RESULTS REPORT

Duration: October 2022 to December 2024

Project ALFA2 (Adaptive Literacy and Language Skills through Feedback Acquisition) of the consortium partners

eKidz.eu GmbH (16INB1003B)

and

Chair for Didactics of German Language and Literature at the University of Regensburg (16INB1003A)

Status: 2025-04-20

The survey was approved by the Bavarian State Ministry for Education and Cultural Affairs (Reference: IV.7BO7106/190/22).

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Europäische Kommission

The project underlying this report was funded by the Federal Ministry of Education and Research under funding reference numbers 16INB1003A and B, respectively. The responsibility for the content of this publication lies with the author(s).

RESULTS REPORT

1. Project Description

1.1 Introduction and problem statement

Extensive research confirms the positive impact of feedback on language acquisition. However, providing differentiated and motivating feedback for reading—a critical skill for academic success and social participation—presents a significant challenge for teachers in daily classroom practice. Consequently, implementing individualized support is often difficult.

Building on the preceding ALFA project implemented in the partnership between the University of Regensburg and the eKidz.eu GmbH, the project's core innovation is the development and evaluation of an automated, adaptive feedback system within a digital reading program for elementary school children. This system features a diagnostic tool with an integrated feedback function, enabling personalized feedback on reading fluency and comprehension based on an automated diagnostic process, in line with the principles of effective feedback (Hattie & Timperley, 2007).

1.2 Scientific and technological foundations

This project addresses a largely unexplored field: the integration of effective feedback for reading fluency into digital applications, particularly the link between diagnostics and substantive feedback for reading aloud. The work is grounded in extensive reading promotion research from the University of Regensburg—evidenced by nationally and internationally recognized projects like ELiS, FiLBY, ADELE, and RATTE—and builds upon the eKidz.eu app, a core German literacy program implemented across four states and numerous school districts. By connecting to the BIRD National Education Platform, the project aims to make this innovative approach widely accessible to schools, teachers, and families. The project leverages the technologically advanced eKidz.eu platform, which organizes content into twelve distinct, didactically differentiated difficulty levels. The app provides teachers with detailed insights by tracking individual student activity and progress. A key feature—the ability for students to record themselves reading aloud for teacher evaluation—serves as the ideal foundation for implementing automated speech recognition. This technology enables the analysis of student performance and the delivery of powerful, adaptive feedback.

1.3 Project procedure and methodology

The project involved the development and evaluation of an automated feedback system to support teachers in promoting reading fluency. Students practiced reading aloud using the eKidz.eu app and received individualized, effective feedback based on a AI empowered implementation of traditional oral reading fluency assessment. This was enhanced with personalized tips based on each student's strengths and weaknesses. The system provided teachers with a clear overview of their class' performance, enabling more targeted support.

A quasi-experimental design was used to compare the effectiveness of the digital reading training with and without the new feedback system, and results were cross-validated with paper-based tests. Data collection was essential to calibrate the speech recognizer for classroom conditions (e.g., background noise, dialects, pauses) and to refine the feedback mechanism for different performance levels.

The data collection process, approved by the Bavarian State Ministry for Education and Cultural Affairs (Ref: IV.7-BO7106/190/22), was organized as follows:

Initial Setup: In collaboration with teachers, project staff distributed research instruments, obtained parental consent, and created pseudonymized student accounts for the eKidz.eu app on school tablets.

Pre-Test (TPE1): In September/October 2023, baseline data was collected using a student questionnaire for demographic information, a reading motivation questionnaire, and the standardized «Salzburg Reading Screening» (SLS) for reading fluency. A teacher questionnaire documented the existing reading curriculum.

Intervention Phase 1: From October to November 2023, the first cohort of students participated in the reading training as part of their German lessons.

Post-Test (TPE2 for Cohort 1): In December 2023, the SLS and reading motivation questionnaire were administered again.

Staggered Intervention: To assess the newly developed feedback functions, a second cohort began the trial in January 2024, following the same pre-test (TPE1) and intervention protocol. This staggered approach allowed for the capture of learning progress across two distinct cohorts under different conditions (with and without feedback).

2. Research Question

F1: How did reading fluency develop in regular reading instruction vs. during training with the eKidz.eu app (with and without implemented feedback) in second grade?

Hypothesis: Training with the eKidz.eu app shows significant advantages over regular reading instruction.

F2: How do different demographic groups benefit from the treatments regarding their reading fluency?

Hypothesis: Weak readers benefit more than strong readers, as do children from educationally disadvantaged families and those with migration backgrounds.

F3: How do reading-related self-concept and reading motivation develop in the groups?

Hypothesis: Self-concept remains stable, while reading motivation increases slightly, especially in the eKidz groups.

3. Design

Glossary:

CG = KG (Control Group)

F = MF (Feedback)

W/O F = OF (Without Feedback)

FiLBY = Fachintegrierte Leseförderung Bayern (Subject-integrated Reading Support Bavaria — a systematic and long-term reading program that sustainably promotes reading competency for students from grades 2-4 with comprehensive materials)

SRS = SLS (Salzburg Reading Screening)
TPE = TPE (Time point of evaluation)

SD = Standardabweichung (Standard Deviation)

LQ = Lesequotient (Reading Quotient)

UREMS = Unbedingte Random-Effects-Modelle (Unconditional Random Effects Model)

Tab. 1: Survey design of the evaluation

	TPE1	Treatment	TPE2[1]	Description
W/OF	2023-10	Training with the eKidz.eu app (without implemented feedback)	2023-12	Second-grade students at Bavarian elementary schools; classes with increased German/reading support needs; training with app in German classes and free learning time and at home; training in the form of reading texts as complete books on 12 levels
F	2024-01	Training with the eKidz.eu app (with newly implemented feedback)	2024-04	Second-grade students at Bavarian elementary school; classes with increased German reading support needs; training with the app in German lessons and in free learning time, and at home; books at 12 levels; AI analysis of decoding accuracy, speed and text comprehension; three-stage feedback on the level of competence achieved, the next competence target and strategy for achieving this target
CG	2024-01	FiLBY-2 / Promoting reading in German lessons according to the curriculum	2024-04	Second-grade students at Bavarian elementary schools; classes predominantly using the evaluated FiLBY-2 training (analog teacher-guided reading program focused on improving reading fluency) during the survey period

Note: 1 The resulting difference in treatment duration, especially in the W/OF group, was taken into account in the evaluations.

2 With the exception of one class, the CG classes carried out the already evaluated FiLBY-2 reading program and therefore formed a very strict comparison group.

4. Sample

Tab. 2: Grade (Jahrgangsstufe 2)

	Classes	Children
CG	19	294
F	39	618
W/OF	10	170
Total	**68**	**1,082**

Tab. 3: Age in years (s_alter)

	M (SD)
CG	8.60 (0.61)
F	8.64 (0.58)
W/OF	7.81 (0.50)
Total	**8.40 (0.67)**

Tab. 4: Gender (s_geschl)

	Girls	Boys
CG	48.54 %	51.46 %
F	41.39 %	58.61 %
W/OF	47.71 %	52.29 %
Total	**45.25 %**	**54.75 %**

Tab. 5: Family language (s_famspr)

	No German	German
CG	17.54 %	82.46 %
F	20.49 %	79.51 %
W/OF	16.99 %	83.01 %
Total	**18.66 %**	**81.34 %**

Tab. 6: Multilingual (s_mehrsprachig)

	No	Yes
CG	57.31 %	42.69 %
F	46.31 %	53.69 %
W/OF	43.79 %	56.21 %
Total	**48.94 %**	**51.06 %**

Tab. 7: Country of origin (s_mgeneration)

	All born in Germany	Child not born in Germany	Parents not born in Germany
CG	47.37 %	01.75 %	50.88 %
F	34.84 %	00.00 %	65.16 %
W/OF	32.03 %	01.31 %	66.67 %
Total	**37.85 %**	**00.88 %**	**61.27 %**

Tab. 8: Books at home (s_medien_niveau)

	Low	High
CG	47.22 %	52.78 %
F	50.96 %	49.04 %
W/OF	50.68 %	49.32 %
Total	**50.00 %**	**50.00 %**

Tab. 9: Socioeconomic status (s_ses)

	Low	High
CG	75.74 %	24.26 %
F	67.77 %	32.23 %
W/OF	68.03 %	31.97 %
Total	**70.25 %**	**29.75 %**

Tab. 10: Note (mzp1_note_d)

	M (SD)
CG	2.07 (0.98)
F	2.34 (1.09)
W/OF	2.47 (1.43)
Total	**2.30 (1.13)**

Tab. 12: Pages read in app (PagesRead)

	M (SD)
F	180.81 (159.77)
W/OF	147.43 (182.40)
Total	**164.70 (160.36)**

Tab. 13: Pages listened to in app (PagesListened)

	M (SD)
F	369.00 (198.71)
W/OF	293.23 (213.71)
Total	**329.10 (205.11)**

Tab. 14: Recorded pages in app (PagesRecorded)

	M (SD)
F	352.80 (193.45)
W/OF	177.38 (171.31)
Total	**293.10 (200.50)**

5. Overview of Key Results of the Project

Key Findings:

1. **Training with the eKidz app is fundamentally effective (d = .30)** and achieves comparable effects to an already positively evaluated reading training program. Training with eKidz can therefore be classified as evidence-based.

2. The further development of the ekidz app during the project period **proved to increase effectiveness**, regardless of the initial motivation of the child or the starting level of reading performance.

3. There were **no differences in terms of use during lessons or at home**. Both forms were equally effective.

4. When controlling for the influence of demographic variables (such as migration background, gender, age) as well as training duration and initial reading fluency level, **the eKidz group with the enhanced app benefited significantly more** than the group that predominantly engaged in paper-based evidence-based reading training.

The project shows that digital learning applications offer a flexible way to specifically promote reading skills. The high proportion of children with migration backgrounds and predominantly low socioeconomic backgrounds of the sample also underscores the importance of such evidence-based support programs to reduce educational inequalities. The study contributes to the growing evidence that personalized, technology-supported learning offerings as the eKidz app have positive effects on reading promotion, especially when they are continuously developed, e.g., using AI and integrating adaptive feedback.

6. Results in Detail

6.1 Descriptive Results and Non-nested Procedures

6.1.1 Reading Fluency (SRS)

Shapiro-Wilk Test on Normal Distribution: *W* = 0.98253, *p* <.0001 (TPE1) bzw. *W* = 0.98652, *p* <.0001 (TPE2).

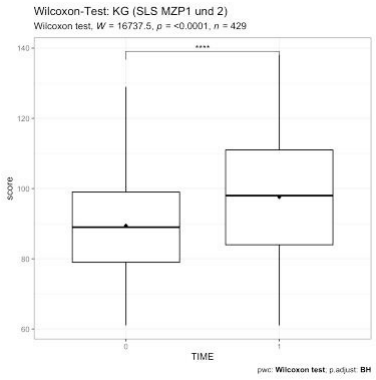
Tab. 15: Mean values of the Salzburg Reading Screening reading quotient [range] (standard deviation)

	TPE1	**TPE2**	***d***	**Wilcoxon**	***p*.adj**
CG	89.47 (14.26)	97.55 (18.39)	*d* = .49	*W* = 16737, *z* = -4.85, *p* < .0001, D*Mdn*= 9	****
F	89.18 (16.65)	94.20 (17.52)	*d* = .29	*W* = 76753, *z* = -4.11, *p* < .0001, D*Mdn* = 7	****
W/OF	84.48 (14.71)	89.40 (14.35)	*d* = .34	*W* = 10252.5, *z* = - 3.24, *p* < .001, D*Mdn* = 7	**
Total	88.18 [61.00; 137.00] (15.68)	94.20 [61.00; 138.00] (17.42)	*d* = .36		

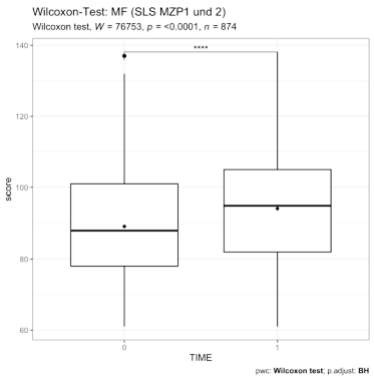
Note.: *d*: Cohens d; *p*.adj: Wilcoxons adj. *p*

Fig. 01: Wilcoxon-Test of the three groups (Reading Fluency)

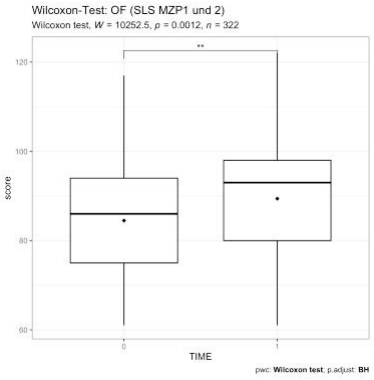
CG (TPE1 and TPE2)



F (TPE1 and TPE2)



W/O F (TPE1 and TPE2)



6.1.2 Self-concept (*mzp1_selbstkonzept*)

Tab. 16: Scale Parameters

Origin	Schaufelberger, R., Kleinkorres, R., Becher, L., Ludewig, U., Lorenz, R. & McElvany, N. (2024). IGLU 2021. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen. Münster; New York: Waxmann. http://doi.org/10.25656/01:30762. [Itemnummer]
Items (4 Items)	*MZP_lm_01_a_lesen_muss* (R) *MZP_lm_01_b_freuen_buch* [ASBR07B] *MZP_lm_01_c_lesen_langweilig* (R) [ASBR07C] *MZP_lm_01_d_lesen_gern* [ASBR07E]
Scale	0: little agreement ... 3: high agreement Value range: 0-12 (higher values mean stronger reading-related self-concept)
Scale Reliability	Cronbach Alpha a = .66 (TPE1), a = .67 (TPE2) McDonalds Omega w = .67 (TPE1), w = .68 (TPE2)

Shapiro-Wilk Test on Normal Distribution: W = 0.94, p <.0001 (TPE1) bzw. W = 0.94, p <.0001 (TPE2).

