Received: 29 December 2021 | Accepted: 5 June 2022

DOI: 10.1002/rev3.3359

STATE OF THE ART REVIEW **Review of Education** BERA

Which types of instruction in writing-to-learn lead to insight and topic knowledge in different disciplines? A review of empirical studies

Aartje van Dijk[1] | **Amos van Gelderen**[1] | **Folkert Kuiken**[2]

[1]Research Center Urban Talent, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

[2]Department for Humanities, University of Amsterdam, Amsterdam, The Netherlands

Correspondence
Aartje van Dijk, Research Center Urban Talent, Rotterdam University of Applied Sciences, Museumpark 40, 3015 CX Rotterdam, The Netherlands.
Email: a.c.van.dijk@hr.nl

Funding information
The funding for the research was provided by Rotterdam University of Applied sciences.

Abstract

This review examines which types of instruction in writing-to-learn lead to effects on insight and topic knowledge in different disciplines, in grades 5–12 and in higher education. Forty-three empirical studies have been selected to answer this question. Four types of instruction are distinguished. Three of them are based on hypotheses proposed by Klein (1999) about the cognitive processes involved in writing-to-learn: Forward Search, Genre Writing and Backward Search. The fourth type, Planning Only, arises from the literature reviewed. Results of the studies show that about two thirds of the (quasi) experimental studies lead to positive effects on insight and topic knowledge for the four types of instruction. However, given the small number of experimental studies conducted, no firm conclusions can be drawn for three types of instruction. For the remaining type of instruction, Genre Writing, a larger number of studies provide positive evidence. Suggestions for future research are discussed.

KEYWORDS
grades 5 to 12, higher education, instruction in writing-to-learn, learning in the disciplines

No sections of this review have been previously published.

Context and implications

Rationale for this study

It is known that instruction is needed for promoting learning by writing. The question is which specific types of instruction can be used for inciting the writing-to-learn process.

Why the new findings matter

The finding that the four types of instruction, especially Genre Writing, may promote learning, can serve as a reference for judging or formulating instruction.

Implications for educational researchers and policy makers

The outcome may be interesting for education, because it entails that one from four types of instruction, that is, Genre Writing, shows the largest positive effects. Subject teachers often are hesitant to use writing-to-learn in their teaching practice, therefore the findings may provide clearness about how to instruct it. They can use Genre Writing instruction for composing writing tasks and select the preferred content of this instruction. The authors indicate mechanisms underlying the process of writing-to-learn that may be stimulated by instruction. For researchers it can be interesting to study these mechanisms, for instance in think-aloud studies.

INTRODUCTION

Language is an important mediator of teaching and learning in all disciplines. Teachers explain subject matter by talking, and students study subject matter by listening, discussing and reading. Students' writing is normally used for *assessing* insight and topic knowledge, but much less for *learning* (Linnemann & Stephany, 2014). Writing assignments are normally not used by teachers as a tool for developing insight into subject matter or for acquiring topic knowledge. Baker et al. (2008) described problems teachers foresee for including writing-to-learn in their teaching, such as the extra time that is required and their insecurity in evaluating students' texts (see also Bean, 2011; Wallace et al., 2007). On the other hand, cognitive activities involved in writing might be an important additional option for arriving at deeper insight and new topic knowledge in the disciplines.

Acquiring insight and topic knowledge as the main goals of writing is what we call *writing-to-learn*, which is the topic of this study. Boscolo and Carotti (2003) describe writing-to-learn as follows: 'all the writing activities aimed at facilitating and/or strengthening recall, understanding or elaboration of concepts and ideas' (p. 200). Examples are writing a learning-journal to reflect on subject matter, writing to explain a subject adapted to the needs of readers that have no prior knowledge, or writing an argument based on several texts.

Emig (1977) stated that next to reading, listening and speaking, writing can be an important tool for learning as well. She considers writing a way of encouraging (critical) thinking and understanding, because it requires the writer to discern conceptual relationships, and to display these by means of syntactical, lexical and rhetorical devices. According to Emig (1977), writing provides unique opportunities for learning, because the slow pace of

the writing process gives extra room for reflection. Furthermore, the result of the writing process remains visible, enabling feedback on text contents and thereby providing new insights.

In research on writing-to-learn, learning concerns higher level thinking processes, which are described as 'integrating new information and prior knowledge' (Rivard & Straw, 2000, p. 567) or as knowledge transforming (Wallace et al., 2007, p. 27, p. 31). The latter is derived from Bereiter and Scardamalia's (1987) writing model that influenced theory on writing-to-learn to date (Klein & Boscolo, 2016). Applebee (1984) states that writing tasks aiming at writing-to-learn should be focused on 'heuristic activity in which subject knowledge is examined and extended' (p. 589).

Apparently, writing-to-learn is considered to be a comprehensive process, directed at insight into the meaning of newly learned knowledge and recall of topic knowledge. In the present review, we take the view that both types of learning by writing are relevant. By *topic knowledge* we mean recall of learned concepts (retrieving concepts from memory). By *insight* we mean the ability to relate new concepts to prior (topic) knowledge. We view *insight* as the ultimate aim of writing-to-learn, because insight is the manifestation of higher order understanding of concepts.

The present review aims at distinguishing types of instruction in writing that may lead to new insight and topic knowledge. The next two sections respectively discuss early studies into the character of writing-to-learn, and more recent theories about it. Following on from that, the types of instruction are described.

THEORETICAL OVERVIEW

Early studies into writing-to-learn

Since the 1970s, researchers were intrigued by the idea that writing evokes a way of thinking and learning, which can be applied in education. The prevailing view was that any writing task leads to learning, which is called the *strong text theory*. Writing was seen as 'a mode of learning' as Emig (1977) called it in the title of her essay. According to Klein and Boscolo (2016), the focus was on two types of writing: analytical writing (arguments, essays) and personal writing (e.g., diaries, poetry, stories and learning journals). Both types were considered relevant for learning in all disciplines and therefore were summarised in the phrase 'Writing Across the Curriculum'.

Applebee (1984) noticed that the relation between writing and learning was an unexamined assumption. He argued that writing-to-learn might be a more complex process than was thought, and that the specific design of writing assignments might determine what type of learning occurs. He observed that the effect of writing might be retention, for instance when students write answers to questions about newly learned topics, or a new insight, for instance when students write an argument to explain their point of view. Consequently, Applebee (1984) proposed experimental research into the interaction between writing tasks and the aims of teachers. Langer and Applebee (1987) elaborated on this interaction and examined which pedagogical conditions contribute to learning. They concluded that treating content in various ways when writing, for instance when students are required to revise their writing, might result in productive learning about that content in terms of topic knowledge as well as insight.

Early studies into effects of writing-to-learn are often based on the strong text theory (Ackerman, 1993; Rivard, 1994). These studies that were conducted in diverse grades and disciplines (Ackerman, 1993), were seldom situated in regular classrooms, and in most cases directed at science (Rivard, 1994). At the most, some studies (e.g., Boyles et al., 1994; Langer & Applebee, 1987) showed a complex picture of positive and negative results. This

caused Ackerman (1993) as well as Rivard (1994) to conclude that the strong text theory was not supported by the outcomes of research.

In Bangert-Drowns et al.'s (2004) meta-analysis on effects of writing-to-learn, one of the research questions is whether particular types of writing assignments, for instance personal expressive writing, result in learning, thus elaborating on Applebee's view. The authors did not find an effect on learning and concluded that the strong text theory does not hold, in accordance with Ackerman (1993) and Rivard (1994). In addition, Bangert-Drowns et al. (2004) assembled studies that were performed in school settings. They examined 48 studies directed at the contexts of primary education, secondary education and university, 34 of which were conducted in regular classrooms, and in different disciplines. They found some support for their hypothesis that writing assignments including prompts on metacognition, such as reflection on students' own learning processes, have an effect on learning, and might stimulate the process of writing-to-learn. The authors suggested that effects might be larger when instruction also comprises training of cognitive writing processes before writing, for instance goal setting and organising.

Theories about the process of writing-to-learn

Three theories about the process of writing-to-learn have been proposed: the knowledge-transforming model by Bereiter and Scardamalia (1987), the dual process model by Galbraith (1992, 2009) and four hypotheses concerning the cognitive processes of writing-to-learn by Klein (1999).

According to the knowledge-transforming model (Bereiter & Scardamalia, 1987), interaction takes place between a content space containing writers' topic knowledge, and a rhetorical space containing their discourse knowledge. In this interaction, writers set rhetorical goals, generate content, revise their rhetorical goals, and repeat these actions, until they consider their text satisfactory. For instance, when writing a text for an audience, a writer describes a theory (content) in abstract terms. When rereading the text, the writer realises that the audience may not understand it, because extra knowledge about the topic is needed (rhetorical goal). Therefore, the writer decides to give an example (content) for making the text more accessible to the audience. As a result of these interactions between content space and rhetorical space, writers may acquire new insights.

Galbraith (1992) distinguished two composing styles. He regards writing including planning activities as a composing style of high self-monitors opposed to the approach of writing without planning, which he considers a style of low self-monitors. High self-monitors pay much attention to rhetorical aspects when planning and reviewing, whereas low self-monitors write spontaneously and only attend to rhetorical aspects after reviewing the content of their texts.

Galbraith (2009) tested his ideas empirically, arriving at the dual process model. This model distinguishes a knowledge retrieval system and a knowledge constituting system in writing. Writers use their knowledge retrieval system, in which explicit knowledge is stored, for retrieving and planning content. When writing, they use their knowledge constituting system in order to make new connections between concepts, of which the writer was not aware previously (implicit knowledge). According to Galbraith (1992, 2009), a characteristic of high self-monitors is that they are inclined to hold to their planning while writing, whereas low self-monitors are inclined to deviate from their planning whenever they want. In the latter situation new insights may arise, whereas high self-monitors may produce well-structured texts, but no new insights.

Klein (1999) derives four hypotheses from the literature for how cognitive and metacognitive processes may contribute to writing-to-learn. The first, known as '*writing at the point of*

utterance' (based on Britton, 1982), states that a text resulting from spontaneous writing reveals the writer's knowledge. The writing-to-learn process is supposed to take place during formulating. This hypothesis corresponds to the strong text theory that was rejected earlier by Ackerman (1993) and Rivard (1994) (see previous section). Therefore, from here on we refer to Klein's other three hypotheses.

The second hypothesis, *Forward Search,* is based on Galbraith (1992), among others, and assumes that writers write down their ideas in a first draft without any preparation and keep on writing until they have written all they can think of. Then, they reread their drafts, draw new inferences or recognise a flaw in the logic, and revise them taking rhetorical goals into account. The assumption is that the process of writing-to-learn takes place in reviewing and revising.

The focus of the third hypothesis, *Genre Writing*, is on the genre of the text to be written. A genre can be characterised by its rhetorical goals and the relations between text elements directed at attaining these goals (Halliday & Martin, 1993). For instance, the rhetorical goal of the genre *argument* is to convince the audience, and the relations between the text elements (opinion, arguments and conclusion) are argumentative. Klein (1999) based the *Genre Writing* hypothesis on the view that writers composing a text of a specific genre have to use their knowledge of that genre. Depending on the genre, writers formulate their rhetorical goals, and the relations between the elements belonging to the genre. By using their genre knowledge, they may recognise relations between concepts they were not aware of before (Newell, 1984). Thus, by paying attention to (reflecting on) the rhetorical requirements of the genre, writers may acquire new insight into conceptual relations between elements of their topic knowledge (Langer & Applebee, 1987).

Finally, the fourth hypothesis, *Backward Search*, implies that writers set rhetorical goals, generate content based on the rhetorical goals (planning), subsequently write their text, and finally revise their text referring to their rhetorical goals and planned content. Just as for the Forward Search hypothesis, the process of writing-to-learn is assumed to take place in reviewing and revising the content. The *Backward Search* hypothesis, however, assumes that writers specifically reflect on goals (planning) *set before writing*. While performing these activities, writers may discover relations between concepts they were not aware of yet, leading to new insights. This—recursory—process is derived from the knowledge-transforming model of Bereiter and Scardamalia (1987).

Types of instruction in writing-to-learn

Since Bangert-Drowns et al. (2004) considered instruction in cognitive and metacognitive writing processes necessary for eliciting the process of writing-to-learn, the following generation of studies explored various types of instruction. Recently, these were reviewed by Miller et al. (2018), Graham et al. (2020) and Hand et al. (2021).

The systematic review of 43 studies by Miller et al. (2018) aimed at exploring the state of research on writing-to-learn in grades 6 to 12. The reviewed studies took place in the disciplines humanities, social studies, science and mathematics. The researchers performed an inductive analysis revealing that instructing writing-to-learn by means of a checklist for organising and generating activities (Science Writing Heuristic) or by journalling led to positive effects. In addition, they found positive effects of inquiry-based instruction of cognitive and metacognitive writing processes. The positive results were found in 46% of the reviewed studies.

In their meta-analysis, Graham et al. (2020) found that 82% of 56 reviewed studies in grades 1 to 12 in science, social studies and mathematics led to positive effects on learning with an average effect size of 0.30. However, the researchers found a large variability in

effect sizes ranging from 1.67 to –0.74, which they could not explain by means of a moderation analysis of any of a large number of variables, such as type and features of writing activities and instruction of cognitive and metacognitive writing processes. Therefore, it was not clear which components of instruction led to the found positive effects.

