How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

Model texts as a feedback instrument in second language writing: A systematic review

Long Quoc Nguyen*

University of Foreign Languages and International Studies, Hue University, 57 Nguyen Khoa Chiem, Hue, Vietnam

FPT University, Ho Chi Minh City, Vietnam

*Correspondence: nlquoc.dhnn23@hueuni.edu.vn

OCRID: https://orcid.org/0000-0002-4674-7199

Phung Dao

Faculty of Education, University of Cambridge, 184 Hills Road, Cambridge CB2 8PQ, UK

Email: pd543@cam.ac.uk

OCRID: https://orcid.org/0000-0002-8612-5589

Bao Trang Thi Nguyen

University of Foreign Languages and International Studies, Hue University, 57 Nguyen Khoa Chiem, Hue, Vietnam

Email: ntbtrang@hueuni.edu.vn

OCRID: https://orcid.org/0000-0002-5285-7255

Funding: This study has not been supported by any funding agency or institution.

Availability of data and materials : All the data used in this study belong to the authors and will be shared upon reasonable request.

Declarations : The authors declare no competing interests.

Model texts as a feedback instrument in second language writing: A systematic review

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

Abstract

The potential of model texts as a feedback instrument (MTFI) in second language (L2) research has been explored for about two decades and continues to receive increasing interest from many researchers. However, to date, there is still an absence of a comprehensive review of studies in this particular research field. The present study aims to fill in this gap by systematically reviewing a total of 25 papers on MTFI dating up to the first of April 2023 with regard to three main focal points: i) study contexts, ii) methodological characteristics, and iii) primary areas of focus as well as key findings. Our dataset revealed salient insights into research on MTFI: a) primary attention to young English as a Foreign Language (EFL) learners at primary schools, b) a wide adoption of three-stage narrative writing tasks performed individually in a classroom-based setting following a quasi-experimental design, c) a predominant focus on learners' noticing and incorporations of features in model texts, and d) consistent reports of the potential of MTFI in facilitating L2 writing, especially in lexical gains. This review enhances our understanding of MTFI and entails important implications for pedagogy. It also offers ways to address current issues in the selected studies and provides empirically established suggestions for advancing this research domain.

Keywords: models, second language writing, feedback, systematic review, issues, suggestions

Introduction

Model texts, also known as ‘comparators’ (Lynch, 2009, p. 199), are referred to as compositions written by competent users of a language (Hanaoka, 2007; Luquin & García Mayo, 2021). In a more detailed manner, Cánovas Guirao et al. (2015, p. 65) defined models as ‘native or native-like texts, which learners compare with their original writing [...] bearing in mind the content and the genre of the composition, as well as learners' age, proficiency level’. In early

research, models were often given to students at the outset of a writing session before they composed their own draft (Cánovas Guirao et al., 2015; Hamano-Bunce, 2022). This practice, however, has been criticized as it often led to superficial imitation rather than learning (Ferris, 2010). It is also likely to result in a low level of cognitive processing and problem solving (Cánovas Guirao et al., 2015). In more recent research, models are presented to learners only after they have completed their own writing. The main aim is to promote comparisons between learners' draft and model texts, which might then facilitate learners' uptake of new L2 knowledge while limiting the possibility of students' copying from these comparators (Hanaoka & Izumi, 2012; García Mayo & Loidi Labandibar, 2017). Studies that provide model texts after learners have finished the first draft often follow three primary stages: composing, comparing, and rewriting. That is, students are initially tasked with composing their own draft in response to a given prompt and noting down any problems they may have (stage one, which serves as a pre-test). Subsequently, they are presented with one or two model texts with which they can compare their writing (stage two). Finally, they are asked to rewrite their original draft (stage three, which is regarded as an immediate post-test).

The use of MTFI for learners' written output, as an alternative to written corrective feedback (WCF), has been under investigation for many years (e.g., Hanaoka, 2007; Kang, 2022). Despite being a young research area (Lázaro-Ibarrola, 2021), MTFI has been shown to play a facilitative role in promoting L2 development (Coyle et al., 2018; Hanaoka & Izumi, 2012). In fact, via exposure to model texts, learners could find solutions to bridge their knowledge gap or hole (Hanaoka, 2007; Hanaoka & Izumi, 2012). Moreover, MTFI can promote learners' metalinguistic awareness, foster their learning autonomy, and help them engage in deeper processing (Hamano-Bunce, 2022; Hanaoka & Izumi, 2012). It also provides them with good samples of writing not only for individual sentences but also for entire discourse (Yang & Zhang,

2010). This exposure has the potential to facilitate writing improvements (Criado et al., 2022; Kang, 2022). Additionally, MTFI enables teachers to reduce the need to provide WCF (Hamano-Bunce, 2022; Luquin & García Mayo, 2021), saving them time for additional writing practice needed for learners' L2 development as called for by Truscott (1996).

