

*Overview of (quasi-)experimental studies included in the review*

No	FA (Year)	RE or QE	SD	# Stud.	Student tasks in intervention	Length of intervention	Type of dependent variable	Type of control group	OC	Type of analysis
Studies										
1A	Hanzawa (2018)	QE	Mixed	50	Note-taking, argumentative writing, debating, public speaking to applying strategic academic language skills related to voice onset time	5-6 hours academic content course + topic + 4-5 hours EAP	Phonetics: Voice onset time English -longer vot at later timepoints	Non-CBI group with instruction in L1	Positive	Repeated measures ANOVA (3 timepoints)
2A	Huang (2016)	QE		238	Animation design task, provided with English material only. Step-by-step instruction with English materials (L2)	1 semester 4 hours weekly	Assigned animation effects exam score	Previous cohort – CBI without instructional streaming video	Positive	t-test, comparing earlier cohort exam scores
3A	Lamri (2018)	QE	Social Sciences	37	Blended CLIL-course; focus on building legal terminology corpus and training reading strategies. Pre-reading tasks and post-reading tasks.	10 hours in the classroom and a minimum of 10 other hours outside the classroom	Reading ability	Business as usual – not blended	Positive	No inferential statistics, only descriptives
4A	Satilmis (2015)	QE	STEM	48	A vocabulary training component by content teacher	4 week program	Content vocabulary	Business as usual (instruction by language teacher)	Positive	Correlation
5A	Vega (2019)	RE	Tourism	15	Written assignments and MC questionnaires based on case studies that emphasized reading skills, group critical thinking exercises and oral presentations.	six hours of class	CEPT – language skills	Non-CLIL – grammar and vocabulary exercises, writing assignment, oral presentations	NS	Z-score mean proficiency
6A	Walenta (2018)	QE	STEM	20	Content-based grammar instruction incorporated into regular class meetings and compromised CBSI activities: right or wrong answer content wise [?]; affective tasks (personal responses in relation to content); discourse level and sentence level tasks. All tasks completed in pairs.	2 x 20 minute sessions incorporated into regular class meetings	Written production of target forms	CLIL without explicit attention to content-based grammar	Positive	Anova (time*instruction interaction), 3 timepoints (pretest, posttest 1 and posttest 2)
7B	Dansereau (2020)	QE	STEM	317	Rewriting of reports; at least one writing tutoring session before resubmitting lab report		Lab reports performance	Voluntary: Without addition of tutoring	Positive for one of the three tasks	t-test (comparison between with and without writing tutoring)

8B	Moon (2018)	QE	STEM	168	Electronically facilitated peer review to enhance self confidence in explaining theoretical concepts in writing	2 weeks	Writing task performance	Homework with traditional physical chemistry problems	Positive for one of the three tasks	Linear regression
9B	Chen (2020)	QE	STEM	184	Academic writing assignments in a four-step program, online simulations with academic writing, prompts to write and develop understanding of chemistry concepts and feedback using content-related examples based on student writing samples	1 course	Conceptual knowledge	Traditional lecture-based approach	Positive	Ancova with pre-post and two conditions
10B	Elder (2016)	RE	STEM	36	Additional writing assignment on a topic students already completed a problem solving assignment for	1 course	Quality of writing assignment	No writing assignment during class	NS	Correlation between exam score and writing assignment
11B	Batova (2019)	QE		522	Structured writing assignment to explain a specific concept to a specific audience		Exam outcomes	Self-enrollment in (1) no writing assignment (2) alternative writing assignment (3) Intervention: structured writing assignment	Positive	Regression with propensity score matching
12B	Kuzmenkova (2022) (1)	RE	FLEBM	192	Reading of authentic research articles, spotting recurrent patterns and studying language in context. Activities based on writing, redrafting and editing.	6 months	Assessment of communicative task realisation	Business-as-usual: extensive reading and analyzing articles and imitating authors' style while writing a proposal	Positive	t-test (difference treatment and control, not controlling for pretest)
	Kuzmenkova (2)	RE	idem	idem		idem	Aspects of academic writing (organization and format)		Positive	t-test (difference treatment and control, not controlling for pretest)
	Kuzmenkova (3)	RE	idem	idem		idem	Language range and control (style, grammar etc)		Positive	t-test (difference treatment and control, not controlling for pretest)

13B	Sherman (2020)	QE	Mixed	82	Realistic problem-solving scenarios to enhance design thinking and oral communication	1 course	Presentation skills	Self-enrollment for different section of 'Design Thinking in Technology'-> control group consisted of students not participating in sections with co-teaching	Positive	t-test: posttest only for group presentations; no individual assessment
14C	Tsai (2019)	QE	FLEBM	33	Writing letters, using dictation texts, teacher and peer feedback	12 weeks, 100 minutes each week	Writing performance	Previous cohort – No ICT addition; English proficiency in both groups comparable at the start	NS	Paired samples t-test
15C	Xu (2018)	RE	FLEBM	200	Reading texts, answering comprehension questions, writing journal entry, presenting oral reports	14 sessions	Vocabulary performance	Two groups (random assignment: (1) reading comprehension plus general vocabulary exercises – control group (2) reading comprehension plus vocabulary enhancement activities with emphasis on content-related concepts– intervention group	Positive	No inferential statistics, only descriptives

Note: Studies 1A-6A = CBA/CBI/CLIL, studies 7B-13B = WAC/WtL, studies 14C-15C = other. FA = first author; QE = quasi-experiment, RE = random experiment; SD = domain of study; FLEBM = finance, law, economics, business, management; STEM = science, technology, engineering, mathematics; OC = Outcome; NS = non-significant