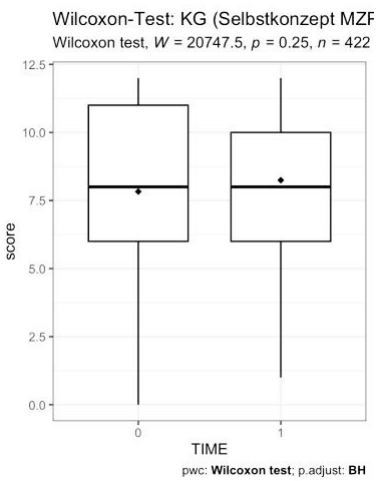
Tab. 17: Mean values self-concept (standard deviation)

	TPE1	**TPE2**	***d***	**Wilcoxon**	***p*.adj**
CG	7.83 (3.13)	8.25 (2.70)	*d* = .14	*W* = 20747.5, *z* = -1.15, *p* = 0.25, D*Mdn* = 0	n.s.
F	7.85 (3.29)	8.01 (2.92)	*d* = .05	*W* = 87234.5, *z* = -0.27, *p* = 0.784, D*Mdn* = 0	n.s.
W/OF	8.43 (2.40)	8.90 (2.80)	*d* = .18	*W* = 10178, *z* = -2.24, *p* = 0.0249, D*Mdn* = 0	*
Total	7.98 (3.07)	8.22 (2.86)	*d* = .08		

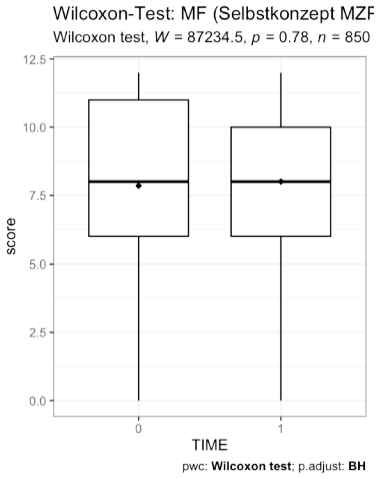
Note: *d*: Cohens d; *p*.adj: Wilcoxons adj. *p*

Fig. 02: Wilcoxon-Test of the three groups (self-concept)

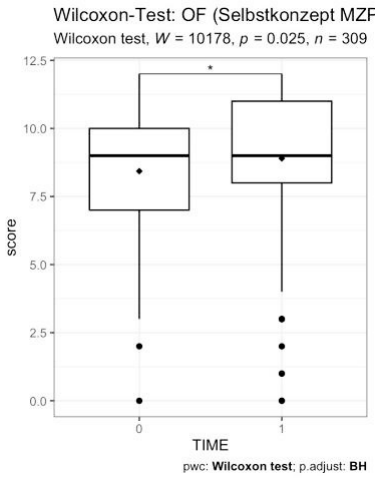
CG (self-concept TPE1 and TPE2)



F (self-concept TPE1 and TPE2)



W/O F (self-concept TPE1 and TPE2)



6.1.3 Reading Motivation 1 (*mzp1_lesemotivation_1*)

Tab. 18: Scale Parameters

Origin	Schaufelberger, R., Kleinkorres, R., Becher, L., Ludewig, U., Lorenz, R. & McElvany, N. (2024). IGLU 2021. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen. Münster; New York: Waxmann. http://doi.org/10.25656/01:30762. [Itemnummer]
Items (3 Items)	*MZP_lm_02_a_lesen_leicht* [ASBR08B] *MZP_lm_02_b_gut_lesen_andere (R)* [ASBR08D] *MZP_lm_02_c_schnell_lesen_andere* (R)
Scale	0: low agreement ... 3: high agreement Value range: 0-9 (higher values mean stronger reading-related self-concept)
Scale Reliability	Cronbach Alpha a = .65 (TPE1), a = .69 (TPE2) McDonalds Omega w = .68 (TPE1), w = .71 (TPE2)

Shapiro-Wilk Test on Normal Distribution: W = 0.94, p <.0001 (TPE1) bzw. W = 0.94, p <.0001 (TPE2).

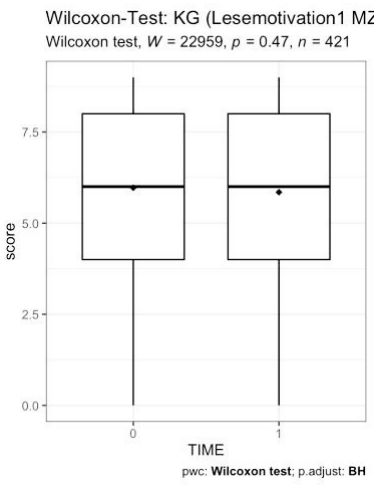
Tab. 19: Mean values for reading motivation 1 (standard deviation)

	TPE1	TPE2	d	Wilcoxon	p.adj
CG	5.97 (2.25)	5.85 (2.21)	d = -.05	W = 22959, z = -0.71, p = 0.475, D*Mdn* = 0	n.s.
F	6.17 (2.22)	6.14 (2.20)	d = -.01	W = 89874, z = -0.37, p = 0.708, D*Mdn* = 0	n.s.
W/OF	6.02 (2.15)	6.10 (2.17)	d = -.04	W = 11654, z = -0.33, p = 0.738, D*Mdn* = 0	n.s.
Total	6.08 (2.21)	6.06 (2.20)	d = -.01		

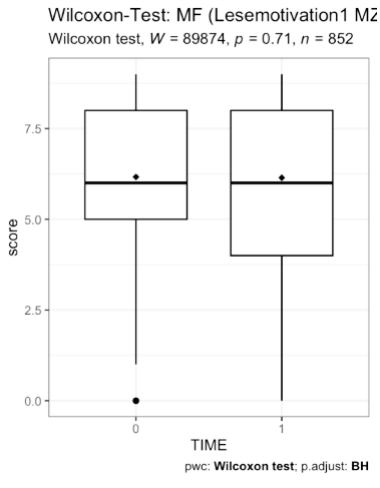
Note: d: Cohens d; p.adj: Wilcoxons adj. p

Fig. 03: Wilcoxon-Test on the three groups (reading motivation 1)

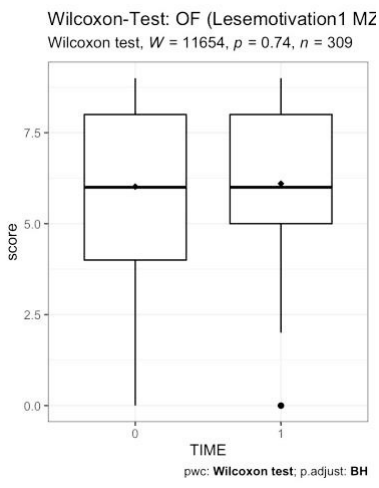
CG (reading motivation 1 TPE1 & TPE2)



F (reading motivation 1 TPE1 & TPE2)



W/O F (reading motivation 1 TPE1 & TPE2)



6.1.4 Reading Motivation 2 (mzp1_lesemotivation_2)

Tab. 20: Scale parameters

Origin	Bos, W., Strietholt, R., Goy, M., Stubbe, T. C., Tarelli, I. & Hornberg, S. (2010). *IGLU 2006. Dokumentation der Erhebungsinstrumente*. Münster, New York, München, Berlin : Waxmann. https://doi.org/10.25656/01:15670 [Itemnummer]
Items (7 Items)	*MZP_lm_03_a_text_verstehen (R)* [SFP021] *MZP_lm_03_b_woerter_schwierig (R)* [SFP022] *MZP_lm_03_c_woerter_neu (R)* [SFP023] *MZP_lm_03_d_texte_schwer (R)* [SFP024] *MZP_lm_03_e_lesen_schnell_gut* [SFP025] *MZP_lm_03_f_geschichten_lesen* [SFP026] *MZP_lm_03_g_lesen_oefter (R)* [SFP027]
Scale	0: low agreement ... 3: high agreement Value range: 0-21 (higher values mean stronger reading-related self-concept)
Scale Reliability	Cronbach Alpha a = .78 (TPE1), a = .79 (TPE2) McDonalds Omega w = .79 (TPE1), w = .79 (TPE2)

Note: (R): right

Shapiro-Wilk Test on Normal Distribution: $W = 0.98$, $p < .0001$ (TPE1) bzw. $W = 0.98$, $p < .0001$ (TPE2).

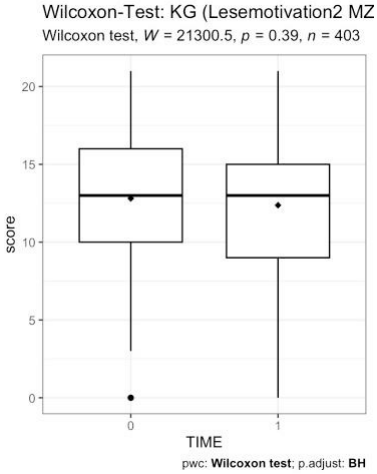
Tab. 21: Mean values for reading motivation 2 (standard deviation)

	TPE1	TPE2	*d*	Wilcoxon	*p*.adj
CG	12.81 (4.64)	12.37 (4.70)	*d* = -.09	*W* = 21300.5, *z* = -0.87, *p* = 0.387, D*Mdn* = 0	n.s.
F	13.00 (4.86)	12.02 (4.82)	*d* = -.20	*W* = 89961, *z* = -3.06, *p* = 0.0022, D*Mdn* = 1	**
W/OF	12.58 (4.03)	13.36 (4.21)	*d* = .19	*W* = 10310.5, *z* = -1.66, *p* = 0.0969, D*Mdn* = 0	n.s.
Total	12.85 (4.61)	12.35 (4.71)	*d* = -.11		

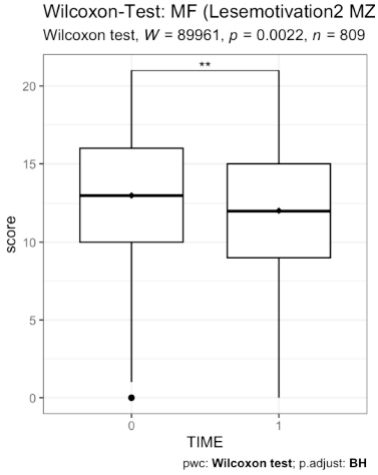
Note: *d*: Cohens d; *p*.adj: Wilcoxons adj. *p*

Fig. 04: Wilcoxon-Test on the three groups (reading motivation 2)

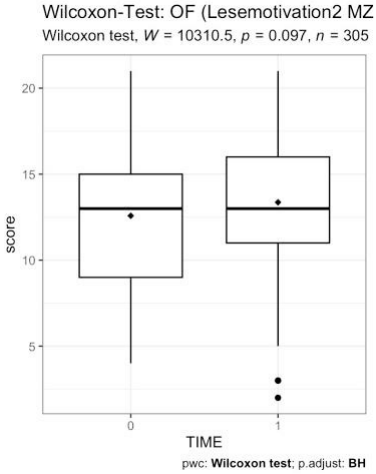
CG (reading motivation 2 TPE1 & TPE2)



F (reading motivation 2 TPE1 & TPE2)



W/O F (reading motivation 2 TPE1 & TPE2)



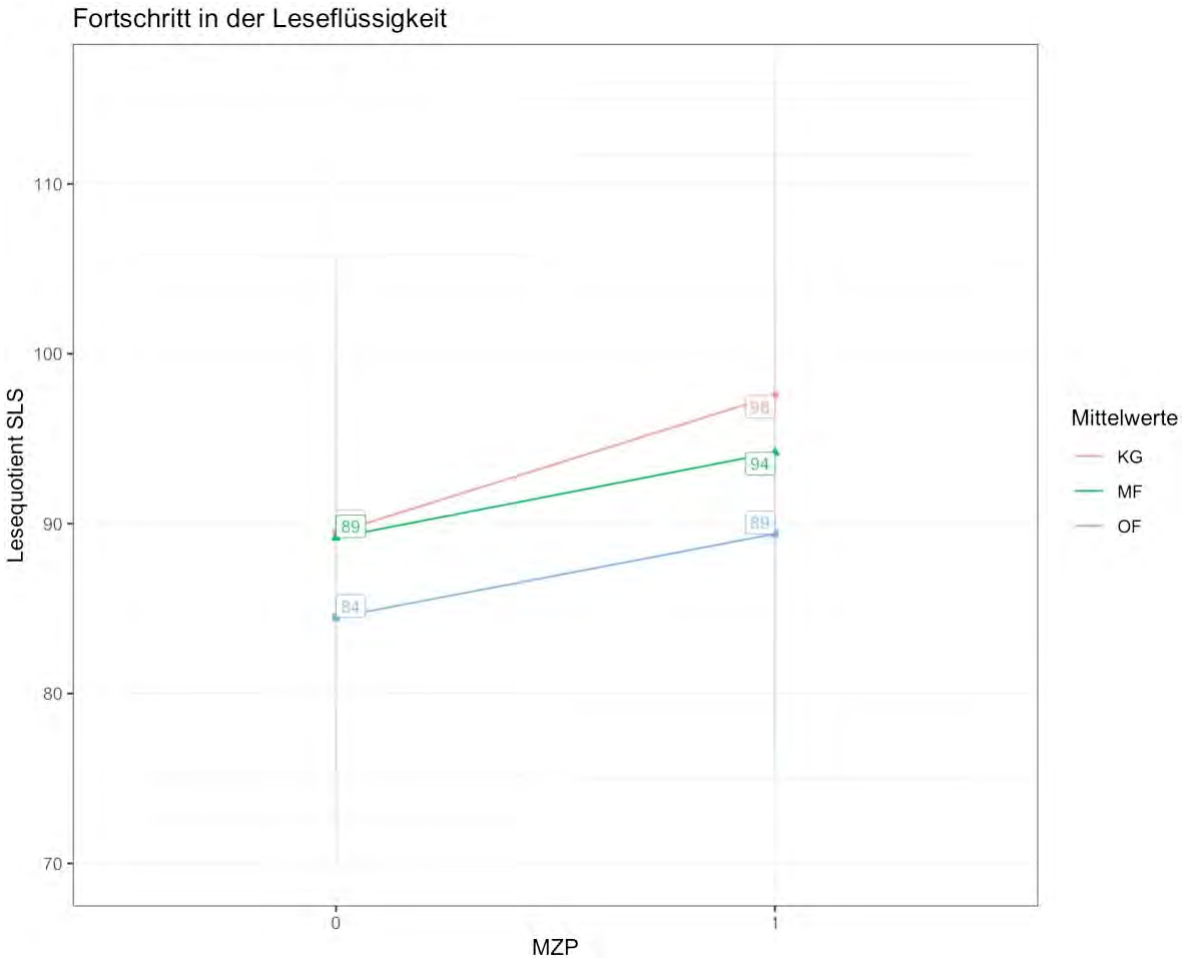
6.2 Reading Fluency

6.2.1 Development of reading fluency by group membership controlling for treatment duration

The model with groups and TPE variables and their interaction fit significantly better than the UREMs (null model) and was therefore selected for analysis (c2 = 271.89, df = 6, p < .001). The ICC was greater than 0.1 (icc = .86).