Hand et al. (2021) reviewed 81 (master and doctoral) theses on the application of instruction using the Science Writing Heuristic. The researchers found that its use resulted in growth of insight and topic knowledge regardless of grade or cultural background. In particular, the reviewers investigated students' knowledge-generating activities for identifying patterns related to the outcomes. Their qualitative analysis showed that the duration of the intervention was a determining factor in arriving at positive effects for students as well as for teachers. Students needed time for mastering knowledge generating and teachers needed time for exploring how to coach students at it. Other influential factors were the teacher's critical questions stimulating students' thinking and the combination of individual, group and class activities for performing generating activities.

Thus, the reviews appear to provide evidence for positive effects of instruction in studies on writing-to-learn (Graham et al., 2020), of specific ways of instructing in writing-to-learn (Hand et al., 2021; Miller et al., 2018), and of conditions contributing to these outcomes (Hand et al., 2021). The three reviews do not provide insight into the relations between the architecture of instruction and underlying writing-to-learn processes. Graham et al. (2020) investigated various modes of instruction, but were not able to explain the found effects. Miller et al. (2018) as well as Hand et al. (2021) show positive effects of instruction by means of the Science Writing Heuristic, but do not explain their outcomes. The three reviews used a data driven approach.

In the present study, we attempt a more theory-guided approach by systematically discriminating types of instruction on the basis of theoretical assumptions about the (meta) cognitive processes that they invoke. Klein's (1999) three hypotheses served as the basis for describing types of instruction. Klein (1999) composed his hypotheses about the process of writing-to-learn by means of (combinations of) the (meta)cognitive processes of planning and reviewing and of the use of genre knowledge during writing. By analysing examples of concrete instruction, it is possible to discern on which of these processes the approach is based. Instruction directed at spontaneous writing followed by inspection, feedback and/or revision can be classified as instruction based on the Forward Search hypothesis. In contrast, instruction that emphasises planning activities before writing, followed by prompts to revise the drafts to ensure a better coverage of the subject or better reception by the readership (instigated by feedback by peers or teachers), can be classified as instruction based on the Backward Search hypothesis. Finally, instruction emphasising that students are familiarised with genre characteristics of the texts they are supposed to write and that they use these characteristics while writing, can be classified as based on the Genre Writing hypothesis.

This review approach allows investigating the relative effectiveness of each of these three types of instruction. In addition, it allows us to see to what degree the three hypotheses of Klein (1999) cover the types of instruction that have been studied until now. In case that instructional approaches are used that clearly deviate from the three types, this is valuable information that may contribute to a better understanding of underlying assumptions in instructional types directed at writing-to-learn.

The present study

Additional to the outcomes of the previous meta-studies, the present study is aimed at clarifying which cognitive (organising, generating, goal setting, using genre knowledge) and metacognitive writing processes (reviewing, revising) are part of effective instruction in

writing-to-learn. Therefore, we conducted an empirical review. Starting from Klein's (1999) hypotheses, we analysed studies using instruction based on one of the three types of instruction discerned and their effects on insight and topic knowledge. In addition, we analysed studies that deviated from the three types and compared them to the other types. Presumably this provides us with better understanding of the cognitive and metacognitive writing processes involved in writing-to-learn and in the types of instruction that can set them off.

Our research question is as follows:

Which types of instruction in writing-to-learn directed at cognitive and metacognitive writing processes result in new insights and topic knowledge in several disciplines?

Because instruction in writing-to-learn is meant to be effective in the context of education, we limited our review to studies taking place in regular classrooms. We were interested in how instruction on writing-to-learn can be embedded in educational practice while improving students' insight and topic knowledge.

METHOD

Criteria for selection

Our main objective was to investigate whether instruction directed at different (meta-)cognitive writing processes, such as specified by Klein's (1999) hypotheses, was effective for acquiring new insight and topic knowledge in educational contexts. Therefore, our criteria for inclusion focus on aspects of instruction and education. Regarding the design of the studies, we preferred a liberal policy and included experimental studies as well as case studies. In doing so, we included empirical studies that applied as many various modes of instruction as possible, in as many types of educational contexts as possible. More specifically, we used the following criteria for including studies:

- The study is an empirical study (experimental, quasi experimental, or a case study).
- The study is aimed at writing-to-learn.
- The study measures effects on insight and/or topic knowledge concerning a disciplinary subject.
- The study is embedded in a regular classroom context.
- The study is directed at grade 5 or higher. Grade 5 students having acquired the basic principles of writing are able to reflect on their own and other students' writing and to revise texts when instructed and stimulated (Van Gelderen, 1997). This reflection can be considered conditional for the processes of writing-to-learn (Bereiter & Scardamalia, 1987).
- Instruction in writing-to-learn is clearly described (*writing assignments* explaining to students *what* to write about and *instruction how* to proceed when writing).
- The pre- and post-tests are clearly described. Because we wanted to investigate effects of instruction on insight and topic knowledge, we needed to know what exactly was measured and how. For tests comprising open-ended questions or multiple-choice items, the aimed outcomes are respectively considered insight and topic knowledge, unless researchers explicitly aimed otherwise (e.g., multiple choice items measuring insight). For other measures, for instance writing assignments, assessment criteria should be directed at insight (e.g., students can provide a summary of lesson contents) and/or topic knowledge (e.g., students can recall lesson contents).

The following criterion was used for exclusion:

- The study is directed at gifted students only or at special needs students only.

Procedure for search and selection of literature

We used automatic and manual search methods. Two searches were undertaken. The starting point of the first search was 1999, the year of publication of Klein's review indicating a reorientation in research on writing-to-learn, resulting in an important role for instruction to evoke learning by writing. Combined in one search, the databases ERIC and PsycINFO were consulted for published studies. In order to ascertain that no relevant studies were missed, we systematically searched the bibliographies of the studies that were found relevant for additional studies (so-called snowballing). This search ended in September 2018. The second search provided an update and covers the period from September 2018 to August 2021.The databases ERIC and PsycINFO were consulted for published studies. Additionally, we systematically searched the bibliographies of the studies that were found relevant (snowballing). For an overview of the searches, see Figure 1 (based on Page et al., 2020).

To secure the validity of the selection procedure (see below), the three authors discussed the selection of studies proposed by the first author and decided on the basis of consensus whether studies should be included and for which of the discriminated types of instruction they were representative. An example is Nevid et al. (2017) using elaborated writing assignments. The question was whether the writing assignments also included some type of instruction directed at writing-to-learn. The conclusion was that this was not the case.

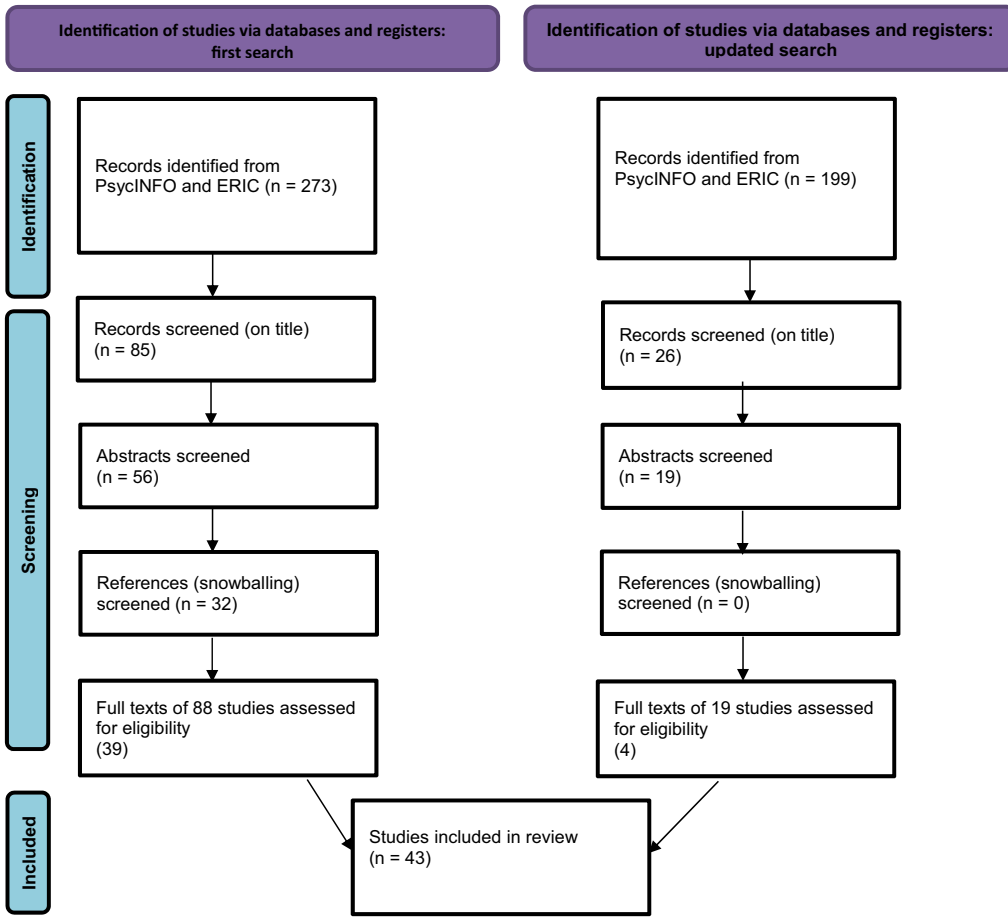

FIGURE 1 Flow diagram of the searches (Page et al., 2020)

The writing assignments described the context of the task precisely but did not provide (meta-)cognitive procedures for the writing process (such as planning, revising or using genre knowledge). Therefore, the study was not included.

Another example is Mateos et al. (2018), a study directed at learning to write syntheses. Students were required to write about contents from sources some of which were contradicting each other. The question was whether the assignment was directed only at learning to write or comprised writing-to-learn as well. It was decided that the latter was the case, because next to writing about contents from sources, students also had to reflect on how to reconcile conflicting contents when writing their synthesis. Therefore, this study was included.

In both searches, we used the following descriptors: *writing-to-learn, learning by writing, writing as a learning…* (the latter is the start of a phrase)*, insight writing, deep learning writing, critical thinking writing, writing in the disciplines, disciplinary writing, written argumentation, argument writing, topic knowledge writing* and *synthesis writing*.

This yielded a total of 273 records, from which a first selection was made based on the titles. This resulted in a selection that consisted of 85 studies. For a further refinement, we consulted all abstracts, which resulted in 56 studies. After snowballing in the references of these studies, we added another 32, leading to a total of 88 studies. Finally, after consulting the 88 full texts, 39 studies satisfied all inclusion criteria.

In August 2021 a new search was undertaken for updating the selection (see Figure 1). From 199 records, a selection based on the titles was made, resulting in 26 studies. Next, we consulted all abstracts, which resulted in a selection of 19 studies. After snowballing in the references no studies were added. Finally, after reading the 19 full texts, four studies were added to the selection from the first search, resulting in a total of 43 studies for this review.

RESULTS

Characteristics of the 43 studies

Table 1 provides an overview of all 43 studies that were selected for our review. The studies are ordered alphabetically according to the name of the first author. From Table 1, it can be inferred that 23 studies are related to knowledge about science, 8 to behavioural sciences, 6 to humanities, 5 social studies, and one to earth sciences.

In the 43 studies, 22 different genres are involved, for instance narrative, laboratory report and argumentative letter. The grade levels of students targeted in the studies vary from below grade 7 (six studies), between grade 7 and 12 (24 studies) to higher education (13 studies). The sample sizes vary from 50 and below (13 studies), between 50 and 100 (18 studies) and above 100 (12). For a few exceptions, in the large majority of selected studies more than five and less than 24 lessons were given. Only in three studies the number of lessons is much larger, namely in two studies it is 40 and in one 56.

The research design of 19 studies is quasi-experimental, and of five studies experimental with randomisation on the individual level. These 24 studies use a control group for comparison. Five of these apply the teachers' usual programme for control (business as usual). In 19 studies, the control group is adapted with regard to time on task and/or the content of the programme in order to achieve more experimental control. In 10 studies there are different experimental groups that are compared to each other. Five of these use randomisation on the individual level. The remaining nine studies are case studies.

