Descriptive data of included studies

Digital appendix of Maskos, K. et al.: Formative assessment in mathematics education. A systematic review.

Table 6 Summarized descriptives of study characteristics and research designs

		N	%
		45	100
Publication year	2015	1	2
	2016	1	2
	2017	5	11
	2018	5	11
	2019	6	13
	2020	3	7
	2021	7	16
	2022	7	16
	2023	10	22
Geographic distribution	Africa	5	11
	Asia	14	31
	Europe	20	45
	North America	6	13
Published in a mathematics education journal		8	18
Education level of participants	Primary (Grades 1–6)	15	33
	Secondary (Grades 7–13)	23	51
	University	7	16
Mathematical learning objectives	Algebra	5	11
	Arithmetic	7	16
	Combinatoric or calculus	2	4
	Geometry	9	20
	General	22	49
Sample size of participating students	20–100	15	33
	100–500	12	27
	500–1,000	8	18
	1,001–3,000	6	13
	> 3,000	4	9
Measurement points	Longitudinal	37	82
	Intervention	*36*	
	Cross-sectional	8	18
	Intervention	*4*	
Intervention		40	89
	By teacher	*30*	
	By researcher	*6*	
	By Intelligent Tutoring System	*2*	
	n/a	*2*	
	With control group	*29*	
	Randomization	*20*	
Content of main effects	Cognitive	24	53
	Noncognitive	9	20
	Both	12	27
Third analytical variable added		23	51
	Mediator	*4*	
	Moderator	*19*	
Type of feedback implemented in FA	Elaborated	39	87
	Right-wrong	1	2
	No info	5	11
Digital media	No digital media	31	69
	Digital media	6	13
	Interactive learning environment	8	18

Table 7 Detailed descriptive data of included studies (*n*=45) investigated effects of formative assessment on students' learning (RQ1)

Author, Year	Title	Journal		Affiliation of 1st author			Students				Teachers
		Name	M.E.[a]	Department & university	Country	Continent	Level	Grade[b]	Age	Sample size	Sample size
Albano et al., 2023	Grasping criteria for success: engaging undergraduate students in formative feedback by means of digital peer workshops	Teaching Mathematics and its Applications: An International Journal of the IMA	1	Department of Information and Electrical Engineering and Applied Mathematics, Università di Salerno	Italy	Europe	University	Freshmen Computer Engineering, Math II	No info	38	0
Alqassab et al., 2018	Training peer-feedback skills on geometric construction tasks: role of domain knowledge and peer-feedback levels	European Journal of Psychology of Education	0	Department of Psychology, Ludwig-Maximilian University Munich	Germany	Europe	University	No info, PST middle school math	22.51 (mean)	43	0
Andersson & Palm, 2017	The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme	Learning and Instruction	0	Department of Science and Mathematics Education, Umeå University	Sweden	Europe	Primary	4	No info	729	25
Ayanaw et al., 2022	The effect of assessment for learning on prospective teachers' learning of algebra through a professional development program	International Electronic Journal of Mathematics Education	1	Science and Mathematics Education, Addis Ababa University	Ethiopia	Africa	University	2nd of 3 years, PST math	No info	129	3
Ayanaw et al., 2023a	The effect of assessment for learning on pre-service mathematics teachers' higher-order-thinking skills in algebra	Journal of Peagogical Research	0	Science and Mathematics Education, Addis Ababa University	Ethiopia	Africa	University	2nd of 3 years, PST math	No info	129	3
Ayanaw et al., 2023b	Professional development on assessment for learning and its effect on pre-service teacher's self-regulated learning	Cogent Education	0	Science and Mathematics Education, Addis Ababa University	Ethiopia	Africa	University	2nd of 3 years, PST math	No info	52	3
Barana et al., 2021	Interactive Feedback for Learning Mathematics in a Digital Learning Environment	Education Sciences	0	Department of Molecular Biotechnology and Health Sciences, University of Turin	Italy	Europe	Secondary	8	No info	399	No info
Barana et al., 2022	Self-Assessment in the Development of Mathematical Problem-Solving Skills	Education Sciences	0	Molecular Biotechnology and Health Sciences, University of Turin	Italy	Europe	Secondary	11	16	182	10
Beesley et al., 2018	Enhancing Formative Assessment Practice and Encouraging Middle School Mathematics Engagement and Persistence	School Science and Mathematics	1	SRI International research institute	United States	North america	Secondary	No info	No info	6177	47
Boström & Palm, 2023	The effect of a formative assessment practice on student achievement in mathematics	Frontiers in Education	0	Department of Science and Mathematics Education, Umeå University	Sweden	Europe	Secondary	7	13	566	29
Bottge et al., 2021	Effects of Formative Assessment Strategies on the Fractions Computation Skills of Students with Disabilities	Remedial and Special Education	0	Department of Early Childhood, Special Education and Counselor Education, University of Kentucky	United States	North america	Secondary	6-12 special education	No info	110	19
Can, 2021	Investigating of Primary School Teacher Candidates' Mathematics Teaching Performance According to Peer Assessments: Example of Teaching Practice	International Journal of Progressive Education	0	Primary Education, Erciyes University	Turkey	Asia	University	PST primary school	20-22 (range)	58	0
Chan et al., 2023	Assessing the effects of a collaborative problem-based learning and peer assessment method on junior secondary students' learning approaches in mathematics using interactive online whiteboards during the COVID-19 pandemic	International Journal Mobile Learning and Organisation	0	Applied Mathematics, The Hong Kong Polytechnic University	Hong Kong (China)	Asia	Secondary	7	No info	87	No info
Chang et al., 2021	Promoting students' cross-disciplinary performance and higher order thinking: a peer assessment-facilitated STEM approach in a mathematics course	Education Tech Research	0	Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology	Taiwan (China)	Asia	University	Junior high school	12 (mean)	112	No info
Chen et al., 2020	Formative assessment with interactive whiteboards: A one-year longitudinal study of primary students' mathematical performance	Computers & Education	0	Chinese Academy of Education Big Data, Qufu Normal University	China	Asia	Primary	2	No info	178	2