Fig. 05: Progress in reading fluency according to groups

Reading quotient (SRS) vs TPE



Tab. 22: Regression coefficients of the analysis of progress in reading fluency as a function of group membership, controlling for treatment duration

	Salzburger Reading Screening LQ			
Fixed Effects	Regression-coefficient	CI	p	
Intercept (CG)	88.51	76.23 – 100.79	<0.001	***
Treatment Duration	-0.01	-0.21 – 0.18	0.903	
TPE	9.04	7.57 – 10.51	<0.001	***
Group F	-1.12	-6.56 – 4.33	0.687	
Group W/OF	-6.81	-16.54 – 2.92	0.170	
TPE × Group F	-2.66	-4.46 – -0.86	0.004	**
TPE × Group W/OF	-4.49	-6.54 – -2.45	<0.001	***
Random Effect				
Residual	40.16			
Student:Class	162.64			
Class	77.85			
Model Fit	R^2_{marg}= 0.07; R^2_{cond}= 0.87			

Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.2 Development of reading fluency by group membership and baseline performance controlling for treatment duration

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 1358.9, df = 12, p < .001). The ICC was greater than 0.1 (icc = .43).

Fig. 06: Progress in reading fluency as a function of group membership and baseline performance controlling for treatment duration

Reading quotient (SRS) vs base line performance of the sub-groups (low/high: 1,2,3,4) of each group: CG, F, W/O F

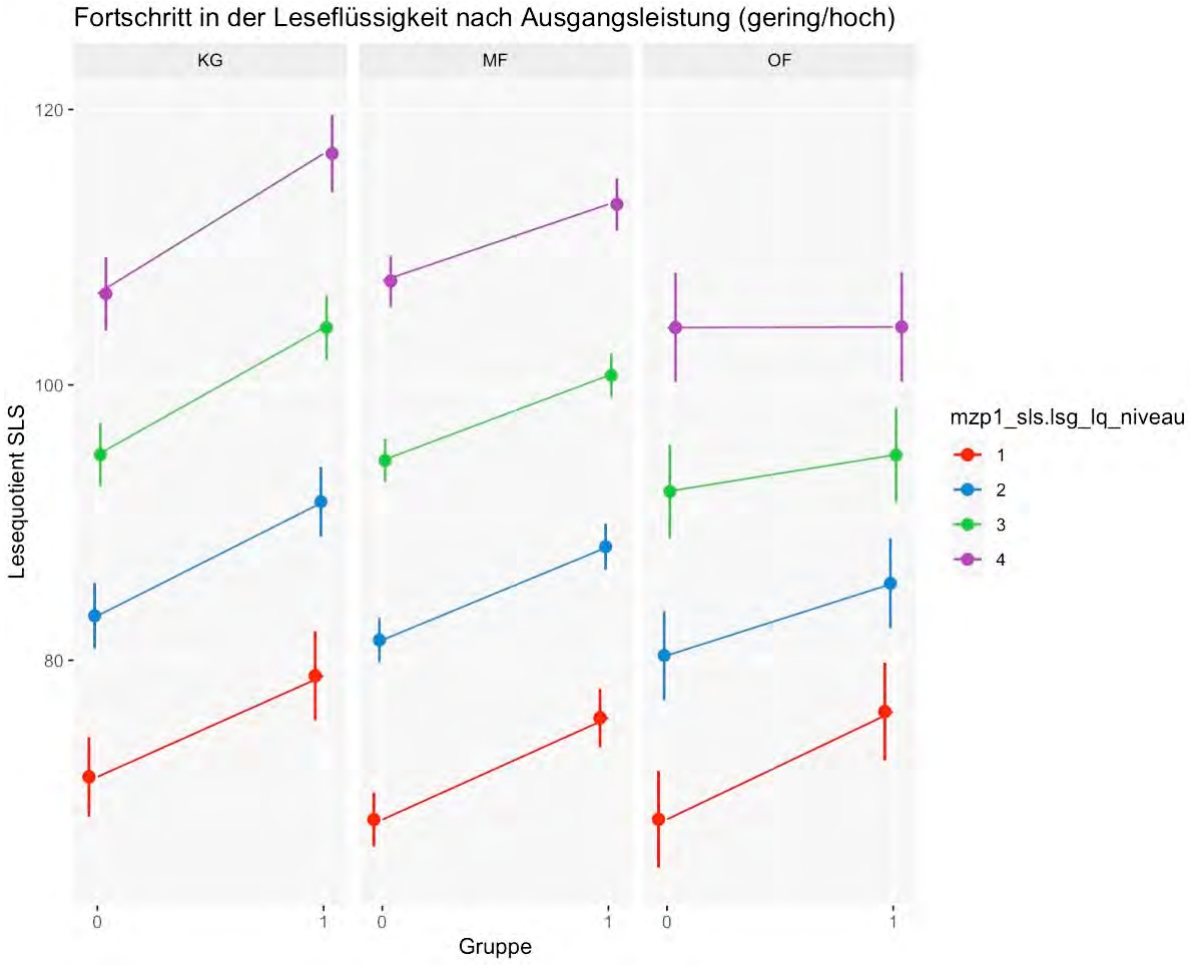
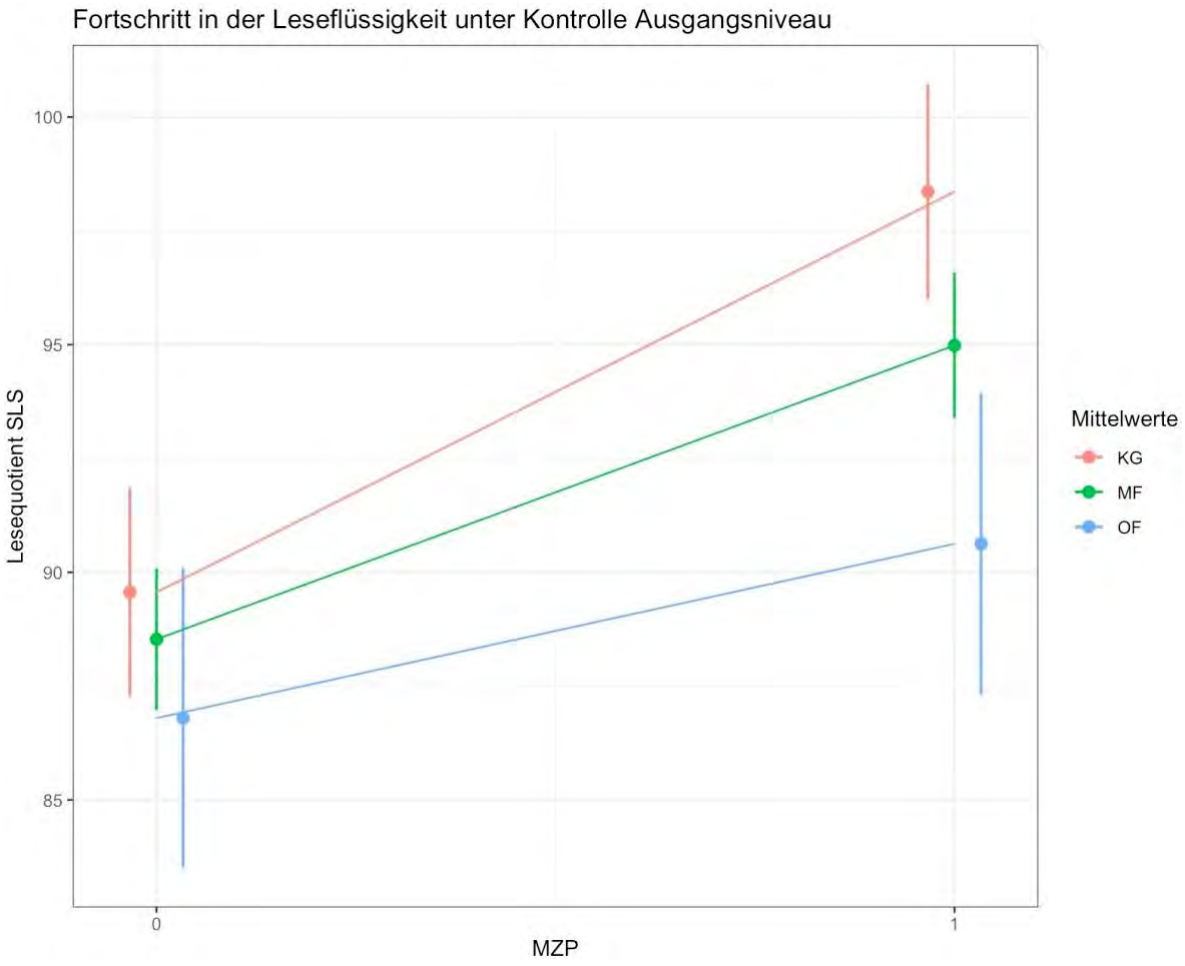


Fig. 07: Progress in reading fluency as a function of group membership and baseline performance controlling for treatment duration

Reading quotient (SRS) vs TPE, controlling for base line performance



Tab. 23: Regression coefficients of the analysis of progress in reading fluency as a function of group membership and baseline performance, controlling for treatment duration

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	61.79	54.54 – 69.04	<0.001	***
Treatment duration	-0.04	-0.14 – 0.06	0.444	
Baseline performance	11.70	10.64 – 12.75	<0.001	***
TPE	6.39	2.54 – 10.24	0.001	**
Group F	-4.45	-8.70 – -0.21	0.040	*
Group W/OF	-3.30	-9.35 – 2.76	0.286	
Baseline performance ×TPE	0.95	-0.36 – 2.26	0.156	
Baseline performance × Group F	1.35	0.05 – 2.64	0.042	*
Baseline performance × Group W/OF	0.21	-1.39 – 1.81	0.797	
TPE × Group F	1.61	-3.04 – 6.25	0.497	
TPE × Group W/OF	4.03	-0.97 – 9.02	0.114	
(Baseline performance × TPE) × Group F	-1.55	-3.14 – 0.03	0.054	
(Baseline performance × TPE) × Group W/OF	-3.54	-5.37 – -1.70	<0.001	***
Random effects				
Residual	38.06			
Child:Class	19.60			
Class	9.16			
Model Fit	R^2_{marg}= 0.76; R^2_{cond}= 0.86			

