher finishes the lesson with a noise, talk, taking a position in class
- phase - organization/transition points, state event: any situation that does not imply effective learning time (fetching chalk, working material, organizing desks, opening windows, printing work results etc.)
- phase - single, state event: any individual student activity on a given task (reading, writing, drawing etc.)
- phase - group, state event: any student activity on a given task together in a group of at least 3 students (reading, writing, drawing etc.)
- phase - class discussion, state event: discussion in class, teacher talks to class/individual/group
- phase - pair: state event: any student activity on a given task together in a team of 2 students (reading, writing, drawing etc.)
- phase - teachers lecture, state event: any teacher's presentation on a certain topic which maybe supported by a PPP, PREZI, notes on board, OHP etc.
- phase - other, state event: not categorizable
- phase - break, state event: e.g. drinking, relaxation exercises

- 883 • phase - external interruption, state event: external interruptions (e.g. fire alarm,
884 technical problems, other teachers coming into the room)
- 885 • speaking time - teacher, state event
- 886 • speaking time - students, state event
- 887 • disruption - chatting with neighbor, state event (perceived/ not perceived, reacted:
888 verbal, non-verbal/ not reacted)
- 889 • disruption - asking a question, state event (perceived/ not perceived, reacted: verbal,
890 non-verbal/ not reacted)
- 891 • disruption - yelling, state event (perceived/ not perceived, reacted: verbal,
892 non-verbal/ not reacted)
- 893 • disruption - looking at phone, state event (perceived/ not perceived, reacted: verbal,
894 non-verbal/ not reacted)
- 895 • disruption - staring out of window, state event (perceived/ not perceived, reacted:
896 verbal, non-verbal/ not reacted)
- 897 • disruption - drawing, state event (perceived/ not perceived, reacted: verbal,
898 non-verbal/ not reacted)
- 899 • disruption - head on table, state event (perceived/ not perceived, reacted: verbal,
900 non-verbal/ not reacted)
- 901 • disruption - clicking pen, state event (perceived/ not perceived, reacted: verbal,
902 non-verbal/ not reacted)
- 903 • disruption - drumming hands, state event (perceived/ not perceived, reacted: verbal,
904 non-verbal/ not reacted)

- disruption - walking around, state event (perceived/ not perceived, reacted: verbal, non-verbal/ not reacted)

First, we coded the speaking time of the teacher and the students to compare all perspectives: coder, observer, students, teacher. The graph below shows the result of the coded speaking duration compared to the estimated speaking duration assessed with the questionnaire.



Eyetracking Data. The Tobii Pro Lab 2 software was used to analyze the teachers' visual attention during each mini-lesson. The software allows for non-screen based recordings of a participants' attention while moving in real-world settings. The recordings of the glasses contain both HD-video from the subjects' perspective as well as the respective gaze data mapped onto the video. In order to map multiple recordings to AOIs, we first imported the eye-tracking recordings into the Tobii Pro Analyzer software. Second, we created dynamic Areas of Interest (AOI) manually to plot the gaze data. Once the

AOIs are created, the gaze recordings of multiple recordings can be mapped and analyzed in aggregated form. Tobii Pro does not allow to do AOI based analyses within Pro Lab. So we exported a tsv. file to do further analyses in the software R.

Gaze relational index (GRI).

The GRI is a measure of visual expertise in information processing. This metric is calculated as the ratio of mean fixation duration to fixation count. The GRI is higher for novices than for experts. (Gegenfurtner et al., 2020)

Table 4

Number and Duration (in msec) of Fixations

Participant	Variable	Fixation Number	Fixation Duration	M Duration Fixation	TOI	GRI
01_01_D	Expert	803.00	316,571.00	394.00	781,978.00	0.49
01_02_A	Expert	1,070.00	385,812.00	361.00	838,026.00	0.34
01_03_B	Novice	617.00	374,315.00	607.00	744,444.00	0.98
01_04_C	Novice	769.00	384,537.00	500.00	723,922.00	0.65
02_01_A	Novice	569.00	101,541.00	178.00	729,762.00	0.31
02_02_B	Novice	1,140.00	520,431.00	457.00	730,565.00	0.40
02_03_C	Novice	1,048.00	469,018.00	448.00	737,604.00	0.43
02_04_D	Novice	613.00	438,655.00	716.00	747,729.00	1.17

Data analysis

We used R [Version 4.0.3; R Core Team (2019)] and the R-packages *dplyr* [R-dplyr], *forcats* [Version 0.5.0; Wickham (2020a)], *ggplot2* [Version 3.3.2; Wickham (2016)], *moments* [Version 0.14; Komsta and Novomestky (2015)], *papaja* [Version 0.1.0.9997; Aust and Barth (2020)], *papayar* (Muschelli, 2016), *psych* [Version 2.0.12; Revelle (2020)], *purrr* [Version 0.3.4; Henry and Wickham (2020)], *readr* [Version 1.4.0; Wickham, Hester, and Francois (2018)], *sjPlot* [Version 2.8.7; Lüdecke (2021)], *stringr* [Version 1.4.0; Wickham (2019)], *tibble* [Version 3.0.4; Müller and Wickham (2021)], *tidyr*