Descriptive data of included studies

Author, Year	Title	Journal		Affiliation of 1st author			Students				Teachers
		Name	M.E.[a]	Department & university	Country	Continent	Level	Grade[b]	Age	Sample size	Sample size
Faber et al., 2017	The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment	Computers & Education	0	Department of Research Methodology, Measurement and Data Analysis, University of Twente	Netherlands	Europe	Primary	3	8-9 (range)	1738	79 (No. of participating schools)
Ganajová et al., 2021	Formative assessment as a tool to enhance the development of inquiry skills in science education	Journal of Baltic Science Education	0	Faculty of Science, Pavol Jozef Šafárik University in Košice	Slovakia	Europe	Secondary	1st year upper sec.	15-16 (range)	121	11
Guiaya & Bueno, 2019	Strategic Questioning as Formative Assessment in Teaching Mathematics Grade 9 Learners	Institutional Multidisciplinary Research and Development Journal	0	Master Teacher I, GHNHS; 2nd: Dean-Editor, G-SPACE	Philippines	Asia	Secondary	9	13-19 (range)	125	No info
Guo et al., 2019	Teacher feedback and students' self-regulated learning in mathematics: A comparison between a high-achieving and a low-achieving secondary schools	Studies in Educational Evaluation	0	Faculty of Education, East China Normal University	China	Asia	Secondary	No info	16.5 (mean)	1282	No info
Hatami, 2015	The effect of collaborative learning and self-assessment on self-regulation	Educational Research and Reviews	0	Faculty of Mathematics, University of Sistan and Baluchestan	Iran	Asia	Secondary	11	No info	75	No info
Kartono & Shora, 2020	Effectiveness of Process Oriented Guided Inquiry Learning with Peer Feedback on Achieving Students' Mathematical Reasoning Capabilities	International Journal of Instruction	0	Faculty of Mathematics, Semarang State University	Indonesia	Asia	Secondary	7	No info	64	No info
Kültür & Kutlu, 2021	The effect of formative assessment on high school students' mathematics achievement and attitudes	Journal of Pedagogical Research	0	Ministry of National Education; 2nd: Faculty of Education, Çukurova University	Turkey	Asia	Secondary	10	15-16 (range)	51	0
Kyaruzi et al., 2019	Students' formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania	Assessment in Education: Principles, Policy & Practice	0	Faculty of Education, University of Dar es Salaam & Ludwig-Maximilian University Munich, Germany	Tanzania	Africa	Secondary	11	16.5 (mean)	2767	No info
Moyo et al., 2022	Evaluating the Impact of Formative Assessment Intervention and Experiences of the Standard 4 Teachers in Teaching Higher-Order-Thinking Skills in Mathematics	Frontiers in Education	0	Science, Mathematics and Technology Education, University of Pretoria	South Africa	Africa	Primary	4	No info	272	9
Ng et al., 2020	Active Learning in undergraduate mathematics tutorials via cooperative problem-based learning and peer assessment with interactive online whiteboards	Asia-Pacific Education Researcher	0	Department of Curriculum and Instruction, Faculty of Education, The Chinese University of Hong Kong	Hong Kong (China)	Asia	University	1st year, undergraduate	No info	830	No info
Patchan et al., 2022	Using peer assessment to improve middle school mathematical communication	The journal of educational research	0	Department of Learning Sciences and Human Development, West Virginia University	United States	North america	Primary	6	11.8-13 (range), 12.1 (mean)	32	No info
Pinger et al., 2018a	Implementation of formative assessment – effects of quality of programme delivery on students' mathematics achievement and interest	Assessment in Education: Principles, Policy and Practice	0	Department of Educational Quality and Evaluation, German Institute for International Educational Research	Germany	Europe	Secondary	9	15.4 (mean)	426	17
Pinger et al., 2018b	Interplay of formative assessment and instructional quality – interactive effects on students' mathematics achievement	Learning Environments Research	0	Department of Educational Quality and Evaluation, German Institute for International Educational Research	Germany	Europe	Secondary	9	15.1 (mean)	859	35
Polly et al., 2017	The Influence of an Internet-Based Formative Assessment Tool on Primary Grades Students' Number Sense Achievement	School Science and Mathematics	1	Reading and Elementary Education, Elementary Math, University of North Carolina at Charlotte	United States	North america	Primary	Kindergarten, 1, 2	No info	13567	906
Rakoczy et al., 2019	Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy	Learning and Instruction	0	Department of Educational Quality and Evaluation, German Institute for International Educational Research	Germany	Europe	Secondary	9	15.1 (mean)	620	26
Rodríguez-Martínez et al., 2023	Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions	British Journal of Educational Technology	0	Lab of Technology Integration in Classroom, University of Castilla-La Mancha	Spain	Europe	Primary	5	No info	94	No info