In 20 studies, the post-tests are aimed at measuring insight as well as topic knowledge (on a combined measure or two separate measures). The post-tests and the delayed post-tests often measure the effects of writing-to-learn by analysing students' texts. In 34 studies,

TABLE 1 Overview of 43 studies

Study	Discipline/ Genre(s)	Grade level	N	Period+lessons	Writing assignments	Teacher, researcher or both
1 Atasoy and Küçük (2020)	Physics answers on open-ended questions	Grade 8	18	6 weeks 24 lessons	6	Teacher Researcher
2 Balgopal et al. (2012)	Science Essay	Undergrads	89	NA	3	Teacher
3 Boscolo and Carotti (2003)	Literature Commentary	Grade 9	50	28 weeks 56 lessons	12	Teacher
4 Corcelles Seuba and Castelló (2015)	Philosophy Argument	Grade 11	6	8 weeks 7 lessons	3	Teacher
5 Finkenstaedt-Quinn et al. (2017)	Chemistry summary	JuniorsSeniors	36	3 weeks	2	Teacher and 2 assistants
6 Granado-Peinado et al. (2019)	Psychology Synthesis	Undergrads	160	6 weeks 6 lessons	2	Researcher
7 Gunel et al. (2006)	Physics Synthesis	Grade 11	132	2 weeks 10 lessons	2	Teacher
8 Gunel et al. (2009)	Biology Explanation	Grade 9, Grade 10	118	10 lessons	1	Researcher
9 Hand et al. (2009)	Physics Explanation	Grade 10	181	21 lessons	2	Teacher

Experimental conditions and treatments	Control group	Randomisation	Post-test +delayed post-test	Covariates or pre-test	Significant results
Writing about disciplinary topics and students' epistemolo- gical views on it.	−	−	1. IN (interview, 5 questions)	IN (interview, 5 questions)	Growth on 1 (5 questions): 30% of the students
Writing three texts, each another aim, from viewpoint of one type of person, using a checklist	−	−	1. IN (WT)	IN (WT)	Growth on 1: 33 % of the students
Writing and discussing different genres aimed at literary comprehension	Business as usual	−	1. TK (WT) 2. IN (WT)	TK (WT) IN (WT)	Exp. cond > ctr. on 2: partial η^2= .22
Collaborative writing using planning guide	−	−	1. IN (WT, individual)	IN (WT, individual)	Growth on 1
Writing for a non-expert audience using double-blind peer review, revision of initial draft	−	−	1. IN (MC) 2. IN (WT)	IN (MC) IN (WT)	Growth of exp. cond from initial draft to final draft on 2
1. Instruction of checklist for synthesis writing and collaborative learning. Collaborative writing 2. Instruction of checklist for synthesis writing. Collaborative writing 3. Collaborative writing, using checklist 4. Collaborative writing	-	+ (individual	1. IN (WT, pairs) 2. IN (WT, individual)	IN (WT, pairs) IN (WT, individual)	Cond 1 > cond 3, on 1 and 2, Cohen's $d = 1.7$ Cond 1 > cond 4 on 1 and 2, Cohen's $d = 2.03$ Cond 2 > cond 4 on 1 and 2 Cohen's $d = 0.79$
Writing for teacher: 1. PowerPoint + script (*sample A*) 2. Plain paper (*sample B*) Writing for grade 10: 3. PowerPoint + script (*sample A*) 4. Plain paper (*sample B*)	−	+ (group)	1.TK (MC) 2. IN (OE)	1.TK (MC) 2. IN (OE)	Cond 1 > cond 2 on 1: Cohen's $d = 0.2$, and on 2: Cohen's $d = 0.6$ Cond 3 > cond 4 on 1: Cohen's $d = 0.8$, and on 2: Cohen's $d = 0.6$
Writing for: 1. 3rd/4th graders 2. parents 3. peers 4. teacher	−	+ (group)	1. TK (MC) 2. IN (OE)	1. Last years' biology grade 2. TK, IN (MC, OE) 3. TK, IN (MC, OE)	Growth of cond 1 − 4 on 2: partial η^2 =0.08
Writing 2 texts using: 1. math, math 2. math, graph 3. graph, graph 4. graph, math	Controlled	+ (group)	TK + IN (MC)	1. TK + IN (MC) 2. TK + IN (MC)	Cond 2 > cond 1: $d = 0.4$ cond 3: $d = 0.5$ cond 4: $d = 1.0$ ctr: $d = 0.5$ Cond 1 > cond 3: $d = 0.2$ cond 4: $d = 0.6$ ctr: $d = 0.2$ Ctr > cond 4: $d = 0.5$

(Continues)

TABLE 1 (Continued)

Study	Discipline/ Genre(s)	Grade level	N	Period+lessons	Writing assignments	Teacher, researcher or both
10 Hand, Hohenshell, and Prain (2004)	Biology 1. Explanation 2. Newspaper article	Grade 10	73	6 weeks 6 lessons	2	Teacher English teacher
11 Hand, Hohenshell, and Prain (2007)	Biology 1. Explanation 2. Newspaper article	Grade 10	87	6 weeks	2	Teacher Researcher
12 Hand, Wallace, and Yang (2004)	Biology Summary	Grade 7	93	8 weeks 40 lessons	1	Teacher
13 Hand, Young, and Bruxvoort (2007)	Chemistry Business letter	Grade 11	52	2 lessons	1	Teacher
14 Hohenshell and Hand (2006)	Cell biology Laboratory report	Grade 9, Grade 10	91	7 weeks	7	Teacher
15 Hunter and Tse (2013)	Economics Essay	University students	1031	8 weeks 6 lessons	2	Teacher Tutors
16 Kabataş Memiş and Öz (2017)	Science Text with embedded formulas or graphs	Grade 5	32	NA	2	Teacher
17 Kieft et al. (2006)	Dutch literature Literary review	Grade 10	113	5 weeks 5 lessons	5	Teacher
18 Kieft et al. (2008)	Dutch literature Literary review	Grade 10	113 (of 220)	5 weeks 5 lessons	5	Teacher

Experimental conditions and treatments	Control group	Randomisation	Post-test +delayed post-test	Covariates or pre-test	Significant results
1. Generating, organising, then writing, two texts 2. Generating, writing, then organising, two texts 3. Generating, organising, then writing, one text 4. Generating, writing, then organising, one text	−	+ (group)	After writing one text: 1. IN (WT) 2. IN (OE) After writing two texts: 3. IN (OE) Delayed: 4. IN (OE)	Last semester's biology grade	Cond 1 and 3*> cond 2 and 4* on 1: $d = 0.49$ Cond 1 and 2*> cond 3 and 4* on 3: $d = 0.70$ on 4: $d = 1.09$
Writing two texts Writing two texts (a, b) with or without instruction: 1a. no instruction, then (*1b*) reviewing by audience 2a. generating, organising, reviewing by audience, then (2*b*) reviewing by audience 3a. reviewing by audience, then (3*b*) generating, organising, reviewing by audience 4a. no instruction, then (4*b*) generating, organising, reviewing by audience	−	−	After writing one text: 1. IN (1 OE) 2. IN (1 OE) 3. IN (1 OE) 4. IN (total OE) After writing two texts: 5. IN (1 OE) 6. IN (1 OE) 7. IN (1 OE) 8. IN (total OE)	Last school year's grade Before writing first text: 1. IN (1 OE) 2. IN (1 OE) 3. IN (1 OE) 4. IN (total OE) Before writing second text: 5. IN (1 OE) 6. IN (1 OE) 7. IN (1 OE) 8. IN (total OE)	Cond 1a > 2a on 1: $\eta^2 = 0.162$ Cond 2a > cond 1a on 3: $\eta^2 = 0.222$ Cond 3a > cond 4a on 2: $\eta^2 = 0.296$, on 3: $\eta^2 = 0.097$ and on 4: $\eta^2 = 0.255$ Cond 1a > 2a and 4b on 1: $\eta^2 = 0.71$ Cond 1b, 3a and 2a+2b*> 4a on 6: $\eta^2 = 0.215$, on 7: $\eta^2 = 0.141$ on 8: $\eta^2 = 0.183$
Writing using: 1. checklist 2. checklist, and writing for an audience	Controlled	−	1. TK (MC) 2. IN (OE)	1. TK (MC) 2. IN (OE)	Cond 1 > ctr on 1: $d = 0.15$ Cond 2> ctr on 1: $d = 0.29$ Cond 2 > ctr on 2: $d = 1.02$
Writing for an audience.	Controlled	+ (group)	1. TK (OE) 2. IN (1 OE) 3. IN (1 OE) 4. IN (1 OE)	1. TK (OE) 2. IN (1 OE) 3. IN (1 OE) 4. IN (1 OE)	Exp. cond > ctr on 4: $d = 0.83$
Writing using checklist, for: 1. teacher 2. peers	Controlled	−	1. TK (MC) 2. IN (OE)	1. TK (MC) 2. IN (OE)	Cond 1 and 2*> ctr on 2: partial $\eta^2 = 0.114$
Writing text 1: discussion of content, instruction on genre structuring Writing text 2: discussion of teacher's feedback on WT1 (= feed forward)	Ctr 1 Ctr 2 Business as usual	−	After four weeks: 1. IN (WT 1) After 8 weeks: 2. In (WT 2)	−	Growth of exp. cond. on 2: Exp cond > ctr 1 on 2: $d = 1.23$ Exp cond > ctr 2 on 2: $d = 0.07$
Writing by using multi-modal representations	Controlled	+ (group)	1. IN (OE)	IN (OE)	Exp. cond > ctr on 1: $\eta^2 = 0.449$
1. First generating and organising, then writing 2. First generating, then writing, then organising	−	+ (individual)	1. IN (OE)	1. IN (OE)	Cond 1 > cond 2 on 1: $d = 0.03$
1. First generating and organising, then writing 2. First generating, then writing, then organising	−	+ (individual)	1. IN (OE)	IN (OE)	−

(Continues)

TABLE 1 (Continued)

Study	Discipline/ Genre(s)	Grade level	N	Period+lessons	Writing assignments	Teacher, researcher or both
19 Klein and Ehrhardt (2015)	Science 1 Argument 2 Discussion	Grade 7, Grade 8	72	10 days	1	Teacher Research assistant
20 Klein and Rose (2010)	Science 1 Argument 2 Explanation	Grade 5, Grade 6	34	20 weeks	16	Researcher Teacher
21 Klein and Kirkpatrick (2010)	Science 1. Argument 2. Explanation	Grade 5, Grade 6	113	24 weeks 20 lessons	16	Teacher
22 Klein and Samuels (2010)	Science Argument	Grade 7, Grade 8	60	20 weeks 40 lessons	20	Teacher
23 Martinez et al. (2015)	History Synthesis	Grade 6	62	5 weeks 15 lessons	3	Researcher Teacher
24 Mateos et al. (2018)	Psychology Synthesis	Undergrads	114	3 lessons	2	Researcher
25 McDermott and Hand (2013)	Chemistry Text with embedded formulas or graphs	Grade 10, Grade 11, Grade 12	70	NA	2	Teacher
26 McDermott and Hand (2013)	Chemistry Text with embedded formulas or graphs	Grade 10, Grade 11, Grade 12	95	NA	1	Teacher
27 Nam et al. (2011)	Geology Explanation	Grade 8	345	20 weeks 8 lessons	4	Teacher
28 Nevid et al. (2012)	Introductory psychology Reflection	Undergrads	135	20 weeks	16	Researcher
29 Nevid et al. (2012)	Introductory psychology Reflection	Undergrads	55	20 weeks	16	Researcher
30 Nückles et al. (2010)	Psychology Learning journal	Undergrads	50	16 weeks 12 lessons	12	Researcher
31 Nückles et al. (2010)	Psychology Learning journal	Undergrads	62	16 weeks 12 lessons	12	Researcher teacher
32 Ritchie et al. (2011)	Biology Stories	Grade 6	55	6 weeks	2	Teacher

Experimental conditions and treatments	Control group	Randomisation	Post-test +delayed post-test	Covariates or pre-test	Significant results
1.Stepwise instruction on elaborating writing subgoals 2. Clustered instruction on writing goals	−	+ (individual)	1. TK (MC)	1. TK (MC)	−
Instruction of genre knowledge and cognitive strategies	Controlled	+ (group)	1. TK, IN (OE)	1. TK, IN (OE)	Exp. cond > ctr on 1: partial $\eta^2 = 0.21$
Instruction of genre knowledge and cognitive strategies	Controlled	+ (group)	1. TK, IN (Cloze, OE)	1. TK, IN (OE)	−
Instruction of genre knowledge	Controlled	+ (group)	1. TK, IN (Cloze, OE)	1. TK, IN (OE)	−
Instruction of strategies on synthesis writing, resulting in checklist for writing	Controlled	+ (group)	1. TK, IN (OE)	1. TK, IN (OE)	Exp. cond > ctr on 1: partial $\eta^2 = 0.64$
1. Instruction of checklist for synthesis writing. Collaborative writing using checklist 2. Self-study of checklist for synthesis writing. Collaborative writing using checklist	-	+ (individual)	1. IN (WT, individual)	2. IN (WT, individual)	Cond 1 > cond 2 on 1: $p < 0.001$ $\eta^2 = 0.13$
Instruction of embedding graphs and formulas in text, resulting in checklist for self-assessing the writing	Controlled	+ (group)	1. TK, IN (MC, OE) (after unit 1) 2. TK, IN (MC, OE) (after unit 2)	TK (MC)	Exp. cond > ctr on 2: $d = 0.53$
Instruction of embedding graphs and formulas in text, resulting in checklist for self-assessing the writing	Controlled	+ (group)	1. TK, IN (MC, OE)	TK (MC)	Exp. cond > ctr on 1: $d = 0.62$
Writing using checklist: 1. sample A 2. sample B 3. sample C	Business as usual	−	TK, IN (OE)	IN (OE)	Cond 1 > ctr $d = 0.61$ Cond 2 > ctr $d = 0.64$
Choice of 16 topics out of a long list	−	−	1. TK (MC)	−	Exp. cond: topics students have written about > other topics
Obligatory topics Choice between two genres	−	−	1. TK (MC)	−	Exp. cond: topics students have written about > other topics
Writing using cognitive and metacognitive hints	Controlled	+ (individual)	1. IN (OE) after 8 weeks 2. IN, (OE) after 16 weeks	−	Cond > ctr on 1: partial $\eta^2 = 0.08$
Writing using: 1. cognitive and metacognitive hints 2. fading cognitive and metacognitive hints	Controlled	+ (individual)	1. IN (OE) after 8 weeks 2. IN, (OE) after 16 weeks	−	Cond 1 and 2 > ctr on 1: partial $\eta^2 = 0.09$ Decrease of cond 1 from 1 to 2; partial $\eta^2 = 0.15$
Writing biology narratives	Controlled	−	1.TK, IN (BioQuiz questionnaire)	1.TK, IN (BioQuiz questionnaire)	Exp cond > ctr on 1: $d = 0.39$