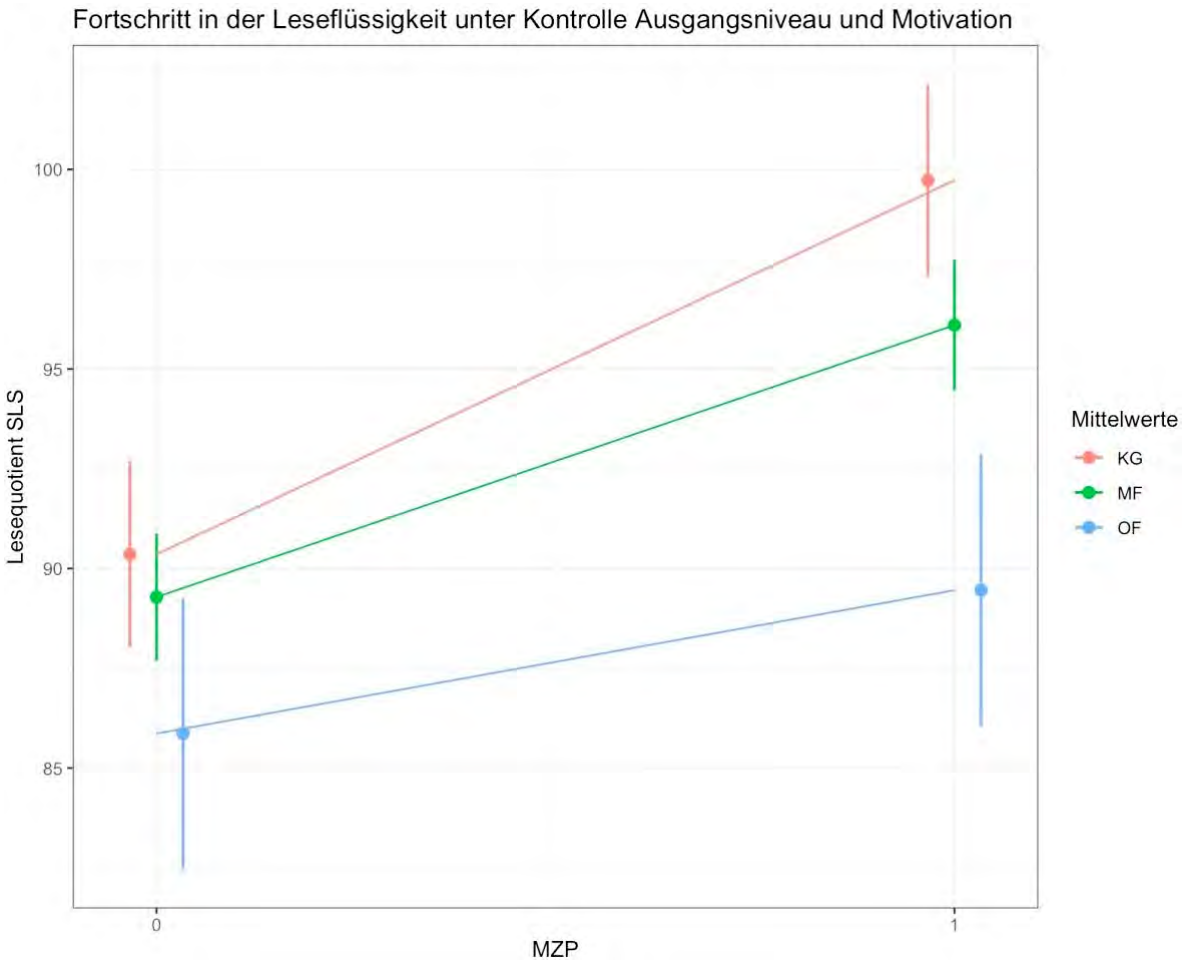
Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.3 Development of reading fluency according to group membership and baseline performance controlling for treatment duration, reading-related self-concept and motivation (1 and 2)

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 1270.3, df = 15, p < .001). The ICC was greater than 0.1 (icc = .42).

Fig. 08: Progress in reading fluency as a function of group membership and baseline performance, controlling for treatment duration, initial self-concept and motivation (1 and 2)

Reading quotient (SRS) vs TPE, controlling for base line performance and motivation



Tab. 24: Regression coefficients of the analysis of progress in reading fluency as a function of group membership and baseline performance, controlling for treatment duration, initial self-concept and motivation (1 and 2)

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	7.98	50.52 – 65.44	<0.001	***
Treatment duration	-0.07	-0.17 – 0.04	0.203	
Self-concept	0.36	0.17 – 0.55	<0.001	***
Reading motivation 1	0.74	0.44 – 1.04	<0.001	***
Reading motivation 2	0.01	-0.14 – 0.16	0.895	
Baseline performance	10.95	9.84 – 12.06	<0.001	***
TPE	7.50	3.37 – 11.64	<0.001	***
Group F	-3.90	-8.33 – 0.54	0.085	
Group W/OF	-4.08	-10.27 – 2.12	0.197	
Baseline performance × TPE	0.72	-0.66 – 2.11	0.306	
Baseline performance × Group F	1.09	-0.25 – 2.44	0.112	
Baseline performance × Group W/OF	-0.16	-1.81 – 1.49	0.849	
TPE × Group F	1.32	-3.70 – 6.34	0.606	
TPE × Group W/OF	1.57	-3.74 – 6.88	0.563	
(Baseline performance × TPE) × Group F	-1.50	-3.19 – 0.18	0.080	
(Baseline performance × TPE) × Group W/OF	-2.85	-4.79 – -0.90	0.004	**
Random effects				
Residual	37.28			
Child:Class	17.04			
Class	9.56			
Model Fit	R^2_{marg}= 0.77; R^2_{cond}= 0.87			

Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.4 Development of reading fluency according to group membership and baseline performance controlling for treatment duration, reading-related self-concept and motivation (1 and 2) as well as demographic characteristics

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 187.6, df = 18, p < .001). The ICC was greater than 0.1 (icc = .42).

Fig. 09: Progress in reading fluency by group membership and baseline performance, controlling for treatment duration, reading-related self-concept and motivation (1 and 2), and demographic characteristics

Reading quotient (SRS) vs TPE, controlling for demographic characteristic,, baseline performance and motivation

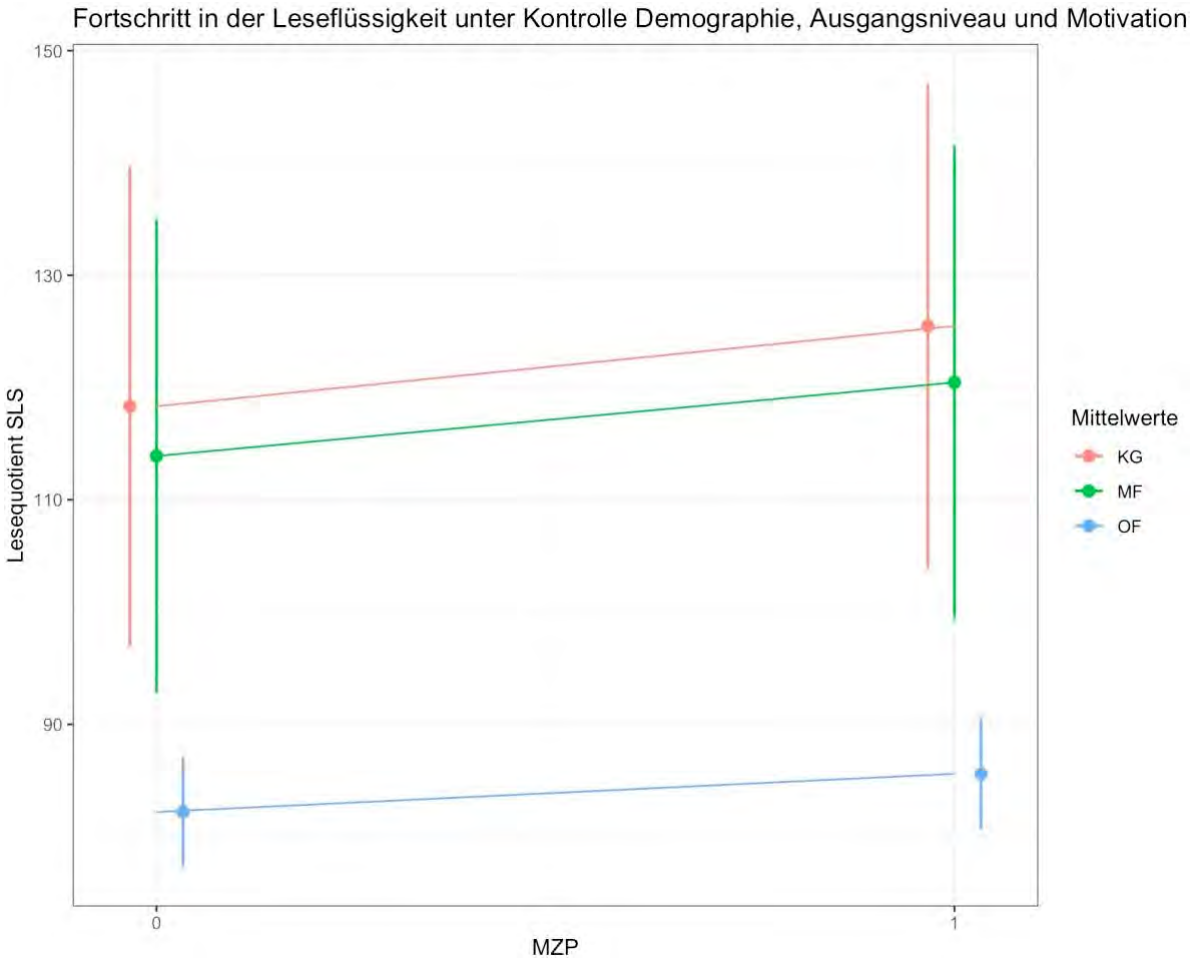
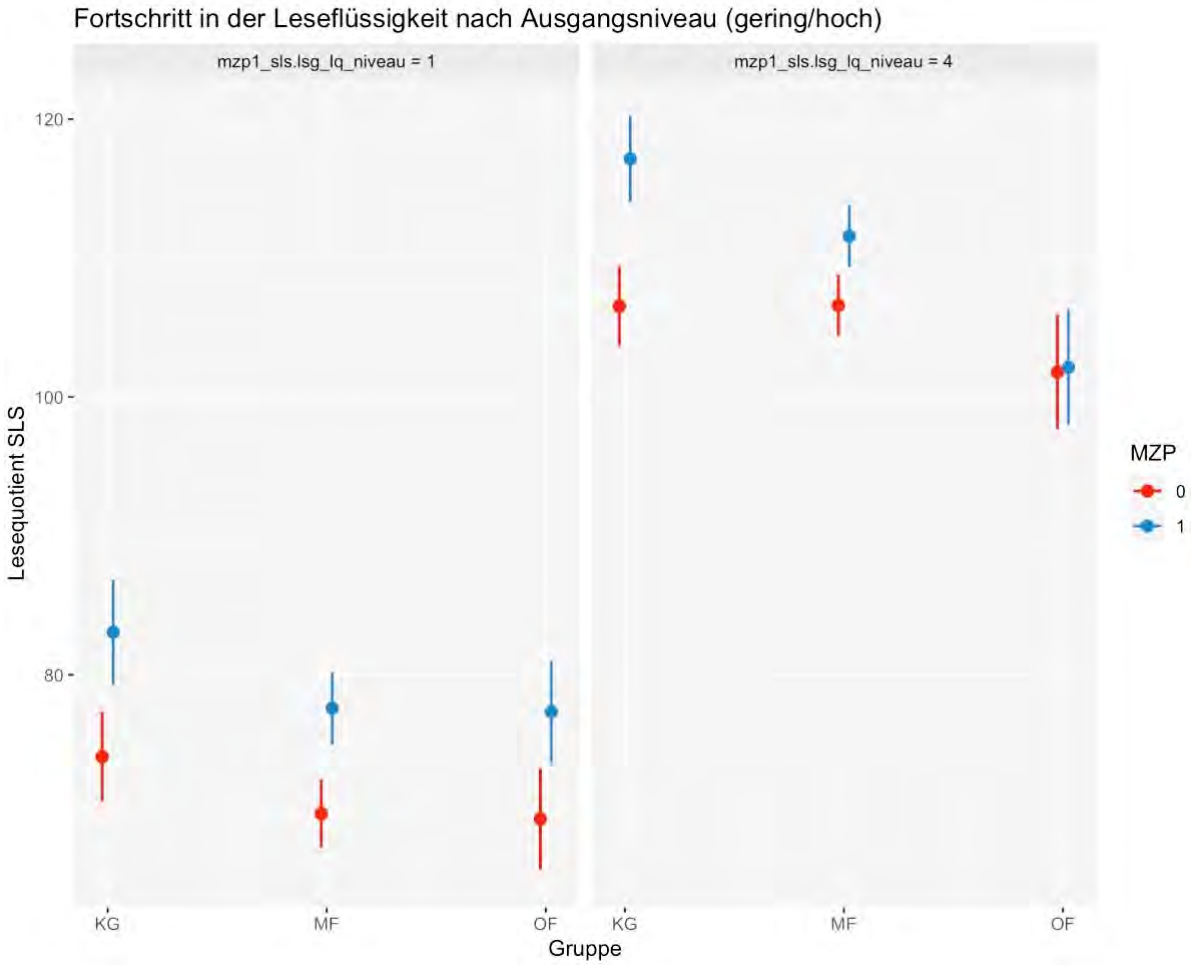


Fig. 10: Progress in reading fluency by group membership and baseline performance, controlling for treatment duration, reading-related self-concept and motivation (1 and 2), and demographic characteristics

Reading quotient (SRS) vs TPE according to base line performance of the sub-groups (low/high: 1,4) of each group: CG, F, W/O F



Tab. 25: Regression coefficients of the analysis of progress in reading fluency as a function of group membership and baseline performance, controlling for treatment duration, initial self-concept and motivation as well as gender, socio-economic status and family language

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	58.87	50.87 – 66.88	<0.001	***
Treatment duration	-0.08	-0.19 – 0.02	0.124	
Socioeconomic status	0.24	-0.08 – 0.56	0.138	
Gender	0.15	-1.02 – 1.32	0.805	
Family Language (G)	1.01	-0.53 – 2.55	0.198	
Self-concept	0.26	0.04 – 0.49	0.023	*
Reading motivation 1	0.80	0.45 – 1.14	<0.001	***
Reading motivation 2	0.03	-0.14 – 0.20	0.728	
Baseline performance	10.81	9.58 – 12.04	<0.001	***
TPE	8.40	3.67 – 13.13	0.001	**
Group F	-5.48	-10.48 – -0.48	0.032	*
Group W/OF	-4.39	-11.01 – 2.22	0.193	
Baseline performance × TPE	0.55	-1.04 – 2.14	0.498	
Baseline performance × Group F	1.39	-0.19 – 2.96	0.084	
Baseline performance × Group W/OF	-0.09	-1.85 – 1.68	0.923	
TPE × Group F	0.05	-5.83 – 5.94	0.986	
TPE × Group W/OF	1.76	-4.23 – 7.74	0.565	
(Baseline performance × TPE) × Group F	-1.42	-3.41 – 0.57	0.161	
(Baseline performance × TPE) × Group W/OF	-3.01	-5.18 – -0.83	0.007	**
Random effects				
Residual	38.69			
Child:Class	15.73			
Class	7.24			
Model Fit	R^2_{marg}= 0.77; R^2_{cond}= 0.87			