Descriptive data of included studies

Author, Year	Title	Journal		Affiliation of 1st author			Students				Teachers
		Name	M.E.[a]	Department & university	Country	Continent	Level	Grade[b]	Age	Sample size	Sample size
Safadi, 2018	Knowledge-Integration Processes and Learning Outcomes Associated with a Self-Diagnosis Activity: The Case of 5th-Graders Studying Simple Fractions	International Journal of Science and Mathematics Education	1	The Academic Arab College for Education	Israel	Asia	Primary	5	No info	46	1
Shih et al., 2023	Mathematics intelligent tutoring system for learning multiplication and division of fractions based on diagnostic teaching	Education and Information Technologies	0	Graduate Institute of Educational Information and Measurement, National Taichung University of Education	Taiwan (China)	Asia	Primary	6	No info	66	No info
Smit et al., 2023	Supporting primary students' mathematical reasoning practice: the effects of formative feedback and the mediating role of self-efficacy	Research in Mathematics Education	1	Institute for School and Profession, University of Teacher Education St. Gallen	Switzerland	Europe	Primary	4, 5, 6	No info	1261	71
Supovitz et al., 2021	Experimental Impacts of Learning Trajectory – Oriented Formative Assessment on Student Problem-Solving Accuracy and Strategy Sophistication	Journal for Research in Mathematics Education	1	Policy, Organizations, Leadership, and Systems Division, Research in Education, University of Pennsylvania	United States	North america	Primary	3 to 5	No info	1161	61 (schools)
Szeibert et al., 2023	A Case Study of Using Test-Enhanced Learning as a Formative Assessment in High School Mathematics	International Journal of Science and Mathematics Education	1	Faculty of Science, Eötvös Loránd University	Hungary	Europe	Secondary	9	No info	66	2
Van den Berg et al., 2016	Developing classroom formative assessment in dutch primary mathematics education	Educational Studies	0	GION Education/Research, University of Groningen	Netherlands	Europe	Primary	2, 3	No info	57	No info
Van den Berg et al., 2018	Testing the effectiveness of classroom formative assessment in Dutch primary mathematics education	School effectiveness and school improvement	0	GION Education/Research, University of Groningen	Netherlands	Europe	Primary	4, 5	No info	786	34
Vattøy & Gamlem, 2023	Students' experiences of peer feedback practices as related to awareness raising of learning goals, self-monitoring, self-efficacy, anxiety, and enjoyment in teaching EFL and mathematics	Scandinavian Journal of Educational Research	0	Faculty of Humanities and Teacher Education, Volda University College	Norway	Europe	Secondary	9	14-15 (range)	1695	No info
Vries et al., 2022	The impact of an assessment for learning teacher professionsl development program on students' metacognition	School effectiveness and school improvement	0	ELAN Teacher Development, Faculty of Behavioural, Management and Social Science, University of Twente	Netherlands	Europe	Secondary	7, 8, 9	No info	803	48
Wafubwa & Csíkos, 2022	Impact of Formative Assessment Instructional Approach on Students' Mathematics Achievement and their Metacognitive Awareness	International Journal of Instruction	0	Doctoral School of Education, Faculty of Humanities and Social Sciences, University of Szeged-Hungary	Hungary	Europe	Secondary	11	No info	164	4
Wongwatkit et al., 2017	Influence of an integrated learning diagnosis and formative assessment-based personalized web learning approach on students learning performances and perceptions	Interactive learning environments	0	Institute for Innovative Learning, Mahidol University	Thailand	Asia	Primary	6	11–12 (range)	63	No info
Yıldırım & Yıldırım, 2019	Predicting mathematics achievement: The role of perceived feedback, teacher support and self-beliefs	Turkish Journal of Education	0	Faculty of Education, Bolu Abant İzzet Baysal University	Turkey	Asia	Secondary	7 to 12	No info	4848	No info
Zheng et al., 2019	Using Instruction-Embedded Formative Assessment to Predict State Summative Test Scores and Achievement Levels in Mathematics	Journal of Learning Analytics	0	Department of Educational Psychology, University of Georgia	United States	North america	Secondary	6, 7, 8	No info	>23,000	0
Zulliger et al., 2022	Observed Quality of Formative Peer and Self-Assessment in Everyday Mathematics Teaching and its Effects on Student Performance	European Journal of Educational Research	0	Institute for school and heterogeneity, University of Teacher Education Lucerne	Switzerland	Europe	Primary	4	10.5 (mean)	634	52

Note.
[a] Mathematics Education Journal (1=yes, 0=no); [b] Preservice teachers (PST)

Detailed results of included studies

Table 8 Results of included studies investigated effects of formative teacher assessment (TA, component 1) on students' cognitive and noncognitive learning (n = 12)

Author, year	Aim	Design [a]	Intervention (by) [b]	PDP [c]	Participants [d] (age)	Math domain [e]	Indep. Variable [f]	Digital media [g]	Feedback [h]	Dependent variables with effect size [i]				Added measures [j]	Main findings
										Cognitive		*Non-cognitive*			
Beesley et al., 2018	Implementation of PDP and its impact on achievement	Long CG	1y (eduT)	9x/1y	6177 S, gr 6-8 47 T	General	TA	0	elab.	District wide math test	n.s. time x group	–		–	No impact of PDP on achievement
Guiaya & Bueno, 2019	Contribution of strategic questioning as FA on achievement	Cross CG	¼ y (extern)	–	125 S, gr 9 (13-19 y)	Combinatorics & calculus	TA	0	elab.	Variation: Calculus:	*** ***	–		–	Sign. increased achievement when using FA strategic questioning
Guo et al., 2019	Relationship between 5 types of T feedback on 6 self-regulated-learning (SRL) strategies in math at high- (1) and low- (2) achieving schools	Cross	–	–	1282 S (15-19 y, 16.5 y)	General	TA	0	elab.	–		DF on intrinsic motivation: TP on self-efficacy:	(1) β= -.12* (2) β= -.17* (1) β=.37*** (2) β=.18*	–	Only effects of T' direct feedback (DF) and T praise (TP) on SRL in both achievement levels, other effects inconsistent
Kyaruzi et al., 2019	Impact of perceived T' FA practices and feedback delivery on feedback use and performance	Cross	–	–	2767 S, gr 11 (16.5 y)	General	TA	0	elab.	Self-reported final math exam made by T	–	–		Med: Percieved feedback use	Effects of perceived FA on self-reported performance mediated by perceived feedback use
Pinger et al., 2018a	Effects of the quality of T' feedback (1) number or (2) specifity of comments, (3) feedback utilisation or (4) as performance evaluation instructed on achievement and interest	Long only EG Random	3w, 13 lessons, predefined (eduT)	2 x ½ d	426 S, gr 9 (15.4 y) 17 T	Geometry	TA	0	elab.	Pythagoras' theorem	(1) β= -.13** (2) β= -.10* (3) β= .40** (4) n.s.	Interest	(1) n.s. (2) n.s. (3) β=.18* (4) β=.21*	–	Positive effects of instruction embedded feedback (3) on achievement and interest
Pinger et al., 2018b	Interplay of FA and instructional quality and its effects on achievement	Long CG Random	3w, 13 lessons, predefined (eduT)	2 x ½ d	859 S, gr 9 (15.1 y) 35 T	Geometry	TA	0	elab.	Pythagoras' theorem	n.s.	–		Mod: Teaching quality, T-S-relationship (n.s.)	N.s. direct effects of FA on achievement, neg. influence of intervention if teaching quality is already high
Rakoczy et al., 2019	Effects of FA on interest and achievement directly and via S' perception of the usefulness of feedback and their self-efficacy	Long CG Random	3w, 13 lessons, predefined (eduT)	2 x ½d	620 S, gr 9 (15.1 y) 26 T	Geometry	TA	0	elab.	Pythagoras' theorem	n.s. (direct/total)	Self-efficacy: Interest:	β=.25* n.s. (direct) β=.52* (total)	Med: Percieved usefulness, self-efficacy (n.s.)	Feedback perceived as more useful in FA condition, greater self-efficacy, interest tended to increase; no difference in learning progress
Supovitz et al., 2021	Impact of FA on problem-solving accuracy and solution strategy sophistication in multiplication and division	Long CG Random	2y (eduT)	3d/y + further sessions for 2y	1161 S, gr 3-5	Arithmetic	TA	0	elab.	–		Accuracy in problem-solving: Strategy sophistication:	β=.04***-.11*** β=.13***-.26**	–	Sig. impact of FA on problem-solving accuracy and sophistication of their strategy
Szeibert et al., 2023	Effects of test-enhanced FA on achievement to narrow achievement gap between S with different socio-economic-status	Cross CG	11 lessons (eduT)	Regularly consults	66 S, gr 9 2 T	Geometry	TA	0	elab.	Final test score	**	–		–	FA impacts achievement positively compared to CG
Van den Berg et al., 2016	Pilot study to develop a classroom FA (CFA) condition to improve T' FA practice and consequently enhance S' achievement	Long CG	2-4w (eduT)	10w	57 S, gr 2-3 6 T	General	TA	1	elab.	Curriculum embedded test	n.s.	–		–	N.s., but higher, performance, in FA condition, researchers discuss that a correct implementation of CFA-model might result in the enhancement of math performance
Van den Berg et al., 2018	Impact of a CFA practice, involving daily TA, weekly computer-based CFA and adaped instruction by T on achievement compared with usual practice; additional analysis of effects only in EG influenced by grade	Long CG Random	21w/1y (eduT)	1y	786 S, gr 4-5 34 T	General	TA	1	elab.	Achievement tests	n.s.	–		Mod: Grade	Neither impact of CFA on achievement nor on interaction between CFA and pretest, but sig. pos. effect of the degree of implementation on achievement in gr 5 (not in gr 4)
Yıldırım & Yıldırım, 2019	Relations between S perception of TA and achievement, self-efficacy, self-concept, anxiety	Cross	–	–	4848 S, gr 7-12	General	TA	0	n/a	PISA 2012 math scores	n.s.	Self-efficacy: Self-concept: Anxiety:	n.s. β=.15** β= -.13**	–	FA influences math self-concept positively and reduces math anxiety