(Continues)

TABLE 1 (Continued)

Study	Discipline/ Genre(s)	Grade level	N	Period+lessons	Writing assignments	Teacher, researcher or both
33 Rivard and Straw (2000)	Ecology Explanation	Grade 8	22	5 weeks 5 lessons	5	Teacher
34 Rivard (2004)	Ecology Explanation	Grade 8	77	5 lessons	5	Teacher
35 Sampson and Walker (2012)	Chemistry Research report	Undergrads	18	20 weeks 15 lessons	4	Teacher
36 Stewart et al. (2010)	Psychology Micro- theme	Juniors, Seniors	73	5 weeks 10 lessons	10	Teacher
37 Van Drie et al. (2015)	History Argumentative letter	Grade 11	42	5 lessons	2	Researcher Teacher
38 Wäschle et al. (2015)	Immunology Learning journal	Grade 7	46	3 weeks 6 lessons	3	Teacher
39. Wäschle et al. (2015)	Philosophy Learning journal	Grade 10	24	8 weeks 6 lessons	6	Teacher
40 Winstead Fry and Villagomez (2012)	Pedagogics Learning journal	Juniors, Seniors	53	15 weeks 15 lessons	15	Researcher
41 Wissinger and De La Paz (2016)	History Essay	Grade 6 Grade 7	151	3 weeks 15 lessons	3	Teacher 3 assistants
42 Wong et al. (2002)	English literature Response journal	Grade 12	48	4 weeks 9 lessons	2	Teacher
43 Yildiz and Akdag (2021)	History Story Column	Teacher training	70	6 weeks 6 lessons	6	Researchers

Abbreviations: * two or more conditions taken together; − means not present; + means present; > outperforms; cond, condition; ctr, control group; exp, experimental; IN, insight; MC, multiple choice; NA, not available; OE, open-ended questions; quest, questionnaire; TK, topic knowledge; WT, writing task.

Experimental conditions and treatments	Control group	Randomisation	Post-test +delayed post-test	Covariates or pre-test	Significant results
Discussion before writing	Controlled	+ (individual)	1. TK (MC, OE, concept maps) 2 IN (MC, OE, concept maps) Delayed: 3. TK (MC, OE, concept maps) 4. IN (MC, OE, concept maps)	1. TK, IN (MC, OE, concept maps).	Exp cond > ctr on 1: $d = 1.0$ and on 3: $d = 1.21$ Exp cond > ctr on 2: $d = 0.60$ and on 4: $d = 1.09$
Discussion before writing	Controlled	+ (individual)	1.TK (MC) Delayed: 2.TK (MC)	TK (MC)	–
Writing four initial drafts and four revised versions using a checklist and a double-blind peer-review guide	–	–	1. TK + IN (WT: initial drafts of 4 WTs) 2. TK + IN (4 revised versions of 4 WTs)	–	Growth on 2 > growth on 1 for all texts
Teacher's feedback and grade after each assignment to: 1. non-honours 2. honours	Business as usual (non-honours only)	–	1. TK (MC) 2. TK, IN (WT)	–	Cond 1 > ctr on 1: $d = 0.89$, and on 2: $d = 0.60$
1.Disciplinary argument writing 2. Non-disciplinary argument writing	–	+ (individual)	1. TK (OE) 2. IN (WT)	1.TK (OE)	Growth of cond 1 and 2 on 1
Writing using cognitive and metacognitive hints.	Controlled	–	1. IN (OE) Delayed: 2. IN (OE) 3. IN (WT)	1.TK, IN (OE)	Exp. cond > ctr on 1: partial $\eta^2 = 0.18$ Exp. cond > ctr on 2: partial $\eta^2 = 0.31$ Exp. cond> ctr on 3: partial $\eta^2 = 0.24$
Writing using: 1. cognitive and metacognitive hints 2. cognitive, metacognitive, hints and motivator	–	–	1.TK, IN (OE) 2. IN (WT)	–	Cond 2 > cond 1 on 2: partial $\eta^2 = 0.41$
Teacher's feedback on each text after each class.	Controlled	–	TK (MC)	TK (MC)	–
Discussing argumentation schemes before writing	Business as usual	+ (individual)	TK (MC)	TK (MC)	Exp. cond > ctr on 1: partial $\eta^2 = 0.23$
Writing using: 1. general analytic prompts 2. specific analytic prompts	Controlled	+ (group)	1. IN (OE) 2. IN (OE)	–	Cond 1 > ctr on 1: $d = 1.30$ cond 2 > ctr on 1: $d = 0.84$ Cond 1 > ctr on 2: $d = 0.59$ cond 2 > ctr on 2: $d = 1.15$
1. instruction on collaboratively writing a story for peers 2. instruction on collaboratively writing a column for peers	–	+ (indivi-dual)	1. IN (OE)	IN (OE)	Cond 1 > cond 2: $\eta^2 = 0.13$

covariates (pre-tests measuring insight and/or topic knowledge) are employed. In three studies, last semesters' grades are added to the pre-test scores or used as the only covariate. In nine studies no covariates are used.

In the last column of Table 1, effects of the treatments in the studies are presented. In case of comparisons between conditions, only significant effects are shown, using the condition numbers involved (as shown in the column named *experimental conditions and treatment*s). In the case that comparisons involve differences in growth between conditions, this is indicated by the term 'growth'. Effect sizes (Cohen's *d,* eta squared or partial eta squared) are reported in the majority of the studies. In the cases that these were not reported, we computed Cohen's *d* when significant effects were found between experimental and control groups or between different experimental groups.

Four types of instruction

Table 2 shows four types of instruction that were distinguished in the selected studies, Forward Search, Genre Writing, Backward Search and Planning Only. Three of them are based on Klein's hypotheses as described in the introductory section. A large part of the studies can be classified by means of the three hypotheses proposed by Klein (1999). However, we found a substantial number of studies that deviated from these hypotheses, because the instruction was based solely on planning for writing. There was no attempt in these studies to combine planning and revision, such as is the case in studies classified as Backward Search. Therefore, we added a fourth type of instruction in our review, called Planning Only.

Table 2 describes the four types of instruction in the 43 studies. We used the processes planning and reviewing based on the writing model of Hayes and Flower (1980) and consistent with Klein's (1999) hypotheses, to specify the more precise processes that were involved in each study. Planning processes involve organising, and/or generating. Reviewing processes involve the following four instructed activities: (1) *feedback by* (type of readers that provide feedback), (2) *feedback on* (the focus of feedback given), (3) *revising* (whether it occurs or not) and (4) *focus of revision* (what students attend to while revising).

The first type of instruction is aimed at eliciting the Forward Search process. Students are told to start writing their ideas without planning and to review their draft afterwards. This instruction focuses on reviewing to reinforce thinking about the contents of the draft. The second type of instruction*,* Genre Writing, aims at genre knowledge relating to the genre that students have to write. Instruction focuses on the genre-specific structure of a model text. Furthermore, instruction may also explain the linguistic elements that need to be used to realise a genre-specific way of writing. The third type of instruction is aimed at Backward Search. This type of instruction aims at planning activities, and at reviewing activities that are focused on the rhetorical and content goals set in the planning activities. The fourth type of instruction in Table 2 concerns instruction on planning activities only.

As Table 2 shows, some studies have characteristics of more than one type of instruction as distinguished above. Nevertheless, these studies are attributable to one dominant type based on the authors' descriptions of their main interests.

Effects of forward search instruction

As can be seen in Table 2, instruction in six studies is characterised as Forward Search. These studies took place in secondary education or university. When students receive Forward Search instruction, they are asked to write down all they can think of. Then, they receive feedback from a reviewer for revising their texts.

In three studies (1, 5, 8), reviewing comprises feedback *and* revising conforming to Klein's (1999) description of the Forward Search process, whereas in the remaining three studies (36, 40, 42), revising is postponed to a next writing assignment. In study 1, students are instructed to write answers on open-ended questions, then review a peer's answers, and revise their own texts using the received feedback. Finally, students and teacher provide feedback in a classroom discussion, and if needed students revise their texts again. In studies 5 and 8, students are instructed to write to various types of audiences (to a non-expert audience, peers, parents, younger students and the teacher). These audiences provide feedback on comprehensibility of the drafts, which students use for revising.

In the other three studies (36, 40, 42), feedback is given in two ways: (1) a teacher provides written feedback on the application of, or reflection on course concepts; (2) a teacher as well as students provide feedback in classroom discussion, focusing on students' responses on prompts about general and specific aspects of a story (e.g., *What do you notice in the story?*; *How do you feel about Daisy?*).

Of the six studies, there are three comparing Forward Search instruction to a business-as-usual control group (36) or a controlled control group (40 and 42). Two of these (36 and 42) report positive effects of Forward Search instruction. In study 36, two of four comparisons show positive effects on topic knowledge (large effect) and on a combined measure of insight and topic knowledge (medium effect). However, the other two comparisons regard a difference between honours and non-honours students, which cannot be considered as experimental comparisons of the intervention effects. Furthermore, initial differences between the (non-honours) experimental and the (non-honours) control group are not accounted for. Therefore, the difference in favour of the non-honours group receiving feedback on its writing may be attributable to other variables than the experimental intervention. In study 42, all four comparisons show positive effects of instruction on insight on two measures of insight (three large effect sizes, one medium). The third study (40) does not show any effect of the comparison on topic knowledge.

The remaining studies (1, 5 and 8) do not compare the results of the experimental conditions with a control group. The experimental conditions in study 1 and 5 show growth of insight, just as all four conditions in study 8.

Summarising, three of six studies compare experimental conditions with a control group, and two of these (36 and 42) show positive effects on insight and topic knowledge, whereas the effects of study 36 may not be attributable to Forward Search instruction. Six of nine comparisons lead to positive results (four large and two medium effects). Additionally, study 1, 5 and 8 provide evidence that feedback by peers or various audiences followed by revision results in growth on insight.

Effects of genre writing instruction

Table 2 shows that in 14 studies, instruction is characterised as Genre Writing. These studies are directed at educational levels varying from grade 5 to university. In all of these studies, instruction contains preparatory lessons on the characteristics of the genre in which students have to write.

The genres concerned are quite varied, such as essay, literature commentary, history synthesis, text with embedded modes, learning journal and narrative (see Table 1). According to Table 2, a number of these studies also contain instruction aiming at planning and/or revising in addition to genre knowledge, as follows. A group of three studies (3, 22 and 37) only focuses on the core business of Genre Writing instruction (genre knowledge). A second group consisting of six studies (2, 19, 30, 31, 38 and 39) contains planning instruction in addition to Genre Writing instruction. A third group consisting of five studies (21, 23, 25, 26 and 32)

TABLE 2 Types of instruction

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
Forward Search instruction						
1 Atasoy and Küçük (2020)	Cond 1	–	Peer, teacher and class	+	+	Students' epistemological views
5 Finkenstaedt-Quinn et al. (2017)	Cond 1:	–	Peers	+	+	Comprehensibility to audience Discipline based characteristics
8 Gunel et al. (2009)	Cond 1: Cond 2: Cond 3: Cond 4:	–	3rd/4th grade Parents Peers Teacher	+	+	Comprehensibility to audience
36 Stewart et al. (2010)	Cond 1: non honours Cond 2: honours	–	Teacher	Topic knowledge, application of course concept	–	–
40 Winstead Fry and Villagomez (2012)	Cond 1	–	Teacher	Reflecting	–	–
42 Wong et al. (2002)	Cond 1: general analytic prompts Cond 2: specific analytic prompts	–	Class and teacher	Students' responses to the prompts	–	–
Genre Writing instruction						
2 Balgopal et al. (2012)	Cond 1	Generating	–	–	–	–
3 Boscolo and Carotti (2003)	Cond 1	–	–	–	–	–
20 Klein and Rose (2010)	Cond 1	Generating Organising	–	–	–	–
21 Klein and Kirkpatrick (2010)	Cond 1	Generating	Self	+	+	Genre-based characteristic

TABLE 2 (Continued)

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
22 Klein and Samuels (2010)	Cond 1	−	−	−	−	−
23 Martinez et al. (2015)	Cond 1	Generating Organising	Self	−	+	−
25 McDermott and Hand (2013)	Cond 1: 2 writing assignments	−	Self An audience other than the teacher	+ −	+ +	Embeddedness −
26 McDermott and Hand (2013)	Cond 1: 1 writing assignment	−	Self An audience other than the teacher	+ −	+ +	Embeddedness −
30 Nückles et al. (2010)	Cond 1	Generating	−	−	−	−
31 Nückles et al. (2010)	Cond 1: (meta) cognitive hints Cond 2: same hints faded	Generating	−	−	−	−
32 Ritchie et al. (2011)	Cond 1:	Generating	Peers	−	+	−
37 Van Drie et al. (2015)	Condi 1: Disciplinary writing Cond 2: Non-disciplinary writing	−	−	−	−	−
38 Wäschle et al. (2015)	Cond 1	Generating	−	−	−	−
39 Wäschle et al. (2015)	Cond 1: (meta)cognitive hints Cond 2: personal utility prompt	Generating	−	−	−	−