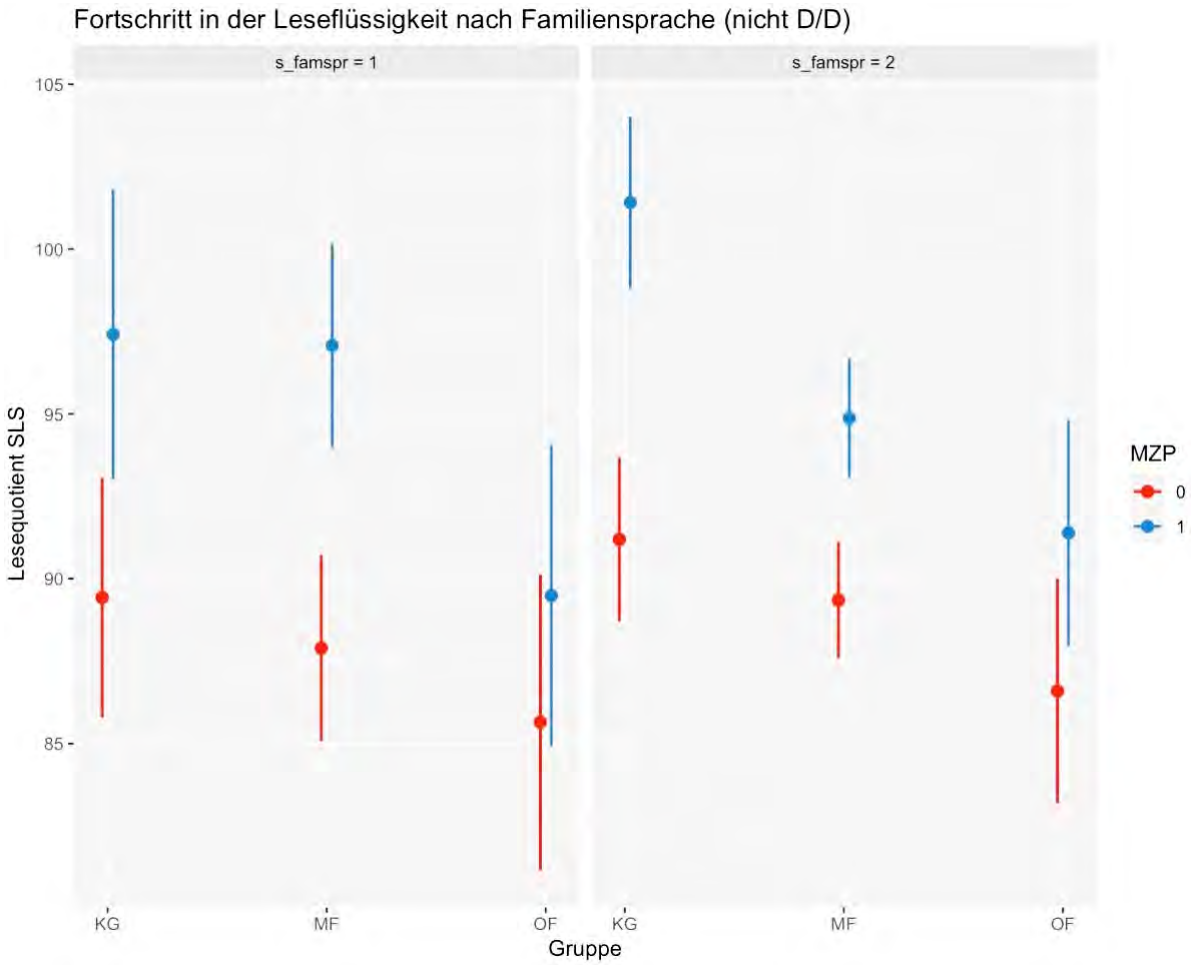
Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.5 Development of reading fluency by group affiliation and family language, controlling for treatment duration, reading-related self-concept and motivation (1 and 2) and demographic characteristics

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 222.64, df = 12, p < .001). The ICC was greater than 0.1 (icc = .79).

Fig. 11: Progress in reading fluency by group membership and family language, controlling for treatment duration, reading-related self-concept and motivation (1 and 2), and demographic characteristics

Reading quotient (SRS) vs TPE according to family language (second language learners/native) of each group: CG, F, W/O F



Tab. 26: Regression coefficients of the analysis of progress in reading fluency as a function of group membership and baseline performance, controlling for treatment duration, initial self-concept and motivation as well as gender, socio-economic status and family language

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	55.95	46.71 – 65.19	<0.001	***
Treatment duration	-0.08	-0.19 – 0.02	0.129	
Socioeconomic status	0.24	-0.09 – 0.56	0.152	
Gender	0.07	-1.11 – 1.25	0.906	
Family Language (G)	1.76	-1.48 – 5.01	0.287	
Self-concept	0.28	0.05 – 0.50	0.018	*
Reading motivation 1	0.81	0.46 – 1.16	<0.001	***
Reading motivation 2	0.02	-0.15 – 0.19	0.801	
Baseline performance	0.98	10.32 – 11.64	<0.001	***
TPE	5.74	-3.31 – 14.79	0.214	
Group F	-1.22	-9.48 – 7.05	0.772	
Group W/OF	-2.96	-13.10 – 7.17	0.566	
Family Language (G) × TPE	2.24	-2.53 – 7.01	0.357	
Family Language (G) × Group F	-0.31	-4.58 – 3.96	0.886	
Family Language (G) × Group W/OF	-0.82	-5.67 – 4.04	0.741	
TPE × Group F	7.09	-4.06 – 18.24	0.212	
TPE × Group W/OF	2.86	-14.85 – 9.13	0.640	
(Family Language (G) × TPE) × Group F	-5.89	-11.79 – 0.00	0.050	
(Family Language (G) × TPE) × Group W/OF	-1.28	-7.64 – 5.08	0.693	
Random effects				
Residual	39.47			
Child:Class	16.04			
Class	7.14			
Model Fi	R^2_{marg}= 0.76; R^2_{cond}= 0.85			

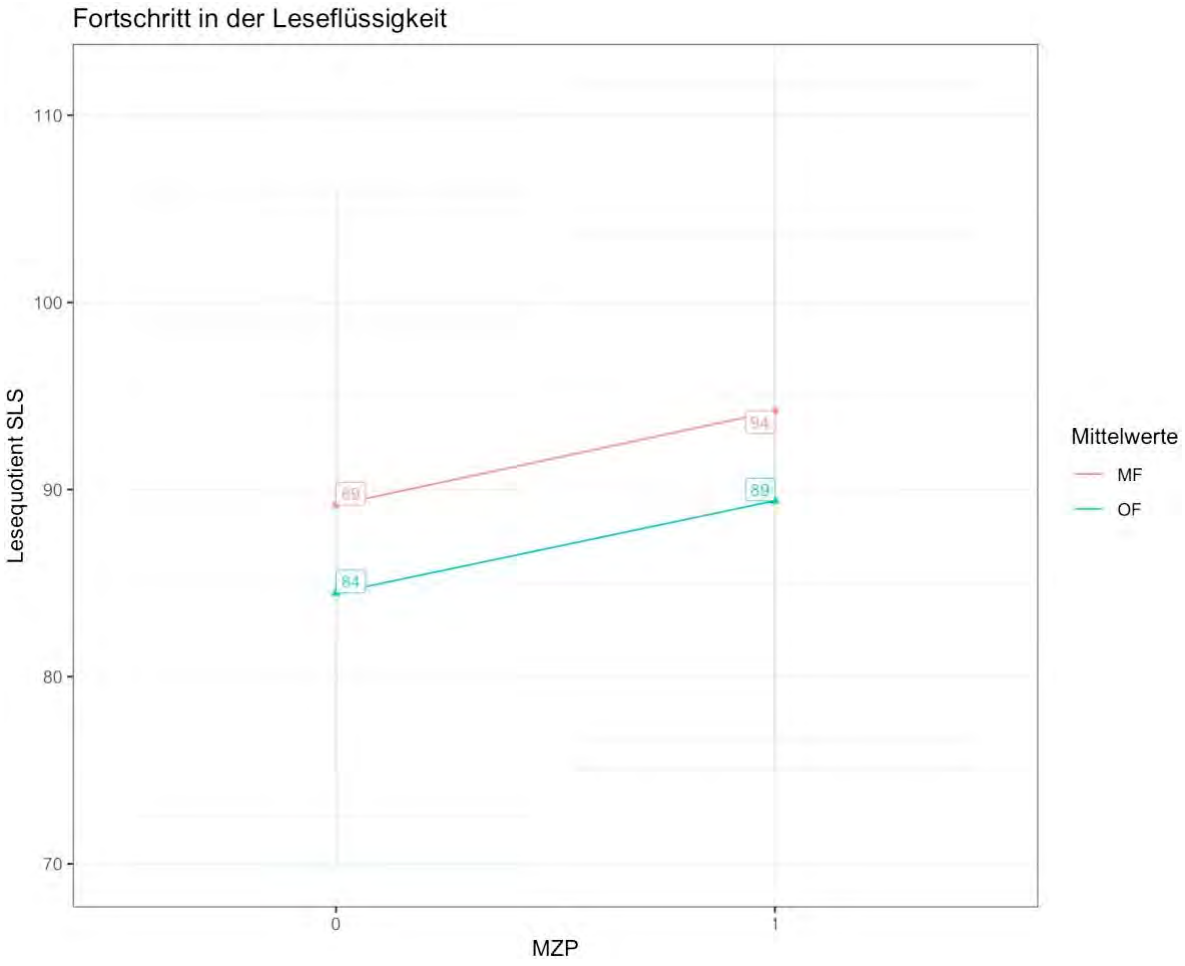
Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.6 Experimental groups only: Development of reading fluency by group membership controlling for treatment duration

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 153.94, df = 4, p < .001). The ICC was greater than 0.1 (icc = .85).

Fig. 12: Progress in reading fluency in the treatment groups by group membership, controlling for treatment duration

Reading quotient (SRS) vs TPE, only experimental groups: F, W/O F



Tab. 27: Regression coefficients of the analysis of progress in reading fluency as a function of group membership (F & W/OF only) controlling for treatment duration

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (F)	85.45	70.82 – 100.08	<0.001	***
Treatment duration	0.02	-0.23 – 0.28	0.864	
TPE	6.39	5.32 – 7.45	<0.001	***
Group W/OF	-4.90	-15.31 – 5.52	0.356	
TPE × Group W/OF	-1.84	-3.64 – -0.03	0.046	*
Random effects				
Residual	42.27			
Child:Class	149.47			
Class	94.63			
Model Fit	R^2_{marg}= 0.061; R^2_{cond}= 0.861			

Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.7 Experimental groups only: Development of reading fluency by group affiliation and baseline performance, controlling for treatment duration, self-concept, motivation (1 and 2), socioeconomic status, gender and family language

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 158.89.48, df = 14, p < .001). The ICC was greater than 0.1 (icc = .37).

Fig. 13: Progress in reading fluency by group membership and baseline performance, controlling for treatment duration, self-concept, motivation (1 and 2) as well as socioeconomic status, gender and family language (by level and group)

Reading quotient (SRS) vs TPE according to base line performance (sub groups 1,2,3,4, low/high performance) , only experimental groups: F, W/O F

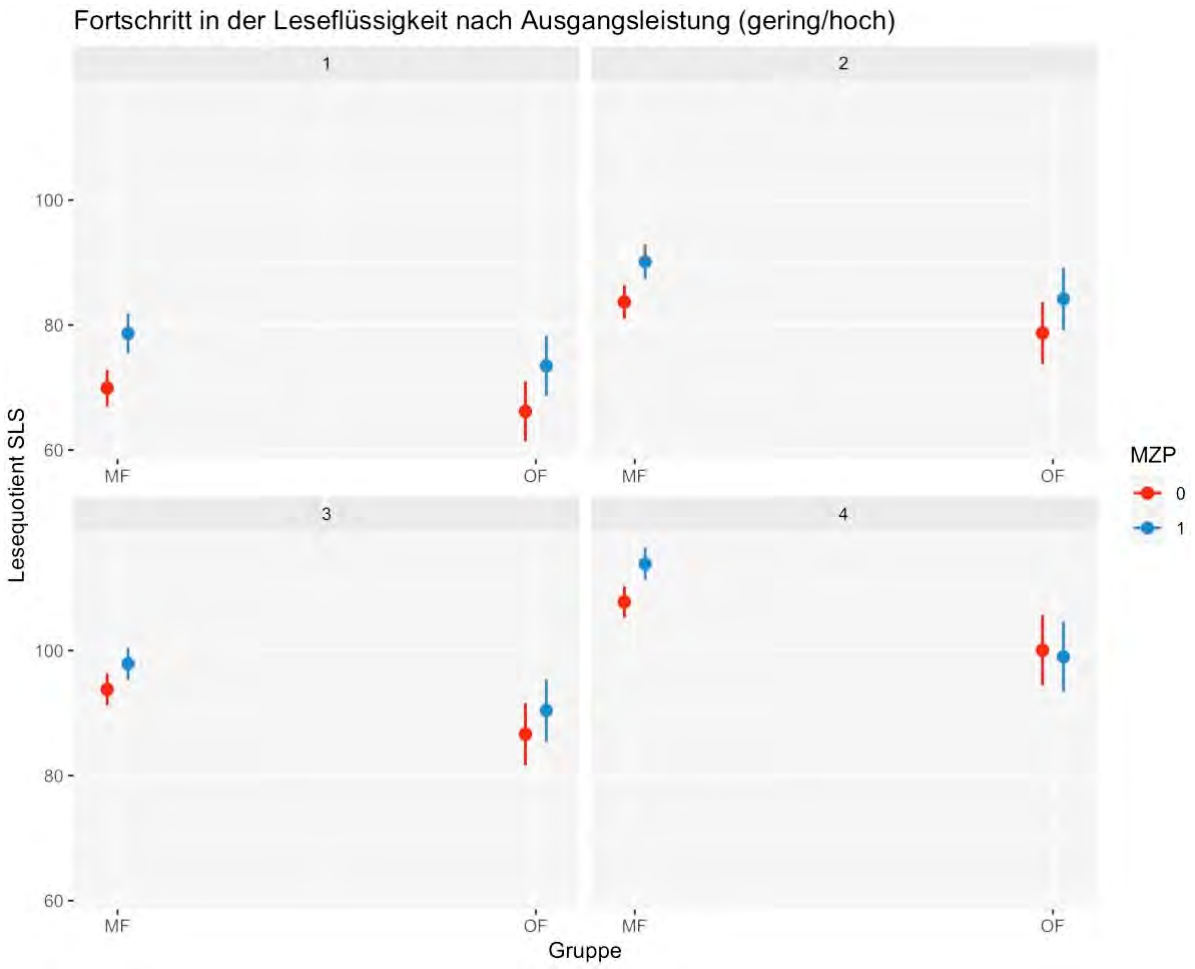


Fig. 14: Progress in reading fluency by groups affiliation and baseline performance, controlling for treatment duration, self-concept, motivation (1 and 2) and socioeconomic status, gender and family language (by groups and level)