Note. Content only written if part of study.
Abb. time periods: y(ears=schoolyears), m(onths), w(eeks), d(ays), '(minutes); agents: T(eacher), Tu(tor), S(tudent)
[a] Longitudinal (long)/cross-sectional (cross); control group (CG); randomization (random)
[b] Duration of Intervention by Teacher (T), T with previous Professional Development Program (eduT), extern e.g., researcher
[c] Length of a Professional Development Program (PDP)
[d] Mean age or range in years; Preservice Teachers (PST)
[e] Domain *general* includes math achievement tests, higher-order-thinking-skills in achievement, non-domain specific problem-solving or reasoning, if not further described.
[f] Components of FA: teacher-assessment (TA), peer assessment (peer), self-assessment (self), Computer-based-assessment (CBA)
[g] no digital media used (0), digital media used (1), Intelligent-Tutor-System used (2)
[h] Elaborated feedback (elab.), right-wrong feedback (r.-w.), not reported (n/a)
[i] *p <0.05, **p <0.01, ***p <0.001, not significant (n.s.), not reported (n/a)
[j] Moderator (mod) or mediator (med) analysis

Detailed results of included studies

Table 9 Results of included studies investigated effects of formative peer or self-assessment (component 2 or 3) on students' cognitive and noncognitive learning (n = 6)

Author, year	Aim	Design [a]	Intervention (by) [b]	PDP [c]	Participants [d] (age)	Math domain [e]	Indep. Variable [f]	Digital media [g]	Feedback [h]	Dependent variables with effect size [i]				Added measures [j]	Main findings
										Cognitive		*Non-cognitive*			
Alqassab et al., 2018	Impact of structured peer-feedback training on feedback provision skills	Long	1 semester (extern)	–	43 PST (22.51 y)	Geometry	Peer	0	elab.	–		Task level: Process-level: Self-reg-level: Self-level: Beliefs:	n.s. n.s. ** n.s. η_p^2=.34***	Mod: Achiev. level	Achievement levels crucial for type of provided peer feedback but not for beliefs about peer feedback
Can, 2021	Developments of PST' math teaching performance according to PA	Long	12w, 8h/w (extern)	–	58 PST (20-22 y)	General	Peer	0	elab.	–		Preparation: Teach-process: Class-management, feedback, communication & evaluation:	η^2=.13* η^2=.23* n.s.	–	PST' lesson preparation and teaching process improved over time, at least 3x teaching practice required
Chang et al., 2021	Impact of peer assessment in online-feedback-platform *Zuvio* on achievement and collaboration tendency (1), critical thinking awarness (2), problem-solving tendency (3), meta-cognition tendency (4) compared to T' feedback	Long CG	3w, 2x50'/w (extern)	–	112 S (12)	Geometry	Peer	1	elab.	Pythagoras' theorem	η^2=.20**	Social-emotional & metacognitive components	(1) η^2=.15** (2) η^2=.13** (3) η^2=.10** (4) η^2=.19**	–	Sig. positive impact of peer assessment on *Zuvio* platform on cognitive and non-cognitive achievement compared to T feedback
Safadi, 2018	Impact of self-diagnosis (SD) on math achievement	Long CG Random	2.5w, 11 lessons	–	46 S, gr 5 1 T	Arithmetic	Self	0	elab.	Fractions	η^2=.20** time x group	–		Mod: SD-level (only EG)	SD enhanced achievement compared to traditional teaching, only benefit for low SD level
Vattøy & Gamlem, 2023	Relation of peer feedback practices to several non-cognitive components	Cross	–	–	1695 S, gr 9 (14-15 y)	General	Peer	0	n/a	–		AR learning goals: Self-monitoring: Self-efficacy: Lack of anxiety: Enjoyment:	β=.28*** β=.10** β= -.08* β= -.14*** β=.19***	–	Awareness raising (AR) of learning goals as the most important variable for S' experiences of peer feedback practices
Zulliger et al., 2022	Independent effects of peer- and self-assessment quality in everyday math instruction on achievement and learning process	Long	–	–	634 S, gr 4 (10.5 y) 52 T	Arithmetic	Peer Self	0	n/a	Division	n.s. n.s.	Strategy variability as metacognition	n.s. n.s.	Med (n.s.): Strategy variability Mod: Math grade	No general effect on performance but high-quality self-assessment as beneficial for lower-performing S moderated with math grade