Backward Search instruction

(Continues)

TABLE 2 (Continued)

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
4 Corcelles Seuba and Castelló (2015)	Cond 1	Generating Organising	Peers, teacher	+	+	Text structure, content
6 Granado-Peinado et al. (2019)	Cond 1: by means of video modelling writing using checklist, and collaborative strategies	Generating Organising	-	-	+	Generating Organising
	Cond 2: by means of video modelling writing using checklist, and collaborative processes	Generating Organising			+	Generating Organising
	Cond 3 writing using checklist, and collaborative processes	Generating Organising			+	Generating Organising
	Cond 4: collaborative processes	Organising				
7 Gunel et al. (2006)	Cond 1: PowerPoint	Organising	Teacher	Organising, concept knowledge	−	−
	Cond 2: paper	Organising	Teacher	Organising, concept knowledge	−	−
	Cond 3: PowerPoint	Organising	Younger students	Comprehensibility	−	−
	Cond 4: paper	Organising	Younger students	Comprehensibility	−	−

TABLE 2 (Continued)

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
9 Hand et al., 2009	Cond 1: math, math	Generating	Younger students,	+	+	Comprehensibility, selecting, organising
	Cond 2: math, graph	Organising	teacher	+	+	idem
	Cond 3: graph, graph	Idem	idem	+	+	idem
	Cond 4: graph, math	Idem	idem	+	+	idem
		Idem	idem			
10 Hand, Hohenshell, and Prain (2004)	Cond 1: explanation,	Generating	Younger students	+	+	Comprehensibility
		Organising				
	newspaper article	Generating	Newspaper editor	+	+	Comprehensibility
		Organising				
	Cond: 2: explanation,	Generating	Self	+	+	Organising
			Younger students	+	+	Comprehensibility
	newspaper article	Generating	Peers	+	+	Organising
			Newspaper editor	+	+	Comprehensibility
	Cond 3: explanation	Generating	Younger students	+	+	Comprehensibility
		Organising				
	Cond 4: explanation	Generating	Self	+	+	Organising
			Younger students	+	+	Comprehensibility
11 Hand, Hohenshell, and Prain (2007)	Condition 1a: textbook assignments	–	–	–	–	–
	Condition 1b: newspaper article	–	Newspaper editor	+	+	Comprehensibility
	Condition 2a: explanation	Generating	Younger students	+	+	Comprehensibility
		Organising	Teacher	+	+	Concepts
	Condition 2b: newspaper article	–	Newspaper editor	+	+	Comprehensibility
	Condition 3a: newspaper	–	Newspaper editor	+	+	Comprehensibility
		Generating	Younger students	+	+	Comprehensibility
	Condition 3b: explanation	Organising	Teacher	+	+	Concepts
		–	–	–	–	–
	Condition 4a: textbook assignments	Generating	Younger students	+	+	Comprehensibility
	Condition 4b: explanation	Organising	Teacher	+	+	Concepts

(Continues)

TABLE 2 (Continued)

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
12 Hand, Wallace, and Yang (2004)	Cond 1	Generating	–	–	–	–
	Cond 2	Generating	Peers	+	+	Comprehensibility
			Teacher	+	+	Comprehensibility
13 Hand, Yang, and Bruxvoort (2007)	Cond 1: business letter	Generating	Younger students	+	+	Comprehensibility, readability
14 Hohenshell & Hand (2006)	Cond 1:	Organising Generating	Self	+	+	Representation of terms, concepts
	Cond 2:	Organising Generating	Peers	+	+	Representation of terms, concepts
15 Hunter and Tse (2013)	Text 1	Generating Organising	Teacher	Student's analysis of content of the future text	–	–
	Text 2	Generating and organising, both based on feedback on text 1			–	–
17 Kieft et al. (2006)	Cond 1	Generating Organising	Peers	–	–	–
	Cond 2	Generating	Self	+	+	Organising
			Peers	–	–	–
18 Kieft et al. (2008)	Cond 1	Generating	Peers	–	–	–
	Cond 2	Organising Generating	Self	+	+	Organising
			Peers	–	–	–
24 Mateos et al. (2018)	Cond 1 checklist (by teacher)	Generating Organising		+	+	Generating, Organising Content of argument
	Cond 2 self-study of checklist	Generating Organising			+	Generating, Organising Content of argument

TABLE 2 (Continued)

Author, date	Experimental conditions	Planning	Reviewing			
			Feedback by	Feedback on	Revising	Focus of revision on
35 Sampson and Walker (2012)	Cond 1	Generating Organising	Peers	+	+	Goals, content of argument
Planning Only instruction						
16 Kabataş Memiş and Öz (2017)	Cond 1	Organising	−	−	−	−
19 Klein and Ehrhardt (2015)	Cond 1 stepwise goals	Generating Organising	−	−	−	−
	Cond 2 clustered goals	Organising	−	−	−	−
27 Nam et al. (2011)	Cond 1: sample A	Generating	−	−	−	−
	Cond 2: sample B	Generating	−	−	−	−
	Cond 3: sample C	Generating	−	−	−	−
28 Nevid et al. (2012)	Cond 1: relating chosen topic to learned subject matter	Organising	−	−	−	−
	Cond 2: relating chosen topic to experiences	Organising	−	−	−	−
29 Nevid et al. (2012)	Cond 1: relating assigned topic to chosen subject matter experiences.	Organising	−	−	−	−
33 Rivard and Straw (2000)	Cond 1 Sample: 11	Generating	−	−	−	−
34 Rivard (2004)	Cond 1 Sample: 39	Generating	−	−	−	−
41 Wissinger and De La Paz (2016)	Cond 1	Generating Organising	−	−	−	−
43 Yildiz and Akdag (2021)	Cond 1: collaboratively story writing	Generating Organising	−	−	−	−
	Cond 2: collaboratively column writing	Generating Organising	−	−	−	−

Note: + means present; − means not present; + in column *Feedback on* means mentioned in *Focus of revision*.

adds planning instruction combined with feedback and revision activities or adds feedback in combination with revision activities.

Plain genre writing instruction

Of the three studies focusing on genre knowledge, two studies (3, 22) compare one experimental condition with a control group. In study 3, students are instructed about literature by writing and discussing several genres (e.g., notes, minutes, synthesis). In comparison to a business-as-usual control group, the experimental condition achieves higher scores on insight, which are found on a post-test writing task (literary commentary, large effect), but not on a measure of topic knowledge based on the same writing task. In study 22, instruction in the genre 'argument writing' is given by means of three consecutive steps: (1) modelling an argumentative text (teacher), (2) shared writing (teacher *and* students), and (3) guided writing (students, with the help of peers or teacher). This is an approach for language teaching called Genre Pedagogics (Rose, 2008). The experimental condition receiving this type of instruction does not outperform the control group on insight and topic knowledge in science (combined measure).

In study 37, two experimental conditions, on historical argument writing and non-disciplinary argument writing, are compared to each other on insight and topic knowledge. Both groups show growth in topic knowledge.

Genre writing instruction complemented with planning

Four studies (20, 30, 31 and 38) compare experimental conditions with controlled control groups (20, 30 and 31) or a business-as-usual control group (38). In study 20, Genre Writing instruction takes place by means of Genre Pedagogics. This is complemented by instruction in planning containing inquiry activities (generating) and setting rhetorical goals (organising). The experimental condition outperforms the control group on insight and topic knowledge in the field of science (one combined measure, large effect). In study 30, instruction in psychology classes is focused on explaining characteristics of learning journals, added with hints for writing (generating). This study reports positive effects on insight (one measure, medium effect) for the experimental condition, when compared to a control group after 8 weeks. However, no effects are found on another, second measure of insight after 16 weeks.

Study 31 (a follow-up of 30) uses the same instruction and measures as study 30, but adds a second experimental condition in which the hints for planning are faded. Fading starts halfway the first eight weeks. After eight weeks, the two experimental conditions, which are taken together, outperform the control group on insight (one measure, medium effect). However, after 16 weeks the faded experimental condition does not show a significant difference with the control group on the measure of insight. The condition receiving permanent hints even shows a significant decrease on insight after 16 weeks. In study 38, instruction consists of an explanation of learning journals combined with hints for generating (just as in studies 30 and 31). The experimental condition outperforms the control group on one direct measure of insight and two delayed measures of insight (three large effects).

Study 39 (a follow-up of 38) compares two experimental conditions: (1) hints for generating, and (2) the same instruction complemented with an explanation of the utility of writing assignments to students (as an extra motivator). The comparison shows that the motivator

condition outperforms the other on a measure of insight (a writing assignment, large effect), but not on a combined measure.

Finally, study 2 (a case study) provides an instruction on essay writing, followed by students' discussion aimed at generating knowledge about the topic studied. In this study, growth in insight is measured by means of three consecutive writing tasks. These tasks prompt students to write about the same topic, from the angle of the same type of person (e.g., a farmer), but each time with a different aim: informing, expressing a feeling, and describing a dilemma. Growth in insight is found for 33% of the sample of 89 students.

Genre writing instruction complemented with planning and/or reviewing

The third group consists of five studies (21, 23, 25, 26 and 32). These studies compare one experimental condition with a controlled control group. In study 21, Genre Pedagogics is used for instruction. This instruction complemented with planning and revising tasks results in negative effects on a combined measure for insight and topic knowledge in the field of science. In study 23, model texts are used to instruct the experimental conditions in writing a synthesis of two source texts on history. Furthermore, students compose a checklist for self-assessing their texts on the characteristics of the genre *synthesis* after writing. Compared to a control group the experimental condition shows significantly higher scores on the combined measure for insight and topic knowledge (large effect). In study 25, the teacher explains a model text on chemistry containing embedded graphs and formulas and compares it to a text without graphs and formulas. After writing, students not only assess their texts themselves by means of an embeddedness checklist, but they also receive feedback from the audience they write for. In the post-test after the second unit, a positive effect is found for the experimental condition compared to a control group on a combined measure of insight and topic knowledge (medium effect). However, no effect is found in the post-test after the first unit. Study 26 is a follow up of study 25. It applies the same type of instruction, but uses a larger sample. In comparison to a control group, it leads to positive results on one combined measure (medium effect). In study 32, the teacher models a narrative text about science. Furthermore, students discuss the contents of their texts (generating). After writing a first draft, students are instructed to give feedback to each other. This study reports a positive effect (small effect) for the experimental condition compared to a control group on one combined measure for insight and topic knowledge.

To summarise the effects found for Genre Writing instruction, 11 studies (of 14) use experimental comparisons with control groups and 9 of these show positive results. Eleven of seventeen comparisons show positive effects on insight and topic knowledge. The remaining three studies not comparing experimental conditions with a control group (2, 37 and 39) show evidence that writing texts about a same topic with different aims (2), using general as well as disciplinary argument writing instruction (37), and the use of a utility prompt as a motivator (39) may help increase insight and topic knowledge.

Effects of backward search instruction

As Table 2 demonstrates, instruction in 14 studies is characterised as Backward Search. All studies are directed at the upper half of secondary education and university. Backward Search comprises instruction in planning and reviewing activities and is applied in various ways.

For planning, students are instructed to work with the Science Writing Heuristic in a number of studies. This heuristic requires students to perform laboratory activities, individually as well as collaboratively. The activities are aimed at selecting and organising contents for students' texts. In some studies, students are required to plan modalities (formulas, graphs) in their texts. Concerning the reviewing activities, in nearly all studies, students are instructed to revise their texts by using feedback from their audiences, which may be peers, younger students, teachers or the writers themselves. In two studies (7 and 15), writers receive feedback from their audiences, but do not revise their texts. However, in study 15, writers discuss the received feedback and use the outcomes for preparing a second text they are required to write just after finishing the reviewed draft.

Five studies (9, 12, 13, 14 and 15) contain a comparison between one or more experimental conditions receiving Backward Search instruction, and a controlled control group, or a business-as-usual control group. In studies 9, 12, 13 and 14, the control group works with an adapted programme, and time on task is controlled. The four experimental conditions in Study 9 receive instruction to write two texts, and to embed one mode (math or graph) in each text, each condition using another sequence of modes. The study demonstrates that Backward Search instruction comprising instruction in the sequences 'math and next graphs', and 'math and next math' leads to positive effects on insight and topic knowledge when compared to a control group (medium and small effect). These sequences show larger effects than the other two investigated sequences.

In study 12, one experimental condition receiving Backward Search instruction, and a second experimental condition receiving instruction in planning only, are compared to a control group. In both experimental conditions, students collaborate using the Science Writing Heuristic for generating contents. Students in the Backward Search condition write to an audience, and revise their texts based on their audiences' feedback. The comparison to a control group is based on two measures (insight and topic knowledge). Results show positive effects on both measures for the Backward Search instruction condition (small effect on topic knowledge and large on insight), while instruction in Planning Only leads to positive effects on topic knowledge (small effect).

In study 13, students are instructed to discuss their ideas (generating) as a preparation for a writing task. Furthermore, students revise their texts with the help of their audiences' feedback. One of four comparisons with a control group shows a significant effect on insight (large effect), but not on topic knowledge.

In study 14, students work with the Science Writing Heuristic in two conditions. In both conditions students write to an audience, a teacher or a peer, receive feedback and revise the text. This study has taken the two experimental conditions together for a comparison to a control group, which leads to positive effects on insight (large effect), but not on topic knowledge.