Reading quotient (SRS) vs TPE according to base line performance (sub groups 1,2,3,4, low/high performance) , only experimental groups: F, W/O F

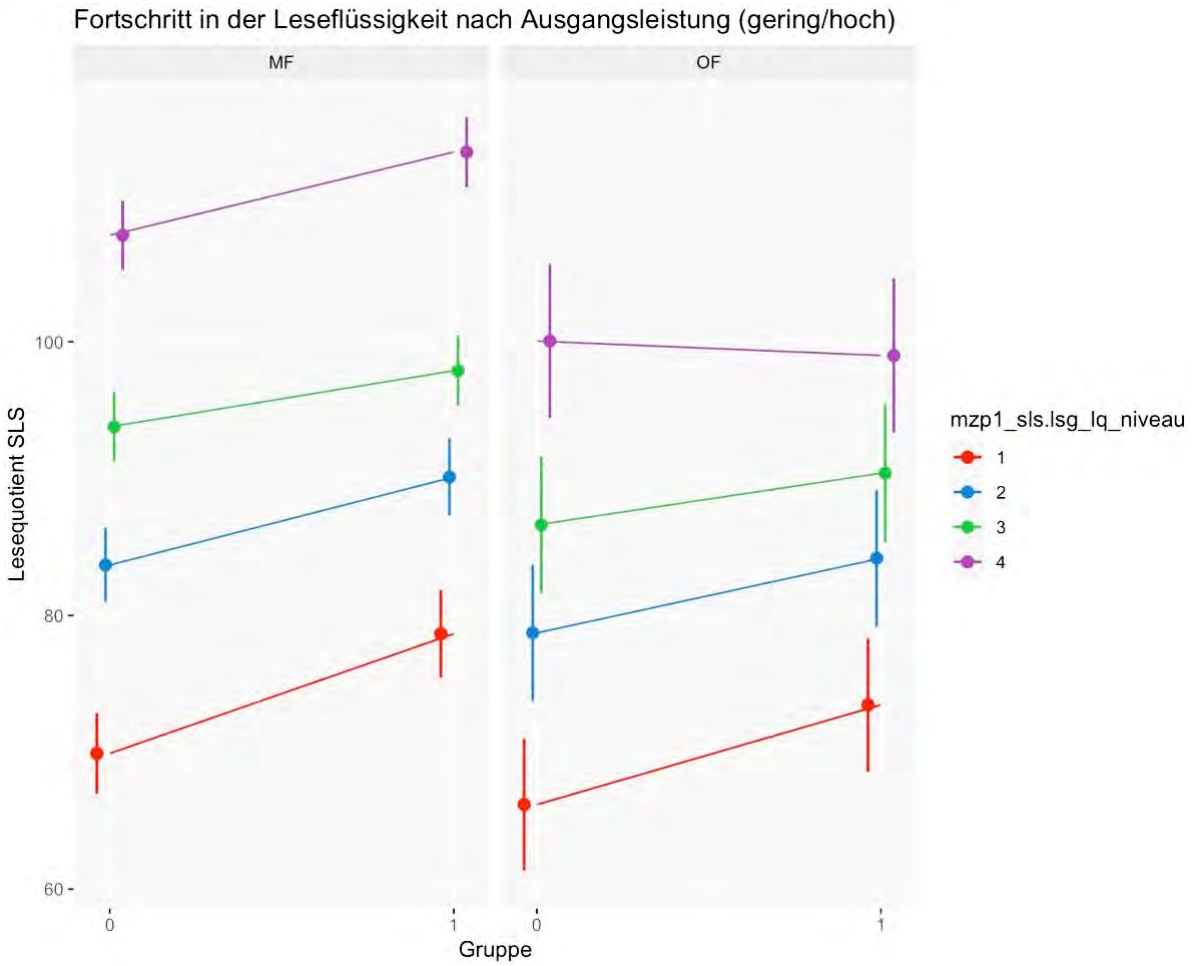
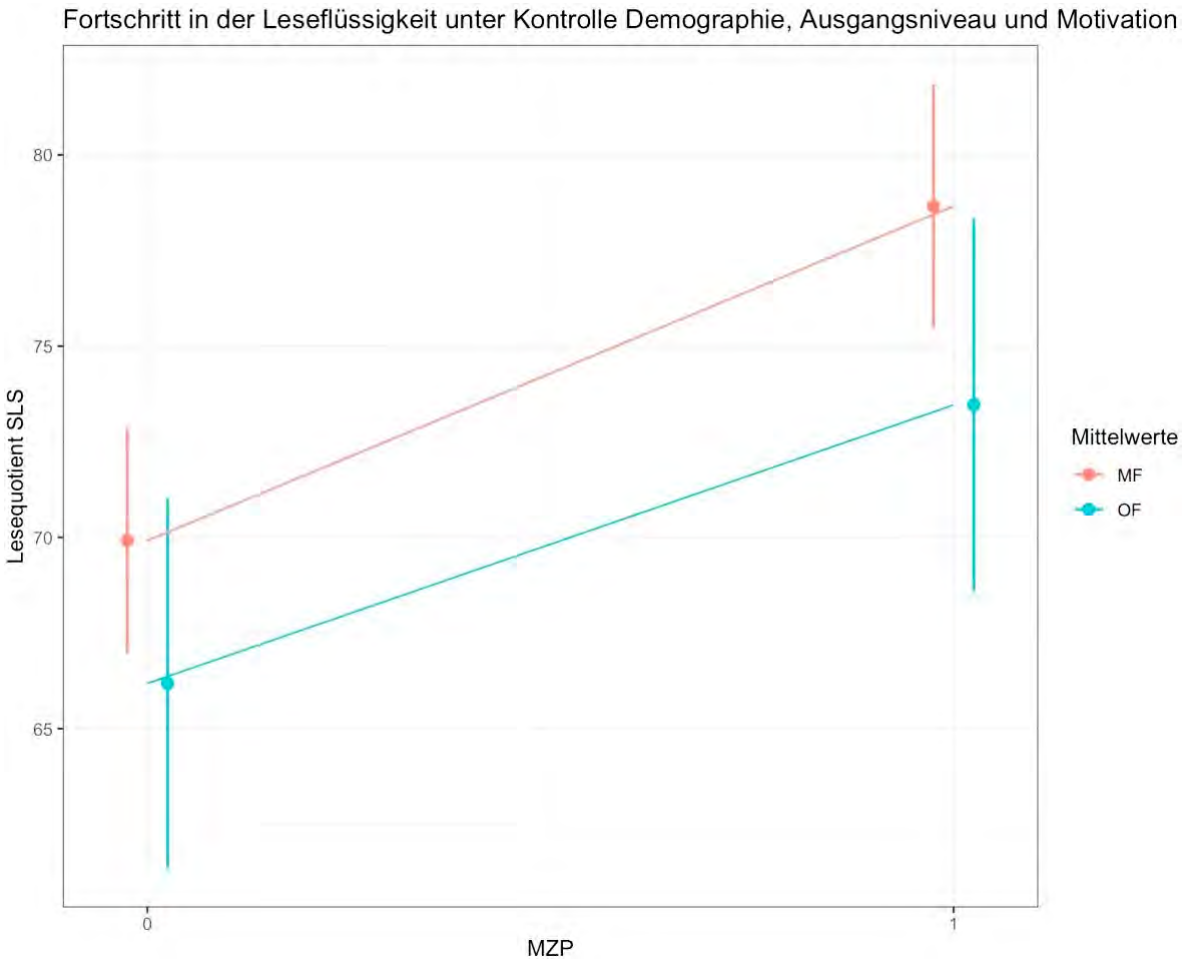


Fig. 15: Progress in reading fluency by group membership and baseline performance, controlling for treatment duration, self-concept, motivation (1 and 2) and socioeconomic status, gender and family language (total)

Reading quotient (SRS) vs TPE controlling for demographics characteristics, base line performance and motivation.



Tab. 28: Regression coefficients of the analysis of progress in self-concept as a function of group membership and baseline performance, controlling for treatment duration, self-concept, motivation (1 and 2) as well as socioeconomic status, gender and family language

	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (F)	125.96	68.15 – 183.77	<0.001	***
Treatment duration	-1.34	-2.51 – -0.17	0.025	*
Socioeconomic status	0.12	-0.55 – 0.79	0.726	
Gender	0.25	-2.21 – 2.72	0.841	
Family Language	-0.02	-2.89 – 2.86	0.991	
Self-concept	-0.21	-0.79 – 0.37	0.478	
Reading motivation 1	0.80	0.08 – 1.52	0.031	*
Reading motivation 2	-0.05	-0.45 – 0.36	0.819	
Baseline performance	9.99	7.39 – 12.59	<0.001	***
TPE	6.30	-2.56 – 15.16	0.162	
Group W/OF	-32.01	-58.89 – -5.12	0.020	*
TPE × Baseline performance	0.11	-2.78 – 2.99	0.942	
TPE × Group W/OF	2.31	-7.79 – 12.42	0.652	
Baseline performance × Group W/OF	0.32	-2.73 – 3.37	0.836	
(Baseline performance × TPE) × Group W/OF	-2.17	-5.65 – 1.31	0.219	
Random effects				
Residual	28.18			
Child:Class	9.81			
Class	6.44			
Model Fit	R^2_{marg}= 0.788; R^2_{cond}= 0.865			

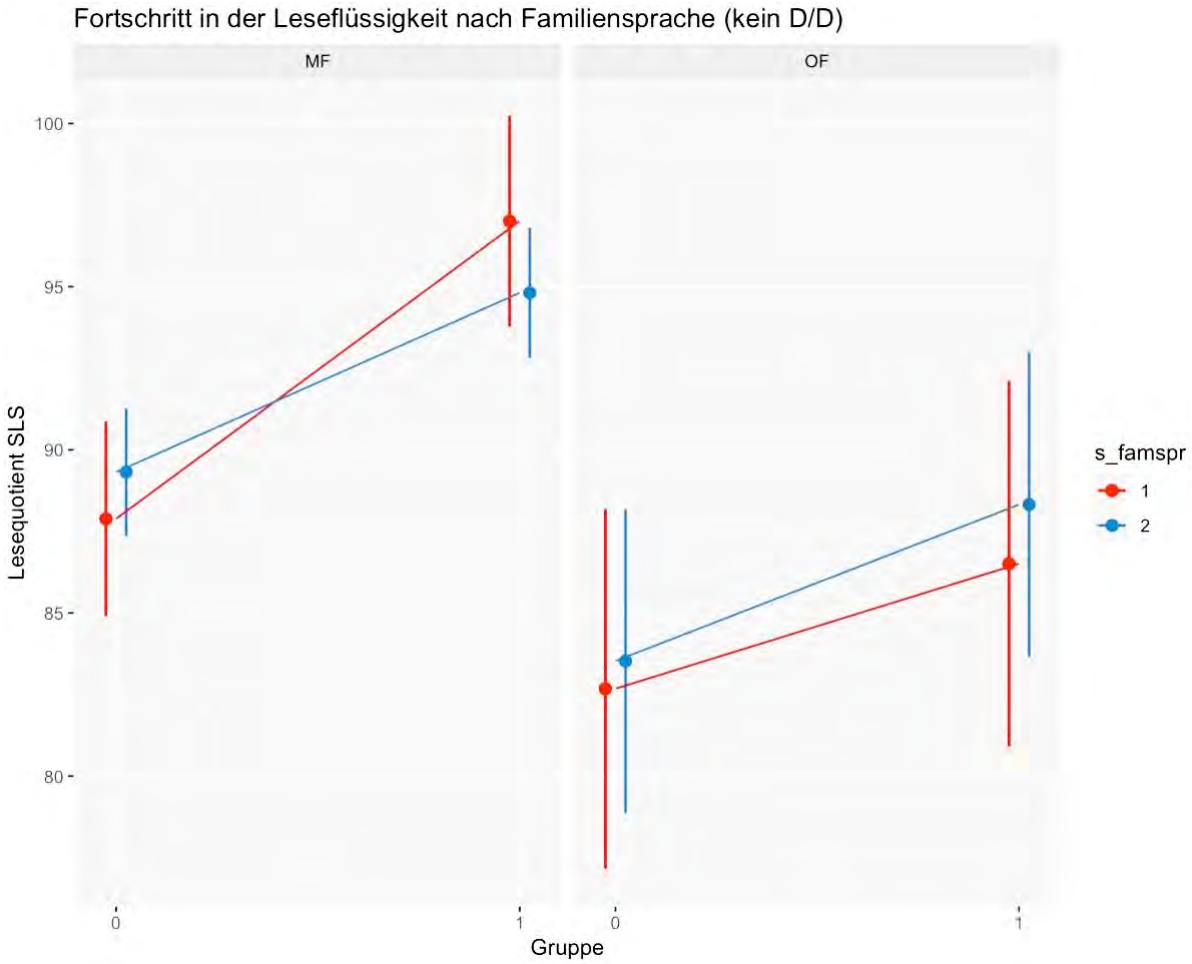
Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.2.8 F & W/OF only: Development of reading fluency by group affiliation and family language, controlling for treatment duration, self-concept, motivation (1 and 2), socioeconomic status, gender and baseline performance

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 605.57, df = 14, p < .001). The ICC was greater than 0.1 (icc = .36).