Note. Content only written if part of study.
Abb. time periods: y(ears=schoolyears), m(onths), w(eeks), d(ays), '(minutes); agents: T(eacher), Tu(tor), S(tudent)
[a] Longitudinal (long)/cross-sectional (cross); control group (CG); randomization (random)
[b] Duration of Intervention by Teacher (T), T with previous Professional Development Program (eduT), extern e.g., researcher
[c] Length of a Professional Development Program (PDP)
[d] Mean age or range in years; Preservice Teachers (PST)
[e] Domain *general* includes math achievement tests, higher-order-thinking-skills in achievement, non-domain specific problem-solving or reasoning, if not further described.
[f] Components of FA: teacher-assessment (TA), peer assessment (peer), self-assessment (self), Computer-based-assessment (CBA)
[g] no digital media used (0), digital media used (1), Intelligent-Tutor-System used (2)
[h] Elaborated feedback (elab.), right-wrong feedback (r.-w.), not reported (n/a)
[i] *p <0.05, **p <0.01, ***p <0.001, not significant (n.s.), not reported (n/a)
[j] Moderator (mod) or mediator (med) analysis

Detailed results of included studies

Table 10 Results of included studies investigated effects of formative teacher and peer or self-assessment (component 1 and 2 or 3) on students' cognitive and noncognitive learning (n = 6)

Author, year	Aim	Design [a]	Intervention (by) [b]	PDP [c]	Participants [d] (age)	Math domain [e]	Indep. Variable [f]	Digital media [g]	Feedback [h]	Dependent variables with effect size [i]				Added measures [j]	Main findings
										Cognitive		*Non-cognitive*			
Barana et al., 2022	Accuracy of self-assessment compared to TA, moderated by different processes in problem-solving	Long	3m (eduT)	10h	182 S, gr 11 (16y); 10 Tu (21-26y)	General	TA Self	0	elab.	–		Accuracy in problem solving	r=.52***-.74***	Mod: Phases of problem-solving process	Correlation between self-assessment and TA, with a tendency of S to underestimate their performance.
Chan et al., 2023	Impact of collaborative problem-based learning and peer assessment implemented with interactive-online-whiteboards (IOWB) on achievement and different deep (1) or surface (2) learning approaches	Long CG	1 lesson (T)	–	87 S, gr 7	Geometry	TA Peer	1	elab.	Angels & lines	n.s.	Learning approaches	(1) r=.52 to .63** (2) n.s.	–	N.s. impact of FA with IOWB on achievement but positive correlation with deep learning approaches
Chen et al., 2020	Impact of IOWB-based FA-activities with *EasiCare* software combined with adapted instruction by T and impromptu feedback by peers on achievement	Long	1y (eduT)	n/a	178 S, gr 2 2 T	General	TA Peer	1	elab.	Mid-term math test	β=.22**	–		–	IOWB-FA-activies influence achievement positively
Kartono & Shora, 2020	Effects of Process Oriented Guided Inquiry Learning (POGIL) with peer feedback on math reasoning abilities	Cross CG Random	1m (T)	–	64 S, gr 7	General	TA Peer	0	elab.	Math reasoning	*	–	–	–	POGIL with peer feedback influences math reasoning ability positively compared to CG
Ng et al., 2020	Impact of collaborative problem-based learning and PA implemented with IOWB on achievement compared to traditional tutor programs	Long CG	10w, 1h/w	–	830 freshmen	Calculus	TA Peer	1	elab.	Calculus	n.s.	–		–	N.s. impact of peer assessment with IOWB on achievement
Patchan et al., 2022	Impact of received peer-feedback compared to T' on problem-solving abilities and revision quality	Long CG Random	4d/2w (T)	–	32 S, gr 6 (12.1y)	General	TA Peer	0	elab.	Problem solving abilities	n.s.	Revision quality	n.s.	–	Peer-feedback influences problem solving ability marginally but n.s., as well as revision quality

Note. Content only written if part of study.
Abb. time periods: y(ears=schoolyears), m(onths), w(eeks), d(ays), '(minutes); agents: T(eacher), Tu(tor), S(tudent)
[a] Longitudinal (long)/cross-sectional (cross); control group (CG); randomization (random)
[b] Duration of Intervention by Teacher (T), T with previous Professional Development Program (eduT), extern e.g., researcher
[c] Length of a Professional Development Program (PDP)
[d] Mean age or range in years; Preservice Teachers (PST)
[e] Domain *general* includes math achievement tests, higher-order-thinking-skills in achievement, non-domain specific problem-solving or reasoning, if not further described.
[f] Components of FA: teacher-assessment (TA), peer assessment (peer), self-assessment (self), Computer-based-assessment (CBA)
[g] no digital media used (0), digital media used (1), Intelligent-Tutor-System used (2)
[h] Elaborated feedback (elab.), right-wrong feedback (r.-w.), not reported (n/a)
[i] *p <0.05, **p <0.01, ***p <0.001, not significant (n.s.), not reported (n/a)
[j] Moderator (mod) or mediator (med) analysis

Detailed results of included studies

Table 11 Results of included studies investigated effects of formative teacher, peer and self-assessment (component 1, 2, 3) on students' cognitive and noncognitive learning (n = 13)