Study 15 compares one experimental condition to two business-as-usual control groups. Students in these groups complete the same writing tasks as the experimental condition, but do not receive Backward Search instruction. Results of the comparison to control group 1 after eight weeks show effects on insight (large effect) just as the comparison to control group 2 (small effect). Comparisons after four weeks do not yield positive effects.

Seven studies on Backward Search instruction do not contain comparisons with a control group, but compare different experimental conditions with each other (6, 7, 10, 11, 17, 18, 24) or are case studies without comparison (4, 35). In study 6, the number of instructional components declines from condition 1 to 4. Instruction in synthesis writing, collaborative learning and writing (the most components) shows the largest effects on insight.

In four comparisons, study 7 shows that in Backward Search instruction, the mode of writing (PowerPoints vs. plain paper) makes a difference for learning: preparing PowerPoints

resulted in positive effects on insight (two medium effects) and topic knowledge (large and medium effect).

Studies 10 and 17 show that conditions containing organising as part of planning outperform conditions in which text organisation is the focus of feedback. In study 10, these positive effects are found on one combined measure and on two measures of insight (small, medium and large effects), while study 17 finds a positive effect on one measure of insight (small effect). Furthermore, study 10 shows that positive effects on insight (two measures) are larger after writing two texts, than after writing one text (medium and large effects).

Study 11 elaborates on the latter outcomes of study 10. Four classes write two texts in different genres. The students receive Backward Search instruction, instruction in reviewing or no-instruction in various sequences. After students have written one text, Backward Search instruction shows larger effects on insight than no-instruction on one measure (medium effect), whereas on another measure, no-instruction shows a larger effect on insight than Backward Search instruction (medium effect). When compared to no instruction, instruction in reviewing leads to positive effects on insight. When instruction in reviewing is used separately as well as combined with Backward Search instruction in two consecutive tasks, positive effects on insight are found on two separate measures (two medium effects) and on a total measure of insight (a combination of scores on three separate measures, medium effect), in comparison to no instruction.

In study 18, a replication of study 17, instruction in neither condition leads to positive effects on insight. Study 24 compares instruction in using a checklist for synthesis writing by means of modelling and a video with instruction for self-studying the checklist. Students in both conditions collaboratively write a synthesis. Instruction in using the checklist leads to larger effects on insight than self-study.

Finally, in two case studies, students are instructed to work with a planning guide (4, 35), followed by a peer review guide in study 35. The studies show that both types of Backward Search instruction lead to growth in insight (4, 35), and topic knowledge (35).

Summarising, five studies (of 14) comparing Backward Search instruction with a control group provide evidence that there are positive effects on insight (four large, one small effect), on a combined measure (one medium, one small effect) and on topic knowledge (two small effects) for 50% of the experimental comparisons. Seven studies comparing different conditions for Backward Search instruction, additionally provide some evidence that elaborate instruction, the mode of writing (PowerPoints vs. plain papers), the number of writing tasks and specific planning instructions (organising vs. generating) are of importance for increasing insight and topic knowledge. In addition, two case studies provide evidence that Backward Search instruction may lead to growth in insight and topic knowledge.

Effects of planning only instruction

The fourth type of instruction in Table 2, Planning Only, is applied in nine studies. These studies are directed at grades 5 to 8, and at university level. Planning Only instruction is directed at planning activities in various ways. For instance, for generating, the Science Writing Heuristic and group discussions are applied, and for organising, (non-textual) modalities, such as formulas or graphs, are applied.

Five studies compare one or more experimental groups with a control group. The control group receives business as usual (27, 41) or controlled instruction (16, 33 and 34). In study 27, three experimental groups are instructed to use the Science Writing Heuristic for generating a research question, first individually and later collaboratively. Two experimental groups show significant effects on a combined measure of insight and topic knowledge (both medium effects). In study 41, directed at generating content, experimental students

are instructed to discuss collecting arguments from historical sources. The teacher uses an argumentation scheme and critical questions to stimulate students' discussion. The experimental group shows a large effect on a measure of topic knowledge.

In study 16, experimental students are instructed in embedding formulas and graphs for organising text contents, whereas control students carry out the writing task, without using such non-textual modalities. This Planning Only instruction leads to a large effect on one measure of insight. Study 33 instructs experimental students in generating by discussion in small groups. Comparisons to a control group performing other tasks (e.g., 'fill in the blanks'), show that generating by discussion leads to effects on two measures (direct and delayed) of insight (medium and large effects respectively) and two measures (direct and delayed) of topic knowledge (large effects). However, because of the small sample ($n = 11$), the positive effects have to be treated with caution. Study 34, a follow-up of study 33, uses the same intervention in a larger sample ($n = 39$). In this case, experimental comparisons do not result in significant effects on two measures of topic knowledge (direct and delayed).

In two studies (19, 43), two experimental conditions are compared to each other. Study 19 compares a clustered instruction providing directions for organising subgoals with instruction guiding students stepwise through generating and organising goals for writing.

No significant differences between the two conditions are found on a measure of topic knowledge. In study 43, students in two experimental conditions are instructed how to plan and write a text for peers collaboratively; in one condition they have to write a story, in the other a column. The study shows a medium effect on students' insight in favour of writing a story compared to writing a column.

The studies 28 and 29, of which 29 is a follow up of 28, do not contain comparisons with another condition. Instruction is directed at organising. Study 28 requires students to choose topics from a list. In study 29, students are instructed to choose between two genres. Both studies show positive effects on students' knowledge of topics they have written about, on one measure.

Summarising, in five studies (of nine), Planning Only instruction shows positive effects on insight (two large effects, one medium effect), on topic knowledge (three large effects) and on a combined measure of insight and topic knowledge (two medium effects), in comparison to a control group. In total, 73% of the experimental comparisons to a control group show positive effects. One study comparing two experimental conditions on insight shows that the genre of the writing task (story vs. column) may determine learning (medium effect). Two studies without comparison groups provide some evidence that making a choice between topics (28) or between two genres of writing (29) may enhance topic knowledge.

CONCLUSIONS

Table 3 compares the experimental effects of the four types of instruction. In the second column, for each type of instruction the number of studies comparing experimental to control groups is shown. In the next column, the total number of comparisons of experimental to control groups in these studies is presented, followed by a column showing the number of comparisons with positive effects. The percentages in the fifth column are computed by dividing the number of comparisons with positive effects by the total number of comparisons. The remaining three columns give the numbers of positive effects for each of the three types of post-tests that are used in the studies and their effect sizes. The comparisons are used in 24 studies. Table 3 involves experimental comparisons only, because these can provide a fairly strong basis for formulating conclusions about effects of the four types of instruction.

TABLE 3 Positive effects of four types of instruction in experimental comparisons with control groups

Type of instruction	Number of studies	Number of comparisons with control group	Number of comparisons with positive effects	Percentage of comparisons with positive effects	On combined measure of insight and topic knowledge	On separate measure of insight	On separate measure of topic knowledge
Forward Search	3 (of 6[a])	9	6	67%	1L[b]	3L 1M	1M
Genre Writing	11 (of 14[a])	17	11	65%	2L 2M 1S	4L 2M	–
Backward Search	5 (of 14[a])	18	9	50%	1M 1S	4L 1S	2S
Planning Only	5 (of 9[a])	11	8	73%	2M	2L 1M	3L

[a]All reviewed studies applying this type of instruction.

[b]Small (S), Medium (M) and Large (L) effect sizes.

It appears from Table 3 that all four types of instruction may lead to positive, mainly large and medium effects on insight and topic knowledge. However, in about one third of the cases, there are no positive effects on learning. The number of studies showing effects varies per type of instruction.

The results of Forward Search are based on a relatively small number of studies (3). Therefore, this type of instruction can be considered as weakly supported by experimental evidence.

When the results of the three types of Genre Writing instruction are taken together, the percentage of positive effects on learning is 65% (six large, four medium and one small effect size), which is nearly two thirds of the experimental comparisons. Given the number of studies (11), the support for Genre Writing instruction as a means to stimulate learning is substantial. In 11 studies on Genre Writing instruction, 11 of 17 comparisons with a control group led to positive effects on insight and combined measures of insight and topic knowledge. Of the three types of Genre Writing, Genre Writing instruction complemented with planning, and Genre Writing complemented with planning and/or reviewing led to more positive effects on insight and topic knowledge than instruction in genre knowledge only.

According to Table 3, five experimental studies testing Backward Search instruction, show that 9 out of 18 comparisons with a control group (50%) led to positive effects on insight and/or topic knowledge (four large, one medium, four small effects). This means that experimental evidence for Backward Search instruction for stimulating writing-to-learn, is rather weak.

Table 3 shows that five studies applying Planning Only instruction comprise eleven comparisons to a control group, of which eight lead to positive effects on insight and/or topic knowledge (five large, three medium effects). Although the percentage of positive effects on learning (73%) is quite high, the number of experimental studies comparing to a control group is relatively small (5). Therefore, we regard support for Planning Only instruction for writing-to-learn as inconclusive.

Additional to the results of the experimental comparisons, results of case studies and studies comparing experimental conditions are of importance, because these studies provide alternative options for how to stimulate reflection by instruction, as described in the Results section. Examples are: using reviewing twice in Forward Search, or adding specific elements to instruction, such as a motivator (utility prompt) in Genre Writing.

Regarding the disciplines involved in the reviewed studies, it appears that writing-to-learn activities can be applied in a large variety of disciplines. The largest part of the studies (23) is directed at science, comprising seven different disciplines. The remaining studies (20) are performed in seven other disciplines belonging to humanities, social studies, behavioural sciences and earth sciences. Concerning the grades, most studies are directed at secondary education (grades 7–12). Much less studies are conducted in higher education, most of which with undergraduate students. The number of studies found in primary education is the smallest (grades 5 and 6).

In the present review, we found that we can classify instruction in studies on writing-to-learn by means of the four types of instruction that have been distinguished. Future research can use these types for characterising and comparing treatments directed at writing-to-learn. This may improve the theoretical and practical use of the distinction between the four types of instruction. In the discussion section below, this idea is elaborated.

DISCUSSION

The results of the present study provide evidence for positive effects of the four types of instruction on learning by writing, though not to the same extent, as explained above. The

results must be considered with caution, because the number of studies for each type is quite small and not all designs are equally strong. As can be seen in Table 1, the sample sizes of the 24 (quasi) experimental studies vary between 32 and 1031 participants, giving quite different weights to each of the studies. Another methodological difference is whether randomisation is used (once in the three studies on Forward Search, eight times in the eleven studies on Genre Writing, twice in five studies on Backward Search, and four times in five studies on Planning Only). In three of the latter, individual randomisation was applied, just as in two studies on Genre Writing. The remaining random assignments were on the level of the group. Additionally, all of the reviewed studies took place in regular classrooms and therefore are dependent on practical issues—for instance, changes in classroom scheduling, rules for testing, or dropout of students and teachers. Such events may have had an influence on the integrity of the designs and results of the reviewed studies.

Mechanisms underlying writing-to-learn

In this section, we provide explanations for our conclusions about the effects of the four types of instruction, by hypothesising about the mechanisms underlying the writing-to-learn process.

Mechanisms underlying forward search instruction

Forward Search instruction requires students to write down all ideas they can think of, and after finishing a first draft, feedback is provided and instruction on revising is given. Galbraith (1999) considers the writing of a first draft as externalising students' knowledge as it is represented in their mind, providing the possibility to reread and reconsider one's own knowledge. This written display of knowledge is supposed to initiate a loop of rereading, feedback and revising directed at recognising and acquiring new insights into conceptual relations and at accommodating rhetorical demands.

To stimulate this constitution of new knowledge as Galbraith (1999) calls it, feedback followed by revision is brought into the writing process by instruction. Klein (1999) and Galbraith and Torrance (2004) consider revision crucial for acquiring knowledge. Expert writers' focus for revision is on the meaning of their texts and on coherence (Klein, 1999), or on identifying potential new insights (Galbraith & Torrance, 2004). Foci for feedback in Forward Search instruction can be derived from these descriptions. Such foci are relating concepts, making inferences, organising content and identifying new insights. These foci stimulate writers to reread their draft critically by asking themselves questions (for instance, did I draw the right inference?) and to revise it.

Wallace et al. (2007) propose that feedback should also pay attention to rhetorical aspects of texts. They argue that writing should be directed at an audience, which has consequences for how ideas are being formulated. When writing for their teacher, students may just write what they know, but writing for an audience requires them to view their topic from their audience's perspective. Then, they have to ask questions such as: what knowledge does my audience have? Which information is appropriate? When revising using feedback focused on such rhetorical aspects, students have to think about how to formulate their ideas. This reformulation may lead to a new perspective and may therefore result in new insights into subject matter (Prain, 2006).

The assumed mechanisms of Forward Search may have been activated in three studies, because three of six studies which contain a comparison to a control group (six out of nine comparisons) provide positive evidence for the idea that Forward Search instruction leads

to learning. Thus, the evidence on the effects of Forward Search instruction on writing-to-learn with only six empirical studies is quite meagre. Therefore, more experimental studies are required to decide whether this type of instruction is generally effective in stimulating the writing-to-learn process.

Mechanisms underlying genre writing instruction

Genre Writing instruction consists of an explanation of the nature of a genre as preparation for writing. According to Newell (1984) writers select the rhetorical goal belonging to the genre in which they are writing. Furthermore, they use the genre structure for constructing relations between ideas and for selecting content. The constraints that a genre poses force writers to rethink the order of their ideas and the relations between them, and to look at them in new ways, which may result in new insights into the topic of their writing.