Due to its considerable potential, MTFI has attracted increasing attention from many L2 researchers for the past two decades (e.g., Coyle et al., 2018; Hanaoka, 2007; Kang, 2022; Luquin & Gárcia Mayo, 2023). As research on MTFI has diversified in many ways and continues to grow, there is a high need to conduct a systematic review of the past literature to provide a comprehensive picture of what has been done and what needs to be further explored. To date, there has been only one study that attempted to review empirical research on MTFI (Cao & Mao, 2022). Through the analysis of 11 MTFI papers, the authors found that L2 learners mostly focused on lexis during the three-stage writing tasks, and that L2 proficiency levels and affective factors may influence how learners process and benefit from models. Although the review by Cao and Mao (2022) has provided significant implications for pedagogy as well as advanced our understanding of the role of MTFI, several aspects still remain underexplored. First, the publications in Cao and Mao's review (2022) only dated up to the year 2021; a large portion of MTFI studies after this time were not included yet. The addition of recent papers would entail a more comprehensive picture of this research domain. Second, what needs to be further examined is the study contexts of MTFI research and the impact of this instructional approach on L2 learners' text quality. The synthesis of such information might well inform institutes and teachers of how to best utilize model texts as a feedback facilitator. Third, there is still an absence of a systematic review of methodologies widely employed in empirical research on MTFI. Details about these methodological features could help researchers understand how studies on MTFI have been conducted, based on which

constructive suggestions for future research can be made. Thus, there is a pressing need to have another review on MTFI to address these gaps.

Theoretical underpinnings of model texts as a feedback instrument

The use of MTFI in L2 writing is mainly anchored to three fundamental theories: the Output Hypothesis, the Noticing Hypothesis, and the Input Hypothesis. The Output Hypothesis (Swain, 1985) argues that production practice, or comprehensible output, plays a vital role in L2 acquisition. This is because learners are less likely to produce more precise language if they do not feel the need to do so. Via writing or speaking, they can realize the problems that refrain them from expressing their intended message. This pushes them to look for solutions to create their desired meaning in a more appropriate and accurate way (Swain, 1985), which is also referred to as pushed output (DeKeyser, 2007). The Output Hypothesis includes three primary functions: noticing, hypothesis testing, and metalinguistic processing, which correlate well with the Noticing Hypothesis (Schmidt, 1990, 2001) and the Input Hypothesis (Krashen, 1982).

First, output is expected to trigger two kinds of noticing: a gap and a hole (Muranoi, 2007). While a gap is considered as something they are able to produce yet in a partially correct way, a hole refers to the complete absence of knowledge in their repertoire (Izumi, 2013). Through written or oral output, learners start to recognize that there is a gap or a hole in their L2 knowledge (Swain, 1985). The role of noticing is vital in second language acquisition (SLA) (e.g., Robinson, 1995; Schmidt, 1995, 2001). This is evident via strong claims on noticing such as 'no noticing, no acquisition' (Ellis, 1995, p.89) or 'those who noticed most, learn most' (Schmidt & Frota, 1986, p. 313). In 2001, Schmidt proposed the updated version of the Noticing Hypothesis which claims that noticing, albeit not a must, can definitely facilitate language learning. This version has been widely supported and accepted by many linguists and researchers (Izumi, 2013).

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

The second function in the Output Hypothesis (Swain, 1985) is referred to as hypothesis testing. After noticing the knowledge gap or hole, learners start to search for solutions and evaluate whether these remedies are effective. They need to have some feedback against which they can test the new information. At this point, the role of the Input Hypothesis (Krashen, 1982) comes into play. Krashen (1982) posits that meaningful input and frequent input exposure are necessary for L2 learning to occur, which is embraced by other scholars (e.g., Ellis & Wulff, 2015). With comprehensible input, especially the information that is slightly beyond their current knowledge, learners can find helpful ways to bridge their knowledge gap or hole. Finally, the metalinguistic function refers to the process that learners attempt to 'use language to reflect on language' (Muranoi, 2007, p. 58), also known as *metatalk* (Swain, 1998), to think about the target language in a conscious way. Swain (1998) argues that this function can deepen learners' linguistic rules and forms, helping them to understand the relationship between forms and meanings in a context-sensitive circumstance.

In L2 writing, the Output Hypothesis, the Noticing Hypothesis, and the Input Hypothesis, when taken together, meaningfully support MTFI throughout its three main stages. In stage one, learners compose a draft on a given topic (output), and often in this stage, they start to notice the gaps or holes in their knowledge (noticing) (Manchón, 2011). In stage two, model texts are provided as a source of comprehensible input (Krashen, 1982) with which learners compare their writing (output - hypothesis testing). During the comparison, they direct their attention to the solutions provided in the models (noticing) to bridge the previous gaps (overt problems) or fill in the holes (covert problems). Finally, noticed features in the models can be incorporated into learners' rewritten texts (output - metalinguistic function) in stage three as uptake or acquisition (Hanaoka, 2007; Hanaoka & Izumi, 2012).