Fig. 16: Progress in reading fluency by group affiliation and family language, controlling for treatment duration, self-concept, motivation (1 and 2), socioeconomic status, gender and baseline performance (by group and family language (schooling language is not spoken at home, schooling language is spoken at home)

Reading quotient (SRS) vs TPE according to family language (second language learners/native), experimental groups only.



Tab. 29: Regression coefficients of the analysis of progress in self-concept as a function of group membership and family language, controlling for treatment duration, self-concept, motivation (1 and 2), socioeconomic status, gender and baseline performance

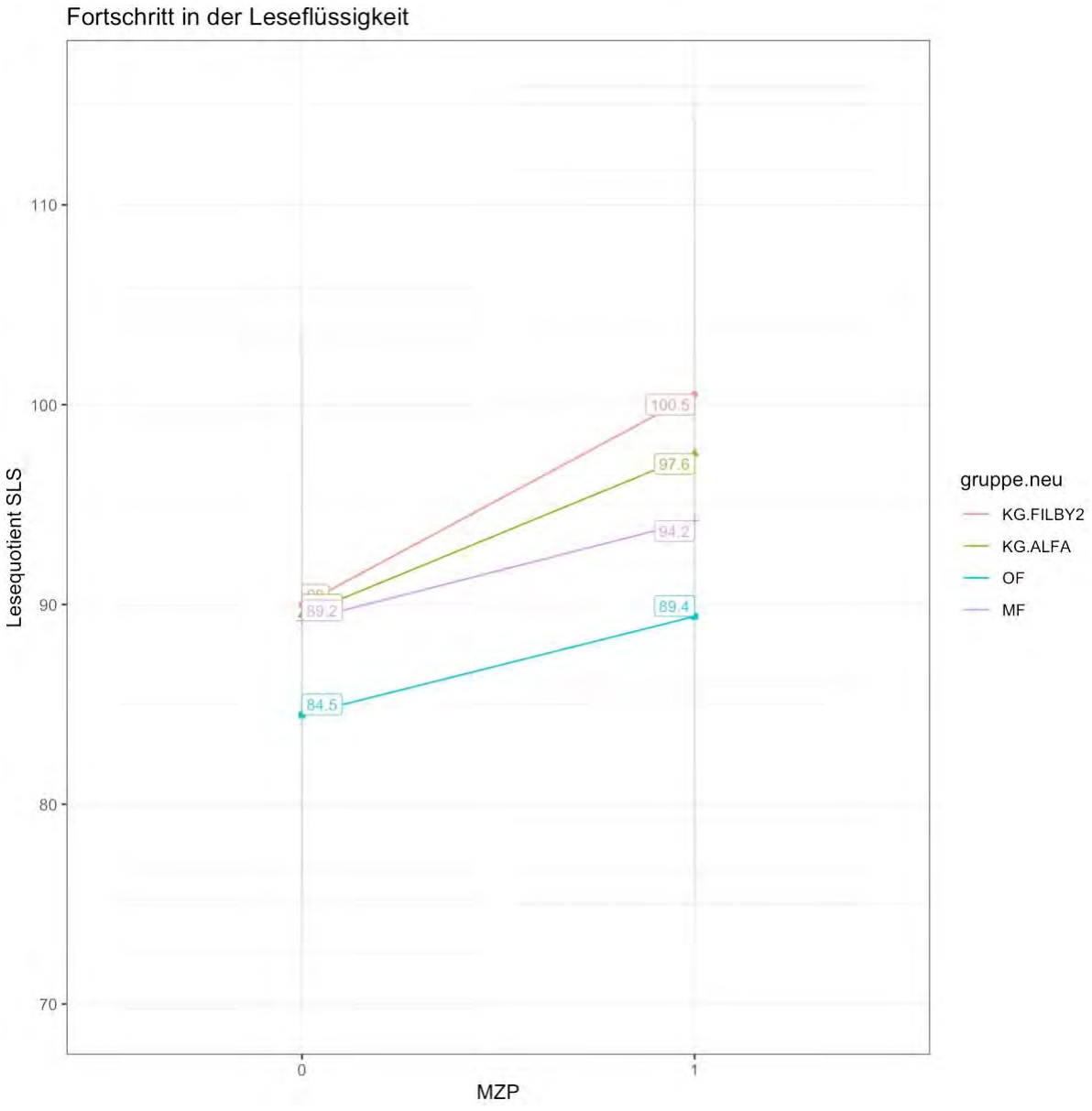
	Salzburger Reading Screening LQ			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (F)	62.46	52.60 – 72.33	<0.001	***
Treatment duration	-0.19	-0.37 – -0.02	0.026	*
Socioeconomic status	0.07	-0.31 – 0.46	0.704	
Gender	0.61	-0.83 – 2.04	0.406	
Baseline performance	11.02	10.21 – 11.83	<0.001	***
Self-concept	0.26	-0.03 – 0.54	0.075	
Reading motivation 1	0.69	0.28 – 1.09	0.001	***
Reading motivation 2	0.06	-0.14 – 0.26	0.577	
Family Language	1.43	-1.41 – 4.27	0.323	
TPE	9.12	5.82 – 12.43	<0.001	***
Group W/OF	-5.21	-11.95 – 1.53	0.130	
TPE × Family Language	-3.63	-7.24 – -0.02	0.048	*
TPE × Group W/OF	-5.29	-10.46 – -0.12	0.045	
Family Language × Group W/OF	-0.58	-5.22 – 4.05	0.805	
(Family Language × TPE) × Group W/OF	4.59	-1.07 – 10.25	0.111	
Random effects				
Residual	42.38			
Child:Class	15.06			
Class	8.58			
Model Fit	R^2 marg= 0.745; R^2 cond= 0.836			

6.2.9 With CG out of FiLBY-2: Development of reading fluency by group membership controlling for treatment duration

The model with groups and TPE variables and their interaction fit significantly better than the UREMs (null model) and was therefore selected for analysis (c2 = 480.24, df = 8, p < .001). The ICC was greater than 0.1 (icc = .85).

Fig. 17: Progress in reading fluency according to groups

Reading quotient (SRS) vs TPE including new groups: CG for an established paper-based reading intervention program, CG in the project ALFA, F, W/O F



Note: *N* = 1.241 (of which 221 from the CG of FiLBY-2: only SLS A/B and complete data for TPE1 and 2 of FiLBY-2)

Tab. 30: Regression coefficients of the analysis of progress in reading fluency as a function of group membership, controlling for treatment duration

	Salzburger Reading Screening LQ		
Fixed Effect	Regression-coefficient	CI	p
Intercept (CG.FiLBY)	91.10	67.45 – 114.74	<0.001
Treatment duration	-0.01	-0.18 – 0.16	0.914
TPE	10.55	9.36 – 11.74	<0.001
Group CG.ALFA	-2.74	-17.34 – 11.85	0.712
Group W/OF	-9.33	-29.33 – 10.68	0.361
Group F	-3.84	-18.62 – 10.93	0.610
TPE × Group CG.ALFA	-1.51	-3.41 – 0.38	0.118
TPE × Group W/OF	-6.00	-7.86 – -4.14	<0.001
TPE × Group F	-4.17	-5.75 – -2.59	<0.001
Random effects			
Residual	40.62		
Child:Class	159.46		
Class	65.97		
Model Fit	R^2_{marg}= 0.094; R^2_{cond}= 0.862		

Note.: ***p < 0.001; **p < 0.01; *p < 0.05

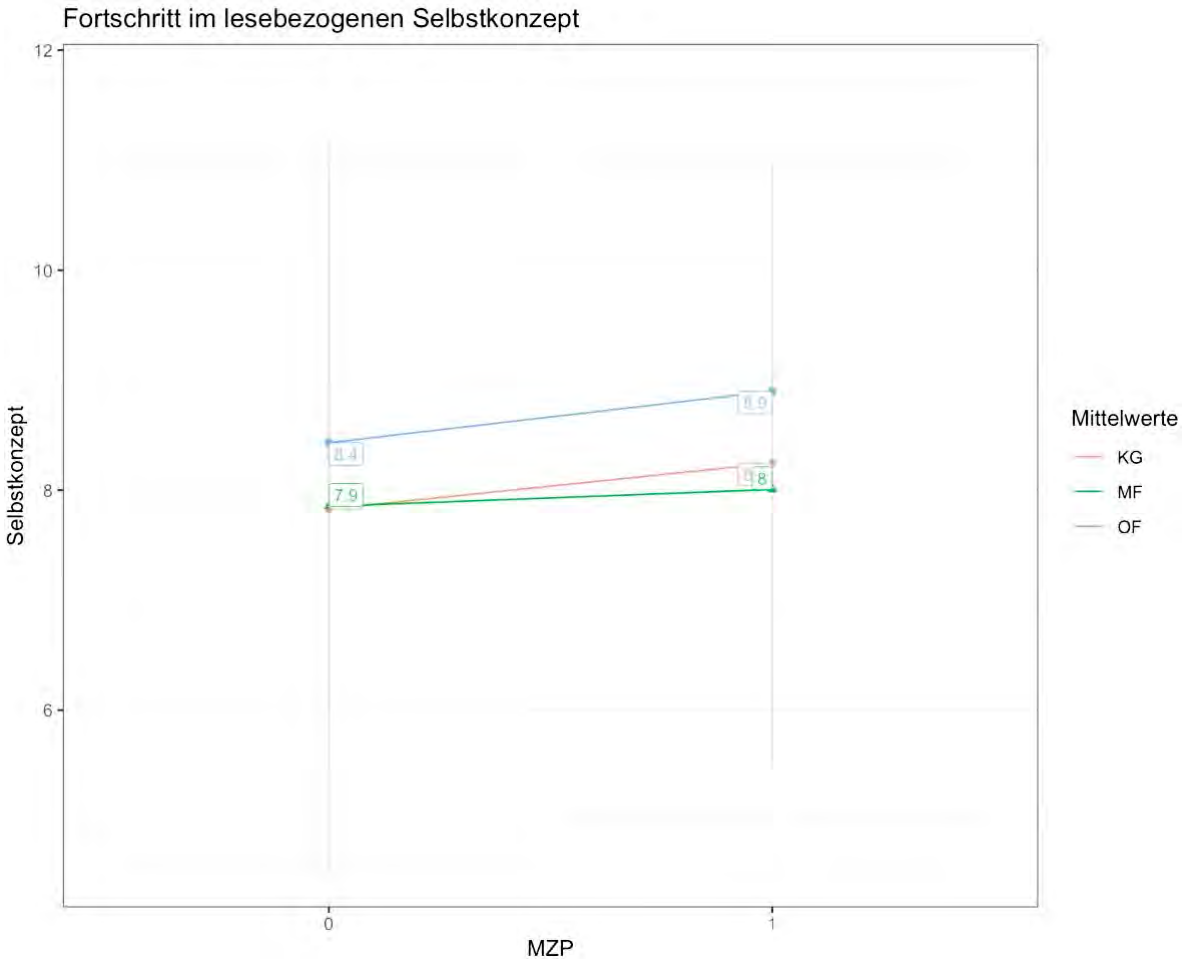
6.3 Reading-related self-concept and reading motivation (1/2)

6.3.1 Development of reading-related self-concept

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 27.31, df = 6, p < .001). The ICC was greater than 0.1 (icc = .61).