For writing in a particular genre, writers need specific genre knowledge. This means that students have to understand the macro-structure of the genre, but they also have to know what type of vocabulary and sentence structure can be used. Because it is not self-evident that students dispose of sufficient genre knowledge (Schleppegrell, 2004), studies classified as Genre Writing instruction provide genre knowledge in preparatory lessons to scaffold students before entering the writing process. Additional support can be given by instruction on planning using the structure of the genre as a coat rack, and by instruction on revising based on feedback that is focused on specific genre characteristics, such as vocabulary, register and sentence structure.

The Genre Writing studies encountered in our review were of three types: 'plain' Genre Writing instruction directed at genre knowledge in advance of writing, Genre Writing added with instruction in planning, and Genre Writing added with instruction in planning and/or reviewing. Results show that the second and third type led to more positive effects on insight and topic knowledge than 'plain' Genre Writing instruction. It therefore seems that additional instruction directed at planning or additional instruction directed at planning and/or revising reinforces the effects of 'plain' Genre Writing instruction on students' insight and topic knowledge. The number of experimental studies directed at Genre Writing instruction using a control group (11) is large in comparison to the other types of instruction in our review. Therefore, the positive evidence that Genre Writing instruction (especially the second and third type) is effective in inciting the process of writing-to-learn in students is relatively strong.

Mechanisms underlying backward search instruction

Backward search instruction consists of instruction in planning as well as reviewing. When reviewing, writers compare their contents to their rhetorical goals and ask themselves whether they have succeeded in their original (planned) intentions. Klein (1999) bases his Backward Search hypothesis on the model for expert writing of Bereiter and Scardamalia (1987). According to this model, writing is a knowledge transforming process, in which new insights arise. Students set rhetorical goals, which serve the interest and knowledge of their audiences. In addition, they generate content from their conceptual knowledge or external sources. There is a constant exchange of ideas between the content space and rhetorical space in order to arrive at a better fit between the two. When writers realise that their ideas from the content space and the rhetorical space do not match, they will adjust their ideas from the content space, or they will reconsider their rhetorical goals. This matching process leads to new insights and topic knowledge.

In accordance with this theoretical mechanism, Backward Search instruction demands from students to perform planning as well as revising activities based on feedback. This sequence of planning, formulating, feedback and revising in Backward Search instruction is intended to stimulate the recursive interaction between rhetorical and content space while writing.

In our review, 9 of 18 experimental comparisons to control groups in five studies show positive effects on insight and topic knowledge, or on insight only. We, therefore, have to conclude that empirical support for the beneficial effects of Backward Search instruction on learning is not convincing. There are relatively few studies allowing comparisons with a control group, and only half of these comparisons show positive effects on topic knowledge or insight.

The remainder of the studies into effects of Backward Search instruction consists of comparisons between experimental conditions (seven studies) and of small case studies (two). Although these studies are useful for optimising conditions for Backward Search instruction, such as the type of audience providing feedback or the focus of feedback provided (see Table 2), they are not suited for evaluating the effects of Backward Search instruction and therefore do not provide evidence for the validity of the hypothesised process underlying writing-to-learn. Given the few studies providing evidence on the effects of Backward Search instruction, there is a need for future studies containing experimental comparisons of Backward Search instruction to control groups on their effect on students' insight and topic knowledge.

Mechanisms underlying planning only instruction

Planning Only instruction entails instruction on planning activities only. Therefore, it can be seen as the opposite of Forward Search instruction that focuses on revision only. Langer and Applebee (1987) state that manipulating contents in various ways will contribute to learning. Wallace et al. (2004) regard the performance of various planning activities by the writer as a way to arrive at insight. These activities may, for instance, be weighing which contents will be part of the text, exchanging ideas with other persons and making notes to organise contents. Planning activities may also include students comparing their own selection of contents with their peers', which may urge them to adapt their own. By carrying out these types of activities, students may reconsider their selection of content elements for their text to be written. When adapting and reconsidering their planning, new insights and topic knowledge may arise. Galbraith (2015) also states that planning activities may lead to new insights, in case writers negotiate contents intensively.

In the reviewed studies, Planning Only instruction is applied by providing students with a checklist (Science Writing Heuristic) for performing various laboratory activities, or by demanding students to plan in peer groups by brainstorming, elaborating and evaluating their ideas. Eight of eleven experimental comparisons to control groups for Planning Only instruction show positive effects on insight and topic knowledge. This may seem quite substantial support for the idea that Planning Only instruction leads to adapting and reconsidering contents and therefore results in writing-to-learn. However, given the fact that only five experimental studies are involved, more experimental research is needed for evaluating the effects of this type of writing instruction.

Suggestions for future research

We have to consider the fact that research into instruction directed at writing-to-learn is of a quite recent date. As remarked previously, it was only after Klein's (1999) seminal article about cognitive processes underlying learning by writing that research turned to the types of instruction that are needed for triggering this process with students of different ages and courses. Therefore, it is not surprising that empirical evidence about the effectiveness of the different types of instruction is still inconclusive. At present, conclusions about the effects of the four types of instruction on learning should be drawn with great care, as results so far are quite mixed and based on relatively few studies that offer hard experimental evidence.

Even for Genre Writing instruction, which is the most studied type of the four distinguished, not more than 11 studies offer experimental comparisons with control groups, taking into account that this group of studies is diversified over subgroups (with planning, or with planning and/or revision). For that reason, it is important that future studies are carried out into each of the types of instruction using strong experimental designs and with a strong theoretical basis, elucidating the assumed underlying processes of writing-to-learn. For example, what is the role of feedback in these processes and how is that role enacted in the classroom context?

The role of instruction on writing-to-learn *in the disciplines* is central in this review. Therefore, we selected studies embedded in classroom contexts. This selection resulted in a large number of different disciplines, for instance: science, humanities, social and behavioral studies. However, some disciplines are missing. Most striking is the absence of mathematics. Although writing-to-learn studies are conducted in mathematics, they do not meet our criteria for inclusion. Either they are conducted before 1999 (Crocker, 1992; Kasparek, 1993), or they are not focused on effects of writing-to-learn instruction on students' insight and topic knowledge, but for instance on the teachers' skills to implement writing-to-learn in math class (Akkus & Hand, 2011; Eaton & Wade, 2014; Kenney et al., 2014). Why are effects of instruction in writing-to-learn barely studied in mathematics? Our explanation is that this may be related to the attitude of teachers in this discipline. Linnemann and Stephany (2014) mention that math teachers hardly reflect on the possible yield of applying writing for learning in their teaching practice, and therefore do not use writing assignments. Furthermore, Brozo and Crain (2018) state that mathematics teachers do not consider writing as relevant for mathematics learning, while Teuscher et al. (2015) report that many math teachers view writing-to-learn activities as consuming too much time. These observations may also be an explanation for researchers' focus on the teacher in studies on mathematics. We suggest research into the question whether instruction in writing-to-learn in mathematics classes leads to positive effects on learning. It is worthwhile to find out whether writing-to-learn in this discipline is an effective way of learning.

Evidence for the process of writing-to-learn is shown only indirectly in the reviewed studies. When instruction leads to more insight and topic knowledge, a mediating effect of the process of writing-to-learn is assumed. Above, we discussed the mechanisms underlying the process of writing-to-learn and determined whether the presence of these mechanisms can be deduced from the found effects of the four types of instruction. However, more insight into the occurrence of these mechanisms is needed. Think aloud studies on the thinking process of individual students while carrying out writing-to-learn tasks under different conditions may offer interesting clues about how learning of topic knowledge and acquiring new insight is brought about by the cognitive processes involved in writing. More of this type of scientific knowledge about the process of writing-to-learn is a valuable source for improving instruction in the use of writing as a learning tool.

Implications for education

The present review discriminates four types of instruction (Forward Search, Genre Writing, Backward Search and Planning Only), all of which may result in increased insight and topic knowledge of students. Although evidence is not as strong as we may wish for, we will explore the question how knowledge acquired in our review can be implemented in education.

The reviewed studies show various ways to apply the types of instruction, such as working with a checklist for planning, or giving feedback on comprehensibility for an audience of readers. These types of instruction can be applied in education as an alternative for more traditional approaches to learning. When teachers understand the basic assumptions underlying each of the four instructional approaches (the so-called mechanisms of writing-to-learn), they may experiment in their classes with writing assignments and instructional support. Then, teachers can find out how the various types function in class. Some students, for instance low self-monitors, may be better served with free writing, guided by no more than a subject to write about (Forward Search instruction), for activating their prior knowledge. Other students may need structured assignments and benefit more from precise direction of their writing process (which is the case with Genre Writing, Backward Search and Planning Only instruction), for instance when writing a lab report.

Not every teacher interested in writing-to-learn, and understanding the mechanisms, will find it easy to provide good instructions. Baker et al. (2008) report teachers' problems with the implementation of writing-to-learn in class. In interviews with the researchers, these (science) teachers state that they are no language teachers and do not know how to evaluate students' texts and how to give them feedback. Furthermore, they are concerned about the amount of time writing-to-learn needs.

When asked for solutions, interviewees suggest that these teachers in disciplines such as science or social studies work together with language art teachers for experimenting with writing-to-learn in their programmes. Such teams can discuss the design of writing tasks, and appropriate types of instruction (Baker et al., 2008). Furthermore, they can compose writing assignments including instruction. Language art teachers can explain that students need concrete directions for writing, meaning that they need to know the goals and the audience they write for, as well as the genre of the text to be written. Instruction in planning (Backward Search, Genre Writing) can be facilitated by choosing an existing checklist (such as the Science Writing Heuristic), of which this review shows positive effects on learning. Implementing Genre Writing instruction requires collecting or designing appropriate model texts. Language art teachers can propose criteria for such texts (e.g., text difficulty, vocabulary, text structure) and provide concrete examples of model texts in a genre (e.g., explanations, arguments, journals etc.).

Teachers will have to prepare for coaching the writing process, particularly for giving feedback. Bean (2011) suggests that the application of a web-based reciprocal peer review system stimulates students' self-support and saves time. Students upload their texts in the system, give feedback to each other, and evaluate their peers' feedback, by means of rubrics. In the end, teachers can add their feedback. The language art teachers and teachers of other subjects can compose the rubrics. These should be compatible with the goals of the writing task, directed at topic knowledge of and insight into subject matter. This type of experimenting with writing-to-learn by collaborating in teams may give teachers more confidence in applying this new way of learning in class.

Positive effects of the four types of instruction that are found in this review give a glimpse of the hypothetical underlying mechanisms of writing-to-learn we have described. This is a promising start, but more experimental evidence for how different types of instruction

stimulate the process of writing-to-learn is needed, because the number of studies providing evidence hitherto is small.

CONFLICT OF INTEREST

There are no potential conflicts of interest for the authors.

ETHICS APPROVAL

Ethics approval was not required for this work.

DATA AVAILABILITY STATEMENT

All the data of this study are presented in the tables of this article.

ORCID

Aartje van Dijk https://orcid.org/0000-0002-4985-8417

REFERENCES

*These titles were used in the review analysis (Table 1).

Ackerman, J. M. (1993). The promise of writing to learn. *Written Communication*, *10*, 334–370. https://doi.org/10.1177/0741088393010003002

Akkus, R., & Hand, B. (2011). Examining teachers' struggles as they attempt to implement dialogical interaction as part of promoting mathematical reasoning within their classrooms. *International Journal of Science and Mathematics Education*, *9*, 975–998.

Applebee, A. N. (1984). Writing and reasoning. *Review of Educational Research*, *54*, 577–596. https://doi.org/10.3102/00346543054004577

*Atasoy, Ş., & Küçük, O. (2020). Development of eighth grade students' epistemological beliefs through writing-to-learn activities. *Journal of Science Learning*, *3*(2), 57–66.

Baker, W. P., Barstack, R., Clark, D., Hull, E., Goodman, B., Kook, J., Kraft, K., Ramakrishna, P., Roberts, E., Shaw, J., Weaver, D., & Lang, M. (2008). Writing-to-learn in the inquiry classroom: Effective strategies from middle school science and writing teachers. *The Clearing House*, *81*, 105–108. https://doi.org/10.3200/TCHS.81.3.105-108

*Balgopal, M. M., Wallace, A. M., & Dahlberg, S. (2012). Writing to learn ecology: A study of three populations of college students. *Environmental Education Research*, *18*, 67–90. https://doi.org/10.1080/13504622.2011.576316

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: a meta-analysis. *Review of Educational Research*, *74*, 29–58. https://doi.org/10.3102/00346543074001029

Bean, J. C. (2011). *Engaging ideas: The professor's guide to integrating writing, thinking and active learning in the classroom*. Jossey-Bass.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Erlbaum.

*Boscolo, P., & Carotti, L. (2003). Does writing contribute to improving high school students' approach to literature? *Educational Studies in Language and Literature*, *3*, 197–224. https://doi.org/10.1023/B:ESLL.0000003685.30925.c4

Boyles, M. P., Killian, P. W., & Rileigh, K. K. (1994). Learning by writing in introductory psychology. *Psychological Reports*, *75*, 563–568. https://doi.org/10.2466/pr0.1994.75.1.563

Britton, J. (1982). Shaping at the point of utterance. In G. M. Pradl (Ed.), *Prospect and retro prospect: Selected essays of James Britton* (pp. 123–129). Boynton/Cook Publishers.