Table 5

Number and Duration (in msec) of Fixations during calibration

Participant	Variable1	Fixation Number	Fixation Duration	M Duration Fixation	TOI	GRI
01_01_D	Expert	9.00	14,372.00	1,597.00	16,470.00	177.44
01_02_A	Expert	10.00	10,194.00	1,019.00	13,335.00	101.90
01_03_B	Novice	17.00	9,234.00	543.00	10,615.00	31.94
01_04_C	Novice	14.00	15,311.00	1,094.00	17,224.00	78.14
02_01_A	Novice	13.00	5,157.00	397.00	17,902.00	30.54
02_02_B	Novice	12.00	10,654.00	888.00	12,325.00	74.00
02_03_C	Novice	18.00	14,151.00	786.00	16,494.00	43.67
02_04_D	Novice	14.00	19,128.00	1,366.00	20,964.00	97.57

[Version 1.1.2; Wickham (2020b)], and *tidyverse* [Version 1.3.0; Wickham et al. (2019)] for all our analyses.

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Questionnaire Data.

Behavioral Data.

Eyetracking Data.

Results

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos

et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Questionnaire Data.

Behavioral Data.

Eyetracking Data.

Discussion

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

References

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*.

Retrieved from <https://github.com/crsh/papaja>

Barnes, D. (2004). The significance of teachers' frames for teaching. In *Teachers and teaching* (pp. 16–38). Routledge.

Barth, V. L. (2017). *Professionelle wahrnehmung von störungen im unterricht*. Springer.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, 35(5), 463–482.

Bogert, N. J. van den. (2016). *On teachers' visual perception and interpretation of classroom events using eye tracking and collaborative tagging methodologies*. Technische Universiteit Eindhoven.

Brophy, J. (1986). Classroom management techniques. *Education and Urban Society*, 18(2), 182–194.

Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science and Mathematics Education*, 13(2), 389–403.

Gegenfurtner, A., Boucheix, J.-M., Gruber, H., Hauser, F., Lehtinen, E., & Lowe, R. K. (2020). The gaze relational index as a measure of visual expertise. *Journal of Expertise*. East Lansing, Michigan: Michigan State University, 2020, Vol. 3, Iss. 1.

Helmke, A., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., Schrader, F.-W., & AdeThurow, M. (2014). Unterrichtsdiagnostik mit EMU. *Aus-Und Fortbildung Der Lehrkräfte in Hinblick Auf Verbesserung Der Diagnosefähigkeit, Umgang*

Mit Heterogenität Und Individuelle Förderung, 149–163.

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154.

Kiel, E., Frey, A., Weiß, S., & Weiss, S. (2013). *Trainingsbuch klassenführung* (Vol. 3992). UTB.

Komsta, L., & Novomestky, F. (2015). *Moments: Moments, cumulants, skewness, kurtosis and related tests*. Retrieved from <https://CRAN.R-project.org/package=moments>

Kounin, J. S. (2006). *Techniken der klassenführung*. Waxmann Verlag.

Kuckartz, U., & Rädiker, S. (2019). *Analyzing qualitative data with MAXQDA*. Springer.

Lachner, A., Jarodzka, H., & Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instructional Science, 44*(3), 197–203.

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science*. Retrieved from <https://CRAN.R-project.org/package=sjPlot>

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Ascd.

Messner, H., & Reusser, K. (2000). Die berufliche entwicklung von lehrpersonen als lebenslanger prozess. *Beiträge Zur Lehrerinnen-Und Lehrerbildung, 18*(2), 157–171.

Muschelli, J. (2016). *Papayar: View medical research images using the papaya*

JavaScript library. Retrieved from

<https://CRAN.R-project.org/package=papayar>

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from

<https://CRAN.R-project.org/package=tibble>

Nolting, H.-P. (2012). *Störungen in der schulklasse*. Beltz.

Pouta, M., Lehtinen, E., & Palonen, T. (2020). Student teachers' and experienced

teachers' professional vision of students' understanding of the rational number

concept.

R Core Team. (2019). *R: A language and environment for statistical computing*.

Vienna, Austria: R Foundation for Statistical Computing. Retrieved from

<https://www.R-project.org/>

Revelle, W. (2020). *Psych: Procedures for psychological, psychometric, and*

personality research. Evanston, Illinois: Northwestern University. Retrieved from

<https://CRAN.R-project.org/package=psych>

Stuermer, K., Seidel, T., Mueller, K., Häusler, J., & Cortina, K. S. (2017). What is

in the eye of preservice teachers while instructing? An eye-tracking study about

attention processes in different teaching situations. *Zeitschrift für*

Erziehungswissenschaft, 20(1), 75–92.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag

New York. Retrieved from <https://ggplot2.tidyverse.org>

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string*

operations. Retrieved from <https://CRAN.R-project.org/package=stringr>

Wickham, H. (2020a). *Forcats: Tools for working with categorical variables*

(factors). Retrieved from <https://CRAN.R-project.org/package=forcats>

- 1039 Wickham, H. (2020b). *Tidyr: Tidy messy data*. Retrieved from
1040 <https://CRAN.R-project.org/package=tidyr>
- 1041 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . .
1042 Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*,
1043 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- 1044 Wickham, H., Hester, J., & François, R. (2018). *Readr: Read rectangular text data*.
1045 Retrieved from <https://CRAN.R-project.org/package=readr>
- 1046 Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. (2017). See and tell: Differences
1047 between expert and novice teachers' interpretations of problematic classroom
1048 management events. *Teaching and Teacher Education*, 66, 295–308.