Fig. 17: Development of reading-related self-concept according to Group

Self-concept vs TPE for each group: CG, F, W/O F



Tab. 31: Regression coefficients of the analysis of progress in self-concept as a function of group membership controlling for treatment duration

	Reading Related Self-concept			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	6.38	4.92 – 7.85	<0.001	***
Treatment duration	0.02	-0.00 – 0.04	0.056	
TPE	0.55	0.13 – 0.97	0.010	**
Group F	-0.12	-0.83 – 0.58	0.731	
Group W/OF	1.36	0.17 – 2.55	0.025	*
TPE × Group F	-0.02	-0.54 – 0.51	0.953	
TPE × Group W/OF	-0.14	-0.74 – 0.46	0.642	
Random effects				
Residual	3.43			
Child:Class	4.55			
Class	0.80			
Model Fit	R^2_{marg}= 0.022; R^2_{cond}= 0.618			

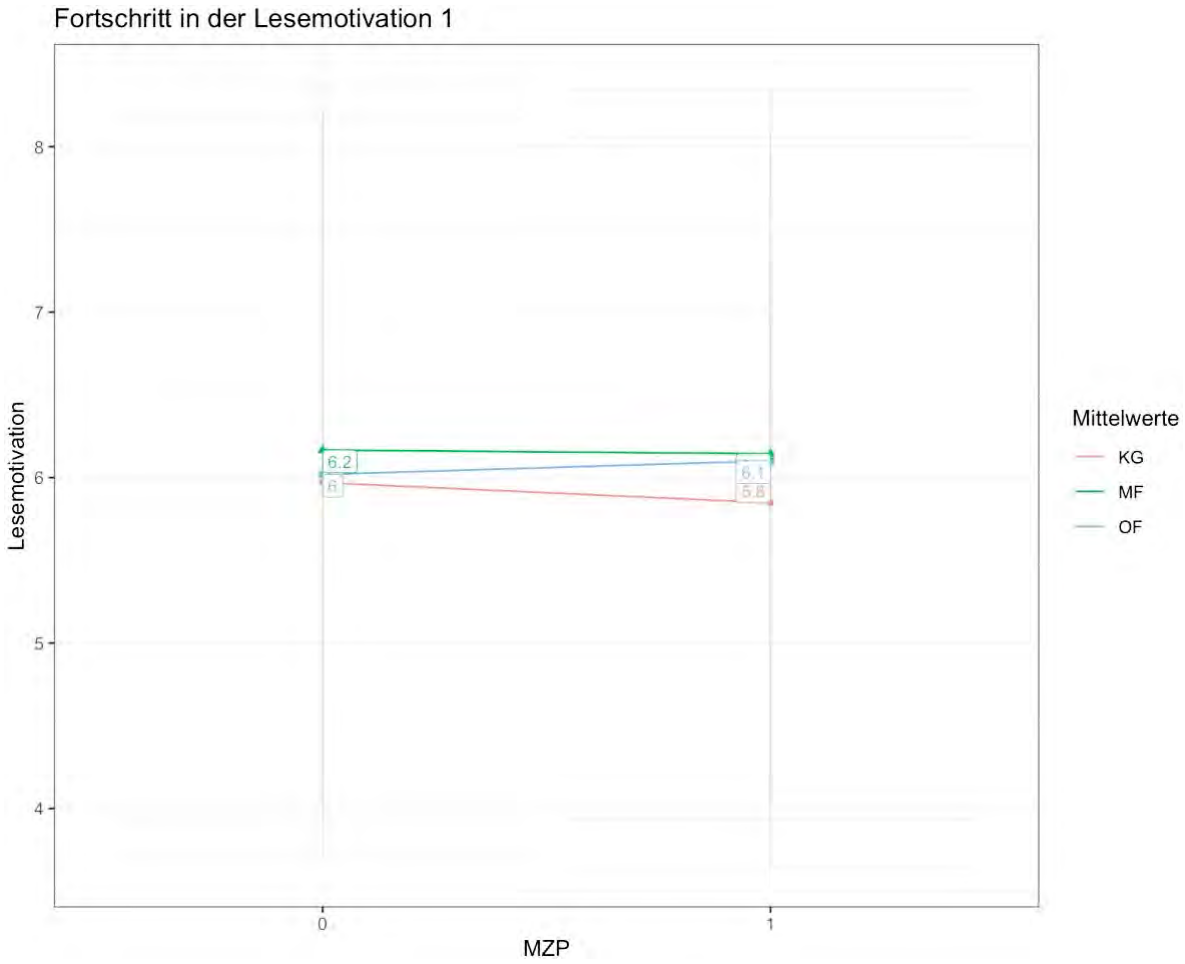
Note: ***p < 0.001; **p < 0.01; *p < 0.05

6.3.2 Development of Reading Motivation 1

The model with groups and TPE variables and their interaction did not fit significantly better than the UREMs (c2 = 2.6642, df = 6, p = .85). The ICC was greater than 0.1 (icc = .58).

Fig. 18: Development of reading motivation 1 by Group

Reading motivation 1 vs TPE for each group: CG, F, W/O F

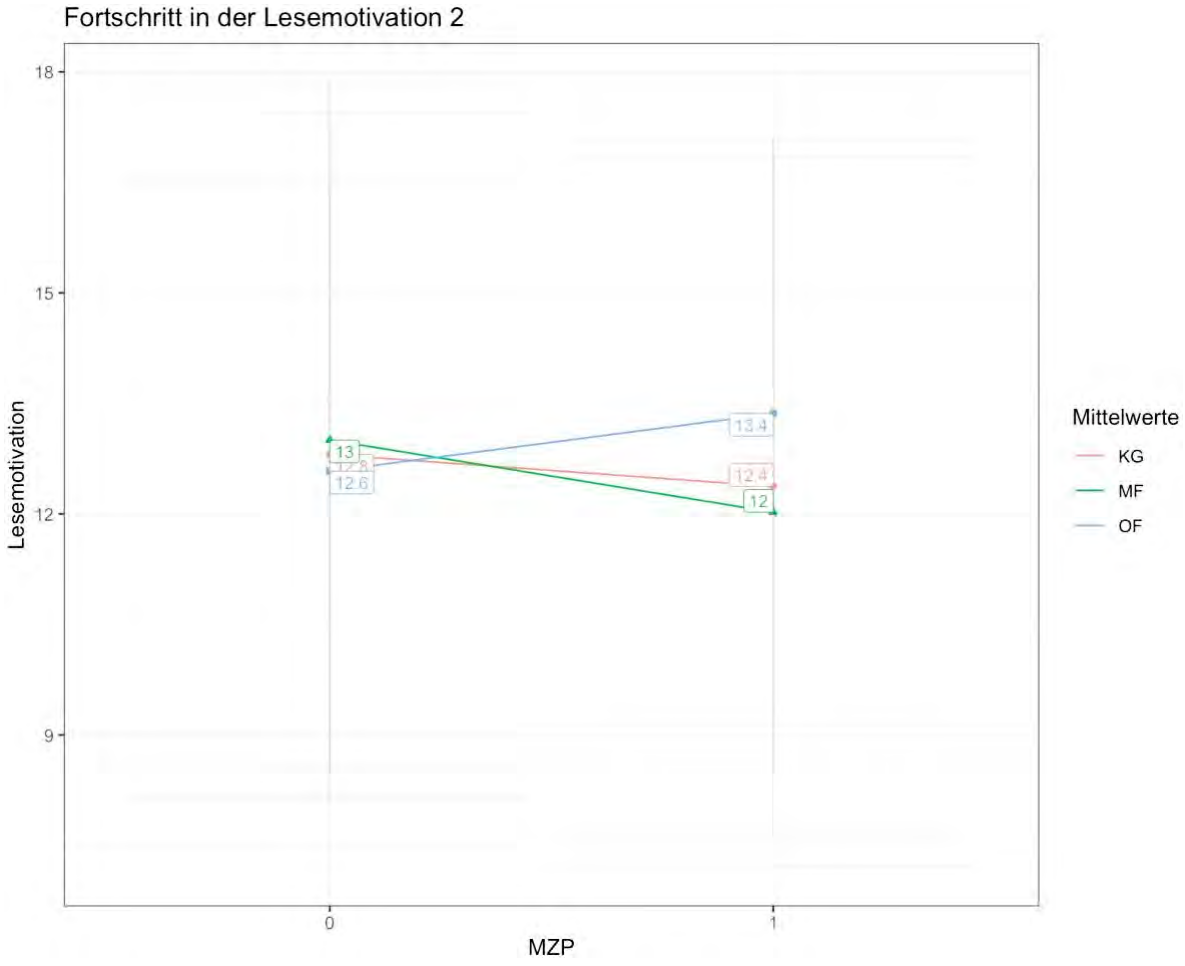


6.3.3 Development of reading motivation 2

The model with groups and TPE variables and their interaction fit significantly better than the UREMs and was therefore selected for analysis (c2 = 17.229, df = 6, p < .01). The ICC was greater than 0.1 (icc = .60).

Fig. 19: Development of Reading motivation 2 by Group

Reading motivation 2 vs TPE for each group: CG, F, W/O F



Tab. 32: Regression coefficients of the analysis of progress in reading motivation 2 as a function of group membership, controlling for treatment duration

	Reading Motivation 2			
Fixed Effect	Regression-coefficient	CI	p	
Intercept (CG)	10.04	8.05 – 12.03	<0.001	***
Treatment duration	0.04	0.01 – 0.07	0.008	**
TPE	-0.17	-0.86 – 0.51	0.617	
Group F	0.31	-0.67 – 1.29	0.532	
Group W/OF	1.28	-0.33 – 2.89	0.119	
TPE × Group F	-0.43	-1.28 – 0.42	0.322	
TPE × Group W/OF	0.93	-0.04 – 1.90	0.060	
Random effects				
Residual	8.76			
Child:Class	11.91			
Class	1.00			
Model Fit	R^2_{marg}= 0.014; R^2_{cond}= 0.601			

Note: ***p < 0.001; **p < 0.01; *p < 0.0

7. Discussion

The ALFA2 project compared the effectiveness of the eKidz.eu app with the already positively evaluated FiLBY - an established paper-based reading program. The FiLBY reading program has proven to be effective in previous studies for improving reading fluency and reading comprehension in elementary school children (Schilcher et al., 2023). It is based on an active view of reading approach that includes elements such as repeated reading, targeted feedback, and strategy training (Schilcher, Wild, & Steinert, 2019). The positive evaluation results of the FiLBY reading program serve as a quality seal and reference point for evaluating new digital reading intervention programs such as the eKidz.eu app (Schilcher, Glondys, & Wild, 2023; Wild & Schilcher, 2019).

The results of the ALFA2 project show that the eKidz.eu app achieves comparable effects to the FiLBY reading program. This indicates that eKidz - a digital learning application is an effective tool for reading promotion. While both programs pursue similar methodological approaches (assisted reading, repeated reading; Wild, & Schilcher, 2019), the eKidz.eu app can specifically utilize digital mechanisms such as adaptive focus on competencies, gamification elements, or automated feedback mechanisms (cf. Schilcher, & Wild, 2021). The fact that the group with feedback in the ALFA2 project showed clear improvements in reading fluency indicates that the feedback mechanisms implemented in the eKidz.eu app were effective. **Reading fluency improved significantly in all groups, with the feedback group (F) achieving the best absolute progress (d = .29 to .49).**

Feedback that is tailored to the individual needs and current performance level of the learner can have positive effects (Kluger, & DeNisi, 1996; Hattie, & Timperley, 2007). It can help to specifically address errors and difficulties, sharpen awareness of strengths and weaknesses, and steer learning efforts more effectively. In the context of reading promotion, personalized feedback could include, for example, hints about incorrectly read words, suggestions for improving reading speed, or strategies for better text comprehension (Wild, Cruse, & Pfister, 2020). The results of ALFA2 show that children can improve their reading skills more efficiently through targeted feedback. The improvement in reading-related self-concept in the group without feedback (d = .18) could indicate that children who learn without external feedback develop a stronger sense of self-regulation and self-efficacy. This could lead to more sustainable motivation in the long term.

The ALFA2 project emphasizes the effectiveness of training that is independent of initial motivation or baseline reading performance i.e. eKidz.eu. This finding points to the potential of such technology and AI powered learning tools to reduce educational inequalities and promote the reading competencies of children from various backgrounds (cf. also AG Leseverstehen, 2023). Overall, the ALFA2 project contributes to the growing evidence that personalized, technology-supported learning offerings have positive effects on reading promotion, especially when they are continuously developed.

References

AG Leseverstehen (2023). Lesen professionell unterrichten – Lehrkräftebildung auf dem Prüfstand. doi.org/10.5283/epub.55097

Bos, W., Strietholt, R., Goy, M., Stubbe, T. C., Tarelli, I. & Hornberg, S. (2010). *IGLU 2006. Dokumentation der Erhebungsinstrumente*. Münster, New York, München, Berlin : Waxmann. https://doi.org/10.25656/01:15670

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), S. 81-112. https:// doi.org/10.3102/003465430298487

Kluger, A. N. & DeNisi, A. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*, 119(2), S. 254-284. https://doi.org/10.1037/00332909.119.2.254

Schaufelberger, R., Kleinkorres, R., Becher, L., Ludewig, U., Lorenz, R. & McElvany, N. (2024). *IGLU 2021. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster; New York: Waxmann. http://doi.org/10.25656/01:30762

Schilcher, A. & Wild, J. (2021). Lese-Apps auf dem Prüfstand. *BiSS-Transfer Journal*, 14, 710.

Schilcher, A., Glondys, M. & Wild, J. (2023). Leseunterricht in den Grundschulen in Deutschland. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 179-196). Münster: Waxmann.

Schilcher, A., Wild, J. & Steinert, M. (2019). *FiLBY-2 Lehrerhandreichung. Fachintegrierte Leseförderung Bayern*. Regensburg: o. V. (online verfügbar online verfügbar den FiLBY-2-Selbstlernkurs unter fibs.alp.dillingen.de)

Schilcher, A., Wild, J., Kraus, E. & Hilbert, S. (2022). FiLBY-2 – ein Leseflüssigkeitstraining für alle Schülerinnen und Schüler? *SLLD-Z*, 2, 1-29. doi.org/10.46586/SLLD.Z.2022.9606

Wild, J. & Schilcher, A. (2019). Grundlagen einer systematischen schulischen Leseförderung. In G. Stückl & M. Wilhelm (Hrsg.), *Lehren und Lernen in der bayerischen Grundschule* (S. 1-17). Kronach, Köln: Carl Link Verlag.

Wild, J., Cruse, M. & Pfister, W. (2024). *Die Entwicklung der Leseflüssigkeit mit lernwirksamen Feedback unterstützen: Praxismaterial*. (Online verfügbar über https://www.lesen.bayern.de/lesefluessigkeit).