Brozo, W. G., & Crain, S. (2018). Writing in math: A disciplinary literacy approach. *The Clearing House*, *91*(1), 7–13. https://doi.org/10.1080/00098655.2017.1342435

*Corcelles Seuba, M., & Castelló, M. (2015). Learning philosophical thinking through collaborative writing in secondary education. *Journal of Writing Research*, *7*(1), 157–200 10.17239/jowr-2015.07.01.07

Crocker, D. A. (1992). Educational reflections. *The Amatic Review*, *13*(2), 66–68. https://doi.org/10.1177/174114329202000201

Eaton, C. D., & Wade, S. S. (2014). Collaborative learning through formative peer review with technology. *Primus*, *24*(6), 529–543. https://doi.org/10.1080/10511970.2014.881442

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, *28*, 122–128. https://doi.org/10.2307/356095

*Finkenstaedt-Quinn, S. A., Halim, A. S., Chambers, T. G., Moon, A., Goldman, R. S., Ruggles Gere, A., & Shultz, G. V. (2017). Investigation of the influence of a writing-to-learn assignment on student

understanding of polymer Properties. *Journal of Chemical Education*, *94*, 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363
Galbraith, D. (1992). Conditions for discovery through writing. *Instructional Science*, *21*, 45–72. https://doi.org/10.1007/BF00119655
Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 139–160). Amsterdam University Press.
Galbraith, D. (2009). Writing as discovery. *British Journal of Educational Psychology Monograph, Series II*, *6*, 5–26. https://doi.org/10.1348/978185409X421129
Galbraith, D. (2015). Conditions for writing to learn. *Journal for Writing Research*, *7*(1), 215–266 10.17239/jowr-2015.07.01.09
Galbraith, D., & Torrance, M. (2004). Revision in the context of different drafting strategies. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 63–86). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-007-1048-1_5
Graham, S., Kiuhara, S. A., & McKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research*, *90*(2), 179226. https://doi.org/10.3102/0034654320914744
*Granado-Peinado, M., Mateos, M., Martin, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*, 2037–2058. https://doi.org/10.1007/s11145-019-09939-6
*Gunel, M., Hand, B., & Gunduz, S. (2006). Comparing student understanding of quantum physics when embedding multimodal representations into two different writing formats: Presentation format versus summary report format. *Wiley InterScience*, *1092*, 1112. https://doi.org/10.1002/sce.20160
*Gunel, M., Hand, B., & McDermott, M. A. (2009). Writing for different audiences: Effects on high-school students' conceptual understanding of biology. *Learning and Instruction*, *19*, 354–376. https://doi.org/10.1016/j.learninstruc.2008.07.001
Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. Routledge.
Hand, B., Chen, Y., & Suh, J. K. (2021). Does a knowledge generation approach to learning benefit students?: A systematic review of research on the science writing heuristic approach. *Educational Psychology Review*, *33*, 535–577. https://doi.org/10.1007/s10648-020-09550-0
*Hand, B., Gunel, M., & Ulu, C. (2009). Sequencing embedded multimodal representations in a writing to learn approach to the teaching of electricity. *Journal of Research in Science Teaching*, *46*(3), 225–247. https://doi.org/10.1002/tea.20282
*Hand, B., Hohenshell, L., & Prain, V. (2004). Exploring students' responses to conceptual questions when engaged with planned writing experiences: A study with year 10 science students. *Journal of Research in Science Teaching*, *41*(2), 186–210. https://doi.org/10.1002/tea.10128
*Hand, B., Hohenshell, L., & Prain, V. (2007). Examining the effect of multiple writing tasks on year 10 biology students' understanding of cell and molecular biology concepts. *Instructional Science*, *35*, 343–373. https://doi.org/10.1007/s11251-006-9012-3
*Hand, B., Wallace, C. W., & Yang, E. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects. *International Journal of Science Education*, *26*(2), 131–149. https://doi.org/10.1080/0950069032000070252
*Hand, B., Yang, O. E., & Bruxvoort, C. (2007). Using writing-to-learn science strategies to improve year 11 students' understanding of stoichiometry. *International Journal of Science and Mathematics Education*, *5*, 125–143. https://doi.org/10.1007/s10763-006-9029-8
Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–27). Erlbaum.
*Hohenshell, L. M., & Hand, B. (2006). Writing-to-learn strategies in secondary school cell biology: A mixed method study. *International Journal of Science Education*, *28*(2-3), 261–289. https://doi.org/10.1080/09500690500336965
*Hunter, K., & Tse, H. (2013). Making disciplinary writing and thinking practices an integral part of academic content teaching. *Active Learning in Higher Education*, *14*(3), 227–239. https://doi.org/10.1177/1469787413498037
*Kabataş Memiş, E., & Öz, M. (2017). The impact of using representation modes within writing to learn activities on the scientific process skills of the fifth-grade students. *Journal of Education and Training Studies*, *5*(2), 31–42 10.11114/jets.v5i2.2001
Kasparek, R. F. (1993). Effects of integrated writing on attitude and algebra performance of high school students. Doctoral dissertation. *Abstracts International*, 54(08), 2931A. (UMI No. 9402483).
Kenney, R., Shoffner, M., & Norris, D. (2014). Reflecting on the use of writing to promote mathematical learning: An examination of preservice mathematics teachers' perspectives. *The Teacher Educator*, *49*(1), 28–43. https://doi.org/10.1080/08878730.2013.848002

*Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2006). Writing as a learning tool: Testing the role of students' writing strategies. *European Journal of Psychology of Education, XXI*, *1*, 17–34.
*Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude treatment interaction approach to writing-to-learn. *Learning and Instruction*, *18*, 379–390. https://doi.org/10.1016/j.learninstruc.2007.07.004
Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, *11*(3), 203–270. https://doi.org/10.1023/A:1021913217147
Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research*, *7*(3), 311–350 10.17239/jowr-2016.07.03.01
*Klein, P. D., & Ehrhardt, J. S. (2015). The effects of discussion and persuasion writing goals on reasoning, cognitive load, and learning. *Alberta Journal for Educational Research*, *61*(1), 40–64.
*Klein, P. D., & Kirkpatrick, L. C. (2010). A framework for content area writing: Mediators and moderators. *Journal of Writing Research*, *2*(1), 1–46 10.17239/jowr-2010.02.01.1
*Klein, P. D., & Rose, M. A. (2010). Teaching argument and explanation to prepare junior students for writing to learn. *Reading Research Quarterly*, *45*(4), 433–461. https://doi.org/10.1598/RRQ.45.4.4
*Klein, P. D., & Samuels, B. (2010). Learning about plate tectonics through argument writing. *The Alberta Journal of Educational Research*, *56*(2), 196–217.
Langer, J.A., & Applebee, A.N. (1987). *How writing shapes thinking: A study of teaching and learning*. NCTE.
Linnemann, M., & Stephany, S. (2014). Supportive writing assignments for less skilled writers in the mathematics classroom. In G. Rijlaarsdam (Series Ed.), P. D. Klein, P. Boscolo, L. C. Kirkpatrick, & C. Gelati (Vol. Eds.), *Studies in writing: Vol. 28, Writing as a learning activity*, 66–93. Brill. 10.1163/9789004265011_005
*Martinez, I., Mateos, M., Martin, E., & Rijlaarsdam, G. (2015). Learning history by composing synthesis texts: Effects of an instructional program on learning, reading and writing processes, and text quality. *Journal of Writing Research*, *7*(2), 275–302 10.17239/jowr-2015.07.02.03
Mateos, M., Martin, E., Cuevas, I., Villalón, R., Martinez, I., & González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction*, *36*(2), 119–138. https://doi.org/10.1080/07370008.2018.1425300
*McDermott, M. A., & Hand, B. (2013). The impact of embedding multiple modes of representation within writing tasks on high school students' chemistry understanding. *Instructional Science*, *41*, 217–246. https://doi.org/10.1007/s11251-012-9225-6
Miller, D. M., Scott, C. E., & McTigue, M. (2018). Writing in the secondary level disciplines: A systematic review of context, cognition and content. *Educational Psychology Review*, *30*, 83–120. https://doi.org/10.1007/s10648-016-9393-z
*Nam, J., Choi, A., & Hand, B. (2011). Implementation of the science writing heuristic (SWH) approach in 8th grade science classrooms. *International Journal of Science and Mathematics Education*, *9*, 1111–1133. https://doi.org/10.1007/s10763-010-9250-3
Nevid, J. S., Ambrose, M. A., & Pyun, Y. S. (2017). Effects of higher and lower level writing to learn assignments on higher and lower level examination questions. *Teaching of Psychology*, *44*(4), 324–329. https://doi.org/10.1177/0098628317727645
*Nevid, J. S., Pastva, A., & McClelland, N. (2012). Writing-to-learn assignments in introductory psychology: Is there a learning benefit? *Teaching of Psychology*, *39*(4), 272–275. https://doi.org/10.1177/0098628312456622
Newell, G. E. (1984). Learning from writing in two content areas: A case study/protocol analysis. *Research in the Teaching of English*, *18*(3), 265–287.
*Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, *38*, 237–258. https://doi.org/10.1007/s11251-009-9106-9
Page, M.J., et al. (2020). The Prisma 2020 statement: an updated guideline for reporting systematic reviews. https://doi.org/10.1136/bmj.n71
Prain, V. (2006). Learning from writing in secondary science: Some theoretical and practical implications. *International Journal of Science Education.*, *28*(2–3), 179–201. https://doi.org/10.1080/09500690500336643
*Ritchie, S. M., Tomas, L., & Tones, M. (2011). Writing stories to enhance scientific literacy. *International Journal of Science Education*, *33*(5), 685–707. https://doi.org/10.1080/09500691003728039
Rivard, L. P. (1994). A Review of writing to learn in science: Implications for practice and research. *Journal of Research in Science Teaching*, *31*(9), 969–983. https://doi.org/10.1002/tea.3660310910
*Rivard, L. P. (2004). Are language-based activities in science effective for all students, including low achievers? *Science Education*, *88*, 420–442. https://doi.org/10.1002/sce.10114
*Rivard, L. P., & Straw, S. B. (2000). The effect of talk and writing on learning science: An exploratory study. *Science Education*, *84*, 566–593. https://doi.org/10.1002/1098-237X(200009)84:5<566::AID-SCE2>3.0.CO;2-U
Rose, D. (2008). Writing as a linguistic mastery: The development of genre-based literacy pedagogy. In D. Myhill, D. Beard, M. Nystrand, & J. Riley (Eds.), *2008*. Sage.

*Sampson, V., & Walker, J. P. (2012). Argument-Driven inquiry as a way to help undergraduate students write to learn by learning to write in chemistry. *International Journal of Science Education*, *34*(10), 1443–1485. https://doi.org/10.1080/09500693.2012.667581

Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Erlbaum. https://doi.org/10.4324/9781410610317

*Stewart, T. L., Myers, A. C., & Culley, M. R. (2010). Enhanced learning and retention through "writing to learn" in the psychology classroom. *Teaching of Psychology*, *37*, 46–49. https://doi.org/10.1080/00986280903425813

Teuscher, D., Hodges Kulinna, P., & Crooker, C. (2015). Writing to learn mathematics: An update. *The Mathematics Educator*, *24*(2), 56–78.

*Van Drie, J., Braaksma, M., & Van Boxtel, C. (2015). Writing in history: Effects of writing instruction on historical reasoning and text quality. *Journal of Writing Research*, *7*(1), 123–156 10.17239/jowr-2015.07.01.06

Van Gelderen, A. (1997). Elementary students' skills in revising: Integrating quantitative and qualitative analysis. *Written Communication*, *14*(3), 360–397. https://doi.org/10.1177/0741088397014003003

Wallace, C. S., Hand, B., & Prain, V. (2007). *Writing and learning in the science classroom*. Springer.

*Wäschle, K., Gebhardt, A., Oberbusch, E. M., & Nückles, M. (2015). Journal writing in science: Effects on comprehension, interest, and critical reflection. *Journal of Writing Research*, *7*(1), 41–64. https://doi.org/10.17239/jowr-2015.07.01.03

*Winstead Fry, S., & Villagomez, A. (2012). Writing to learn benefits and limitations. *College Teaching*, *60*, 170–175. https://doi.org/10.1080/87567555.2012.697081

*Wissinger, D. R., & De La Paz, S. (2016). Effects of critical discussions on middle school students' written historical arguments. *Journal of Educational Psychology*, *108*(1), 43–59. https://doi.org/10.1037/edu0000043

*Wong, B. Y. L., Kuperis, S., Jamieson, D., Keller, L., & Cull-Hewitt, R. (2002). Effects of guided journal writing on students' story understanding. *The Journal of Educational Research*, *95*(3), 179–191. https://doi.org/10.1080/00220670209596588

*Yildiz, E., & Akdag, S. (2021). The effects of cooperative learning and writing to learn applications on academic achievement. *International Journal of Progressive Education*, *17*(1), 196–209. https://doi.org/10.29329/ijpe.2020.329.13

How to cite this article: van Dijk, A., van Gelderen, A., & Kuiken, F. (2022). Which types of instruction in writing-to-learn lead to insight and topic knowledge in different disciplines? A review of empirical studies. *Review of Education*, *10*, e3359. https://doi.org/10.1002/rev3.3359