Electronic Supplementary Material 1

Figure S1

Mean Number of Teachers' Assessments and Experts' Judgment for 10 Texts Split by Scale

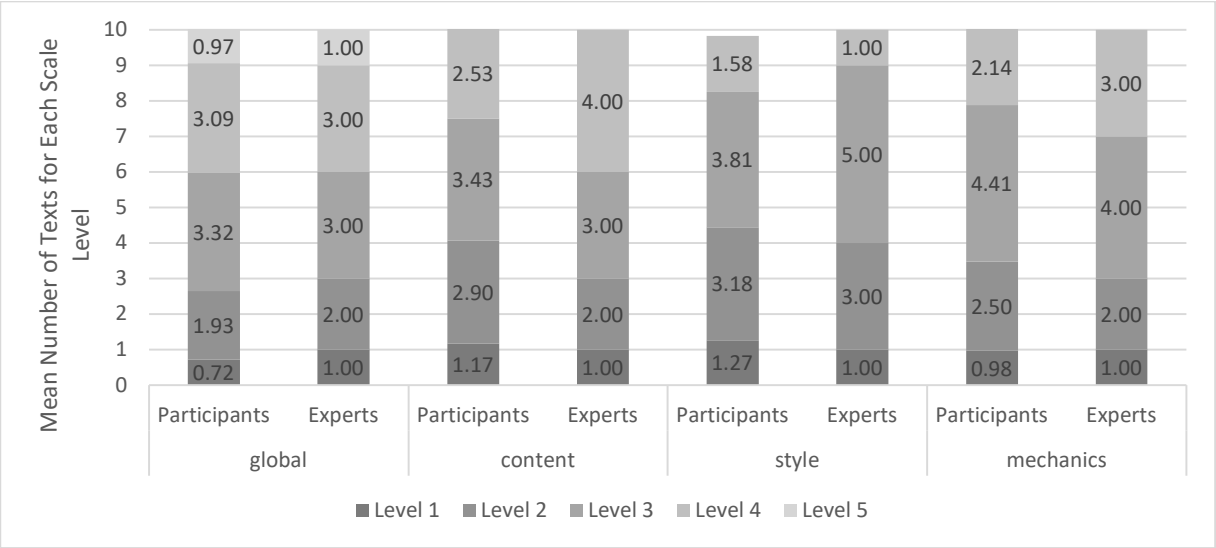
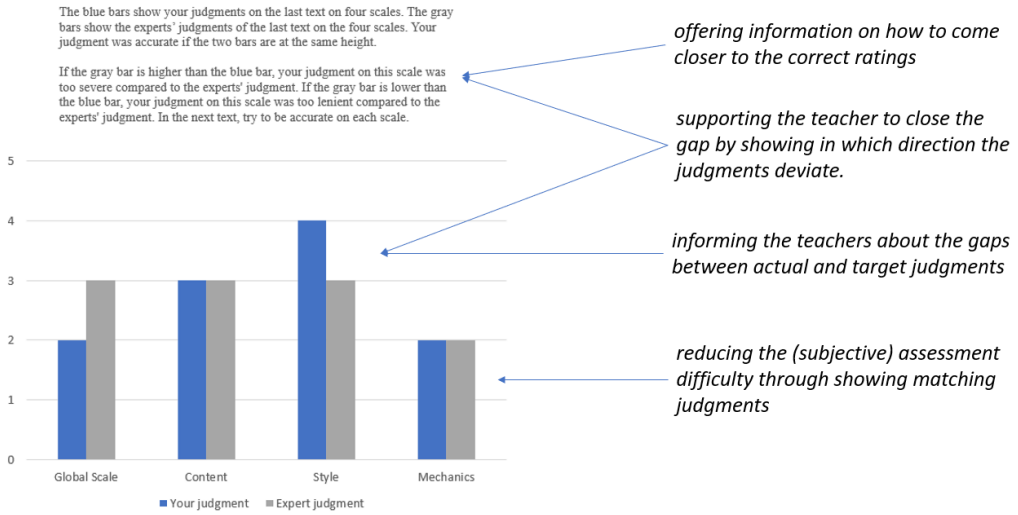