Author, year	Aim	Design [a]	Interven-tion (by) [b]	PDP [c]	Participants [d] (age)	Math domain [e]	Indep. Variable [f]	Digital media [g]	Feed-back [h]	Dependent variables with effect size [i]				Added measures [j]	Main findings
										Cognitive		*Non-cognitive*			
Albano et al., 2023	Peer assessment-workshops with the aim to promote self-assessment on math activity and their impact on accuracy of S' feedback compared with T'	Cross	12w, 6h/w (Extern)	–	38 freshmen	General: Math II course	TA Peer Self	0	elab.	–		Feedback ac-curacy in 4 criteria	Increase of r 0.1-0.2	–	Accuracy of peer- and self-assess-ment measured by correlation with TA increases in all criteria at a low to medium level (r=0.3-0.52)
Andersson & Palm, 2017	Effects of changes in FA classroom practice followed a PDP on achievement	Long CG Random	1y (eduT)	144h/1 term, 6h/w +72h prep.	729 S, gr 4 25 T	General	TA Peer Self	0	elab.	Achievement	d=.66*	–		Mod (n.s.): procedure/pro-cess required tasks	Classes in FA condition sig. out-performed classes in CG in post-test controlled for pretest
Ayanaw et al., 2022	Effects of FA on achievement, moderated by achievement levels	Long CG Random	1 semester (eduT)	12w	129 PST 3 T	Algebra	TA Peer Self	0	elab.	Achievement test (self-made)	η_p^2=.58***	–		Mod (n.s.): Achiev. levels	FA condition outperformed CG in posttest, n.s. difference in posttest scores between achiev. levels
Ayanaw et al., 2023a	Effect of improved implementation of FA on higher-order-thinking skills (HOTs), moderated by achievement levels	Long CG Random	1 semester (eduT)	12w	129 PST 3 T	Algebra	TA Peer Self	0	elab.	HOTs in achievement test	η^2=.36*	–		Mod (n.s.): Achiev. levels	FA condition outperformed CG in posttest, n.s. difference in posttest scores between achiev. levels
Ayanaw et al., 2023b	Effects of FA on self-regulated learning skills (SRL), moderated by achievement levels	Long only EG Random	1 semester (eduT)	12w	52 PST	Algebra	TA Peer Self	0	elab.	–		SRL	d=.88***	Mod: Achiev. levels	SRL posttest scores higher than in pretest, difference between achiev. levels
Boström & Palm, 2023	Effects of changes in FA classroom practice followed a PDP on achievement, moderated by achievement levels	Long CG Random	1y	96h/1 term + 120h prep.	566 S, gr 7 (13y) 29 T	General	TA Peer Self	0	elab.	Achievement test	n.s.			Mod (n.s.): Achiev. levels	N.s. effects of FA on achieve-ment, neither correlated with number of new FA activities, nor between achiev. levels
Ganajová et al., 2021	Effect of FA classroom techniques (FACT) on inquiry skill de-velopment, moderated by achieve-ment levels in overall assessment	Long CG	5m (eduT)	1x before + 3x during in-tervention	121 S, gr 11 (15-16y) 11 T	General	TA Peer Self	0	elab.	Inquiry skills in data analysis	**	–		Mod: Achiev. levels	FACT pos. influence inquiry skills compared to CG, sig. differ-ence in development of individual inquiry skills for medium-/low-performing S
Hatami, 2015	Impact of collaborative learning peer assessment and self-assessment on achievement and self-regulation (SRL)	Long CG	8x 90' (T)	–	75 S, gr 11	Geometry	TA Peer Self	0	n/a	Achievement test (self-made)	n.s. time x group	Motivation & SRL	*** time x group	–	FA influenced achievement score and their self-motivation in the posttest positively
Kültür & Kutlu, 2021	Effect of FA classroom practices on achievement and attitudes towards math	Long CG Random	42x 40' (ex)	–	51 S, gr 10 (15-16y)	Algebra	TA Peer Self	0	elab.	Functions	η^2=.11** time x group	Attitudes to math	η^2=.29*** time x group	–	FA pos. influences achievement and attitudes from pre- to posttest compared to CG
Moyo et al., 2022	Effects of FA classroom practices on higher-order-thinking skills (HOTs) in math	Long	8w (eduT)	2d	272 S, gr 4 9 T	General	TA Peer Self	0	elab.	HOTs in achievement test	d=3.0***	–		–	Large effect of FA on HOTs from pre- to posttest
Smit et al., 2023	Effects of formative feedback on the development of reasoning com-petence via self-efficacy beliefs	Long	10w (eduT)	10w, 1h/w	1261 S, gr 4-6 71 T	General	TA Peer Self	0	elab.	Math reasoning	n.s. (ind & class)	Self-efficacy	β=.19** (ind) β=.62* (class)	Med: Self-efficacy	FA influences self-efficacy on in-dividual and class level, and per-formance on class level mediated by self-efficacy
Vries et al., 2022	Effects of changes in FA classroom practice followed a PDP on meta-cognitive skills	Long CG Random	1y (eduT)	4x/1y	803 S, gr 7-9 48 T	General	TA Peer Self	0	elab.	–		Planning Predicting Evaluating	n.s. β=.23* β=.22*	–	PDP influenced FA activities in-fluence Ss predicting and evaluat-ing skills
Wafubwa & Csíkos, 2022	Impact of FA classroom practices on achievement and metacognitive awarness, influence of gender	Long CG Random	14x 40'/4w (eduT)	3d	164 S, gr 11 4 T	Algebra	TA Peer Self	0	elab.	Proportional reason-ing	η^2=.38*	Meta-cogni-tion	η^2=.47***	Mod (n.s.): Gender	FA influences achievement and their metacognitive awarness in-dependet from gender

Note. Content only written if part of study.
Abb. time periods: y(ears=schoolyears), m(onths), w(eeks), d(ays), '(minutes); agents: T(eacher), Tu(tor), S(tudent)
[a] Longitudinal (long)/cross-sectional (cross); control group (CG); randomization (random)
[b] Duration of Intervention by Teacher (T), T with previous Professional Development Program (eduT), extern e.g., researcher
[c] Length of a Professional Development Program (PDP)
[d] Mean age or range in years; Preservice Teachers (PST)
[e] Domain *general* includes math achievement tests, higher-order-thinking-skills in achievement, non-domain specific problem-solving or reasoning, if not further described.
[f] Components of FA: teacher-assessment (TA), peer assessment (peer), self-assessment (self), Computer-based-assessment (CBA)
[g] no digital media used (0), digital media used (1), Intelligent-Tutor-System used (2)
[h] Elaborated feedback (elab.), right-wrong feedback (r.-w.), not reported (n/a)
[i] *p <0.05, **p <0.01, ***p <0.001, not significant (n.s.), not reported (n/a)
[j] Moderator (mod) or mediator (med) analysis

Detailed results of included studies

Table 12 Results of included studies investigated effects of computer-based assessment (CBA, component 4) on students' cognitive and noncognitive learning (n = 8)

Author, year	Aim	Design [a]	Intervention (by) [b]	PDP [c]	Participants [d] (age)	Math domain [e]	Indep. Var. [f]	Digital media [g]	Feedback [h]	Dependent variables with effect size [i]				Added measures [j]	Main findings
										Cognitive		*Non-cognitive*			
Barana et al., 2021	Impacts of automatic interactive feedback (IF) in a digital learning environment on achievement, influence of different socio-economic-states (SES)	Long	n/a (T)	–	399 S, gr 8	General	CBA	2	elab.	Grade in FA tool	***	–		Mod: SES	IF is more effective than other kinds of activities to engage learners in actions aimed at improving their results, stronger effects in low SES contexts
Bottge et al., 2021	Impact of technology-assisted prompts combined with instructional program (EG1, CBA to S+T) on achievement versus only instructional program (EG2, TA) (1); influence of additional time with FA against business-as-usual-teaching (2)	Long 2 EG Random	7 lessons, predefined (eduT)	2h	110 S with disabilities, gr 6-12 19 T	Arithmetic	CBA TA	2	elab.	Computation (a) & fractions (b, c)	(1) n.s. (2a) β=.29* (2b) β=.37* (2c) β=.70*	–		–	N.s. difference between EG1 and EG2; additional time with instructional program increases achievement sig. compared to add. time with business-as-usual-teaching
Faber et al., 2017	Impact of a digital FA tool *Snappet* perceiving simple feedback and adapted assignments on achievement and motivation, influence of pretest scores; additional analysis about extent of use (1) and observation score (2) only in EG	Long CG Random	5m (eduT)	½ d + optional sessions	1808 S, gr 3 (8-9y)	General	CBA	2	r.-w.	Stand. math test	β=.21** (1) β=.16** (2) β=.06*	Motivation	β=.16* (1) β=.19** (2) n.s.	Mod: Achiev. in pretest	Positive effects of *Snappet* on achievement and motivation, intensity of use supports these effects, achievement effects were higher for in pretest high-performing S
Polly et al., 2017	Impact of internet-based FA tool *AMC Anywhere* with resulting adapted instruction by T with PDP measured by frequency of use per S (1) or frequency of use per class mean (2) on achievement	Long CG	1y with at least 3x FA	80h/1y	13567 S, Kindergarten & gr 1-2 906 T	Arithmetic	CBA TA	2	elab.	Number sense with representation (a) or only mentaly (b)	(1a) n.s. (1b) * (2a) ** (2b) n.s.	–		Mod (n.s.): PDP	More frequently use of *AMC Anywhere* is pos. related to achievement, effect of FA on achievement is not increased by teachers' PDP
Rodríguez-Martínez et al., 2023	Impact of personalized homework based on digital learning-analytics (error analysis) compared to non-personalized homework	Long CG Random	4x 50' (T)	–	94 S, gr 5	Arithmetic	CBA	2	elab.	Fractions	d=.35*	–		–	Personalized homework based on errors relates to better understanding of the concept of fractions
Shih et al., 2023	Impact of adaptive dialogue-based instruction by intelligent-tutor-system (ITS) based on error analysis	Long CG Random	2h	–	66 S, gr 6	Arithmetic	CBA	2	elab.	Fractions & word probl.	η^2=.74*	–		Mod: Achiev. level	Using math ITS outperformes understanding of math problems compared to traditional classroom instruction, sig. effects for low-performing S
Wongwatkit et al., 2017	Impact of an integrated learning diagnosis and FA-based personalized web learning system on achievement, influence of S' learning perceptions	Long CG	2.5h	–	63 S, gr 6 (11-12y)	Geometry	CBA	2	elab.	Circle area	η^2=.097*	–		Mod: Perceived usefulness, assistence to learn, learning style (n.s.)	Learning with CBA revealed sig. better learning achievement than conventional web-based system, sig. relationship of S' learning perceptions with achievement
Zheng et al., 2019	Impact of user time (1), solved problems (2) and number of workspaces (3) in *MATHia* intelligent-tutor-system on achievement, influence of previous knowledge (KC)	Long	1y	–	23'000+ S, gr 6-8	General	CBA	2	n/a	District-wide stand. math test	(1) β= -.25*** (2) β=.30*** (3) β=.25***	–		Mod: Achiev. in pretest, no. of mastered KC	Number of solved problems and workspaces in *MATHia* increases achievement, not only more time; high achieving in pretest promotes achievement, but high no. of knowledge components mastered influences achievement negatively

Note. Content only written if part of study.
Abb. time periods: y(ears=schoolyears), m(onths), w(eeks), d(ays), '(minutes); agents: T(eacher), Tu(tor), S(tudent)
[a] Longitudinal (long)/cross-sectional (cross); control group (CG); randomization (random)
[b] Duration of Intervention by Teacher (T), T with previous Professional Development Program (eduT), extern e.g., researcher
[c] Length of a Professional Development Program (PDP)
[d] Mean age or range in years; Preservice Teachers (PST)
[e] Domain *general* includes math achievement tests, higher-order-thinking-skills in achievement, non-domain specific problem-solving or reasoning, if not further described.
[f] Components of FA: teacher-assessment (TA), peer assessment (peer), self-assessment (self), Computer-based-assessment (CBA)
[g] no digital media used (0), digital media used (1), Intelligent-Tutor-System used (2)
[h] Elaborated feedback (elab.), right-wrong feedback (r.-w.), not reported (n/a)
[i] *p <0.05, **p <0.01, ***p <0.001, not significant (n.s.), not reported (n/a)
[j] Moderator (mod) or mediator (med) analysis