Figure S2

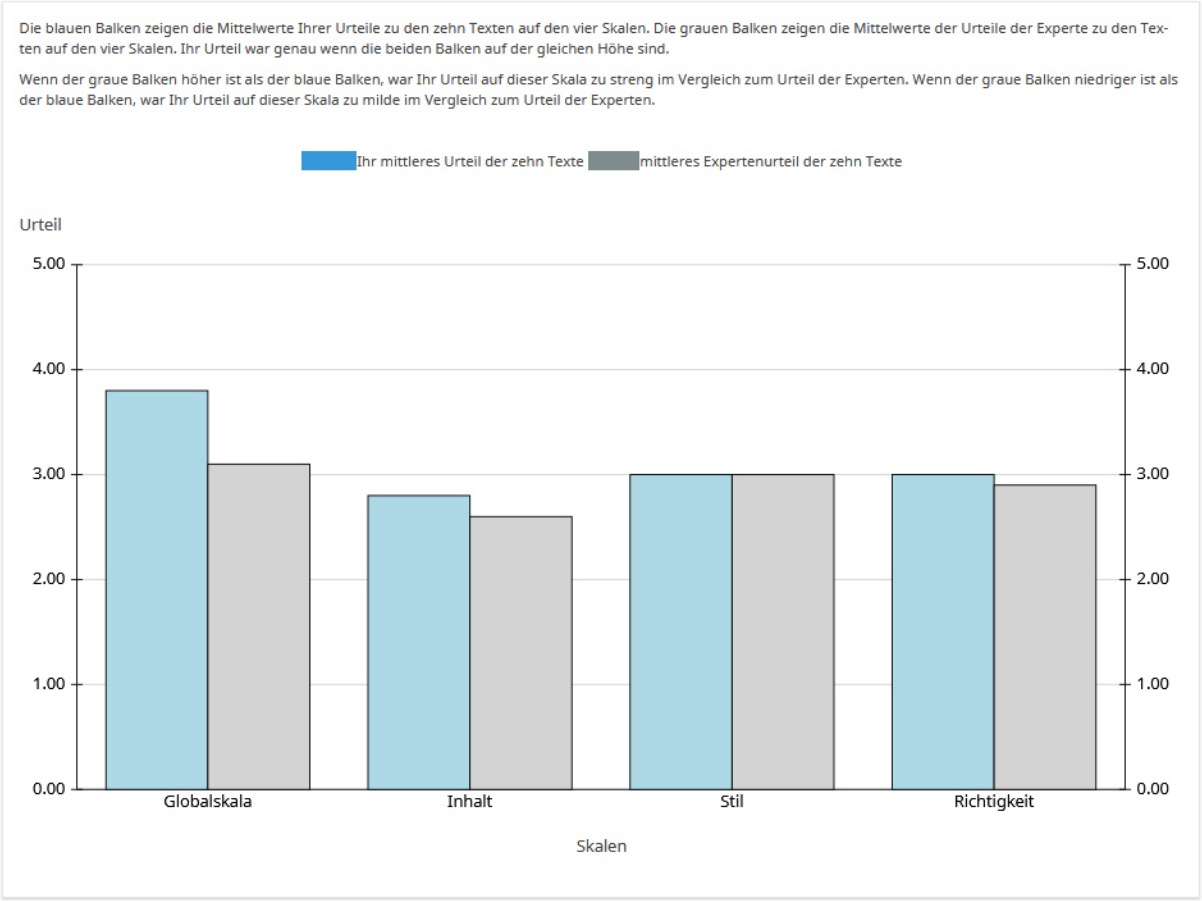
Linking Theoretical Assumptions To the Feedback



https://doi.org/10.1024/1010-0652/a000365

Figure S3

Feedback After the First Five Texts



The blue bars show the mean values of your judgments on the 10 texts on the four scales. The gray bars show the mean values of the experts' judgments on the texts on the four scales. Your judgment was accurate if the two bars have the same height.

If the gray bar is higher than the blue bar, your judgment on this scale was too strict compared to the experts' judgment. If the gray bar is lower than the blue bar, your judgment on this scale was too lenient compared to the experts' judgment.

https://doi.org/10.1024/1010-0652/a000365

Figure S4

Group Differences on Multiple Components of Assessment Accuracy for the First Five Texts

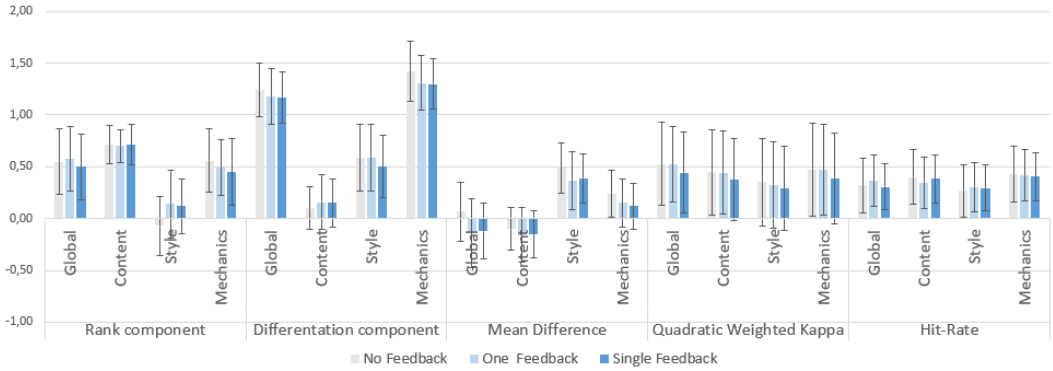
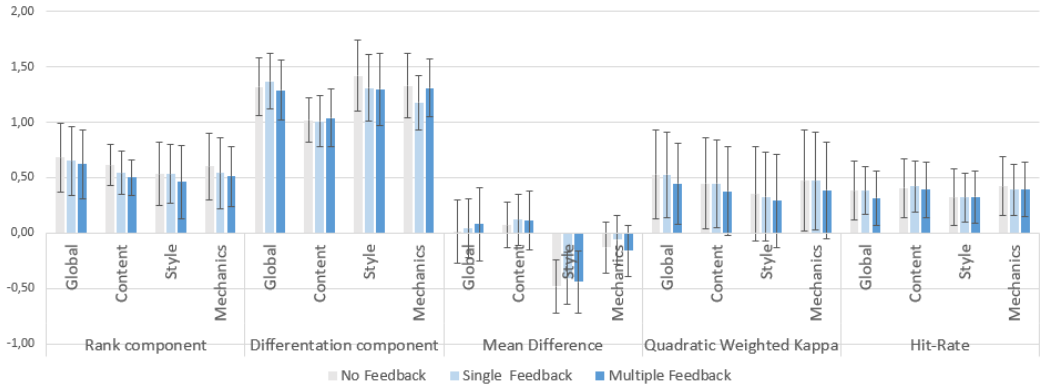


Figure S5

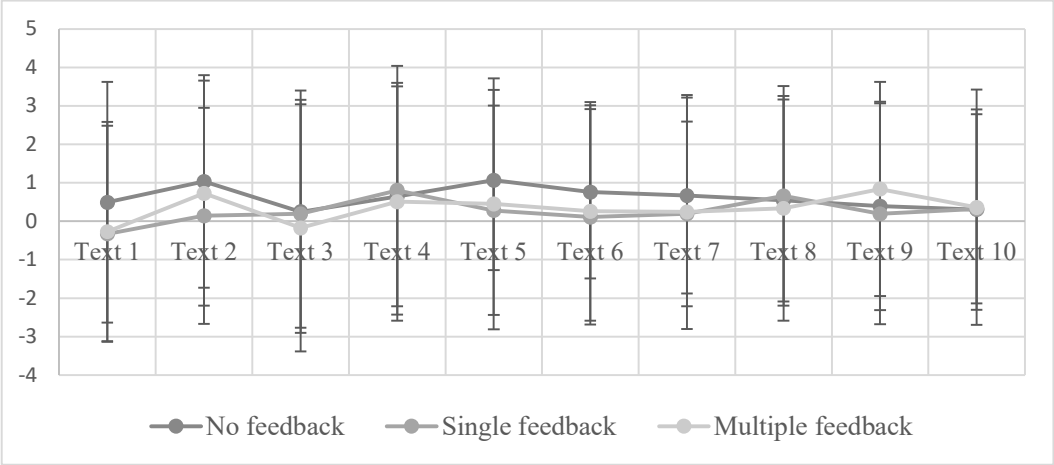
Group Differences on Multiple Components of Assessment Accuracy for the Last Five Texts



https://doi.org/10.1024/1010-0652/a000365

Figure S6

Assessment Accuracy, Split from Text 1 to Text 10



https://doi.org/10.1024/1010-0652/a000365

Table S1

Randomization Check For the Covariates

	N	*n* female	*n* experienced teachers	Mean (*SD*) years of experience	*n* student teachers	No. of semesters studied
No feedback	85	56	31	8.37 (10.69)	54	8.63 (9.57)
Single feedback	112	80	46	11.33 (12.03)	66	8.53 (9.87)
Multiple feedback	98	65	37	10.94 (11.55)	61	8.99 (9.96)

Table S2

Means (Standard Deviations) for Assessment Accuracy of the First Five Texts, Split by Group and Scale

	Mean Difference				Hit Rate			
	Global	Content	Style	Mechanics	Global	Content	Style	Mechanics
No Feedback	0.07 (0.57)	-0.10 (0.41)	0.49 (0.48)	0.24 (0.46)	0.32 (0.53)	0.40 (0.53)	0.27 (0.51)	0.42 (0.54)
Single Feedback	-0.14 (0.66)	-0.16 (0.53)	0.37 (0.56)	0.15 (0.46)	0.31 (0.49)	0.38 (0.50)	0.29 (0.47)	0.39 (0.49)
Multiple Feedback	-0.12 (0.53)	-0.15 (0.46)	0.39 (0.47)	0.12 (0.44)	0.36 (0.44)	0.34 (0.46)	0.31 (0.44)	0.39 (0.46)

Note: Correct judgments would result in a mean of zero for the mean difference and a value of one for the hit rate.

https://doi.org/10.1024/1010-0652/a000365

Table S3

Test Statistic for Assessment Accuracy

	Mean Difference				Mean Hit Rate			
	Global Scale	Content	Style	Mechanics	Global Scale	Content	Style	Mechanics
F (2, 288)	0.264	0.249	0.471	1.116	3.512	0.548	0.014	0.313
p	.768	.779	.625	.329	.031	.579	.986	.731

https://doi.org/10.1024/1010-0652/a000365