Summarized results of included studies

Table 13 Compact results of included studies (*n*=45)

Author, year	Specific math context	Control group [a]	Randomiza-tion	Educational level [b]	Sample (range) [c]	Lenght [d]	PDP [e]	Digital media [f]	Elaborated feedback	Effects on cognitive learning			Effects on noncognitive learning		
										Positive significant	*Negativ significant*	*Not significant*	*Positive significant*	*Negativ significant*	*Not significant*
Albano et al., 2023	0	—	—	U	1	3	—	0	1	—	—	—	1	0	0
Alqassab et al., 2018	1	—	—	U	1	3	—	0	1	—	—	—	1	0	1
Andersson & Palm, 2017	0	1	1	P	3	3	3	0	1	1	0	0	—	—	—
Ayanaw et al., 2022	1	1	1	U	2	3	3	0	1	1	0	0	—	—	—
Ayanaw et al., 2023a	1	1	1	U	2	3	3	0	1	1	0	0	—	—	—
Ayanaw et al., 2023b	1	—	1	U	1	3	3	0	1	—	—	—	1	0	0
Barana et al., 2021	0	—	—	S	2	?	—	2	1	1	0	0	—	—	—
Barana et al., 2022	0	—	—	S	2	3	2	0	1	—	—	—	1	0	0
Beesley et al., 2018	0	1	0	S	5	3	3	0	1	0	0	1	—	—	—
Boström & Palm, 2023	0	1	1	S	3	3	3	0	1	0	0	1	—	—	—
Bottge et al., 2021	1	1*	1	S	2	2	1	2	1	1	0	1	—	—	—
Can, 2021	0	—	—	U	1	3	—	0	1	—	—	—	1	0	1
Chan et al., 2023	1	1	0	S	1	1	—	1	1	0	0	1	1	0	1
Chang et al., 2021	1	1	0	S	2	2	—	1	1	1	0	0	1	0	0
Chen et al., 2020	0	—	—	P	2	3	n/a	1	1	1	0	0	—	—	—
Faber et al., 2017	0	1	1	P	4	3	1	2	0	1	0	0	1	0	1
Ganajová et al., 2021	0	1	0	S	2	3	2	0	1	1	0	0	—	—	—
Guiaya & Bueno, 2019	1	1	0	S	2	3	—	0	1	1	0	0	—	—	—
Guo et al., 2019	0	—	—	S	4	—	—	0	1	—	—	—	1	1	0
Hatami, 2015	1	1	0	S	1	2	—	0	n/a	0	0	1	1	0	0
Kartono & Shora, 2020	0	1	1	S	1	2	—	0	1	1	0	0	—	—	—
Kültür & Kutlu, 2021	1	1	1	S	1	3	—	0	1	1	0	0	1	0	0
Kyaruzi et al., 2019	0	—	—	S	4	—	—	0	1	—	—	—	—	—	—
Moyo et al., 2022	0	—	—	P	2	3	2	0	1	1	0	0	—	—	—
Ng et al., 2020	1	1	0	U	3	2	—	1	1	0	0	1	—	—	—
Patchan et al., 2022	0	1	1	P	1	2	—	0	1	0	0	1	0	0	1
Pinger et al., 2018a	1	—	1	S	2	2	1	0	1	1	1	1	1	0	1
Pinger et al., 2018b	1	1	1	S	3	2	1	0	1	0	0	1	—	—	—
Polly et al., 2017	1	1	0	P	5	3	3	2	1	1	0	1	—	—	—
Rakoczy et al., 2019	1	1	1	S	3	2	1	0	1	0	0	1	1	0	1
Rodríguez-Martínez et al., 2023	1	1	1	P	1	2	—	2	1	1	0	0	—	—	—
Safadi, 2018	1	1	1	P	1	2	—	0	1	1	0	0	—	—	—
Shih et al., 2023	1	1	1	P	1	1	—	2	1	1	0	0	—	—	—
Smit et al., 2023	0	—	—	P	4	3	3	0	1	0	0	1	1	0	0
Supovitz et al., 2021	1	1	1	P	4	3	2	0	1	—	—	—	1	0	0
Szeibert et al., 2023	1	1	0	S	1	2	1	0	1	1	0	0	—	—	—
Van den Berg et al., 2016	0	1	0	P	1	2	3	1	1	0	0	1	—	—	—
Van den Berg et al., 2018	0	1	1	P	3	3	3	1	1	0	0	1	—	—	—
Vattøy & Gamlem, 2023	0	—	—	S	4	—	—	0	n/a	—	—	—	1	1	0
Vries et al., 2022	0	1	1	S	3	3	2	0	1	—	—	—	1	0	1
Wafubwa & Csíkos, 2022	1	1	1	S	2	2	2	0	1	1	0	0	1	0	0
Wongwatkit et al., 2017	1	1	0	P	1	1	—	2	1	1	0	0	—	—	—
Yıldırım & Yıldırım, 2019	0	—	—	S	5	—	—	0	n/a	0	0	1	—	—	—
Zheng et al., 2019	0	—	—	S	5	3	—	2	n/a	1	1	0	—	—	—
Zulliger et al., 2022	1	—	—	P	3	—	—	0	n/a	0	0	1	0	0	1

Note. If not defined otherwise, 1=yes, 0=no, — not included, n/a=no information in the article

[a] *Two experimental groups compared

[b] P(rimary school), S(econdary school), U(niversity)

[c] 1 = 20-100 students, 2 = 100-500 students, 3 = 500-1000 students, 4 = 1000-3000 students, 5 = >3000 students

[d] 1 = <1 day, 2 = >1 day, ≤16 lessons, 3 = >3 months with regular FA use

[e] 1 = ≤1 day, 2 = >1 day, <5 sessions, 3 = >1 month, ≥5 sessions

[f] 0 = no digital media, 1 = digital media exept interactive learning environments, 2 = Interactive learning environments