While these theories may also be the foundations for other kinds of feedback such as WCF, model texts are unique in that they correlate well to L2 learners' internal syllabus (Qi & Lapkin, 2001; Sachs & Polio, 2007). Indeed, learners are given the freedom to focus on what they want and need in the models. This could be more helpful for them than error corrections (Qi & Lapkin, 2001) because the foci of teachers in their feedback do not always align with what learners want to concentrate on (Long & Robinson, 1998). Furthermore, unlike WCF which is considered as negative feedback, models provide learners with positive evidence which may be necessary for them to fill in their knowledge gap or hole (Muranoi, 2007).

Despite its great potential in facilitating L2 learners' writing development and a growing body of research on this topic, there has been no review that concentrates solely on MTFI yet. The study by Cao and Mao (2022) entails important contributions, but it delved into the combination of two different kinds of feedback: MTFI and reformulations (i.e., learners' draft rewritten by an L1 speaker of the target language), not particularly on MTFI. Moreover, there are various aspects that need to be further explored: the inclusion of studies after the year 2021, the methodological design features employed in previous MTFI research, and the synthesis of emerging areas of focus as well as their findings. Thus, another systematic review is highly needed.

The current study

To fill in these gaps mentioned, this present study aims to provide a review of L2 writing research employing model texts as a feedback instrument. In particular, it synthesizes the research contexts, the methodological characteristics, and the major areas of focus as well as key findings in past MTFI studies dating up to the first of April, 2023. Therefore, three research questions (RQs) are formulated to guide this review:

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

RQ1. What are the study contexts in empirical L2 research on MTFI?

RQ2. What are the methodological characteristics of empirical L2 research on MTFI?

RQ3. What are the primary areas of focus and key findings in empirical L2 research on MTFI?

Method

Creating the report pool

To choose relevant studies for this review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) updated guidelines (Page et al., 2021), which include three stages: identification, screening, and inclusion/exclusion (Figure 2). Following Cao and Mao's (2022) review, we developed a combination of keywords (i.e., model OR comparator, AND second language OR foreign language, AND writing OR written, AND feedback). After that, we searched the keywords in four main databases: Web of Science, Scopus, Education Resources Information Center (ERIC), and ProQuest (Education Database, Research Library: Literature & Language, Research Library: Social Sciences). The search resulted in 620 records of relevant articles dating up to the first of April, 2023. We then refined the pool to 476 articles by removing 144 duplicates.

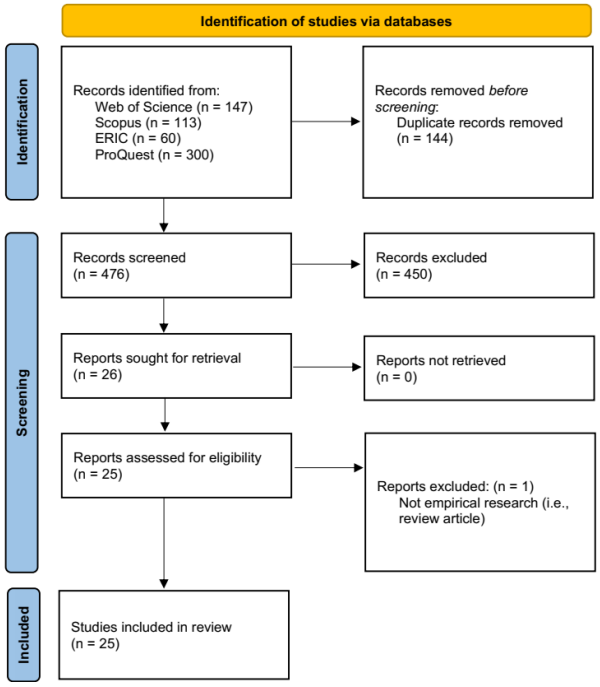
Inclusion and exclusion criteria

The pool of 476 papers were screened based on a set of inclusion and exclusion criteria. First, we scanned the title and abstract of each article and only retained those meeting four criteria: a) being published in journals listed in the Social Sciences Citation Index (SSCI) or the Emerging Sciences Citation Index (ESCI), b) being conducted in the L2 setting , c) being written in English, and d) employing model texts as a feedback instrument. The first inclusion criterion was established for quality assurance (for a detailed rationale, see Zhang, 2020) while the others were

adopted because they reflected the main foci of the current paper. Based on these criteria, we obtained a total of 26 studies for full-text screening. Second, we carefully read through each paper and decided to eliminate one review article since it did not contain any empirical data. Finally, a pool of 25 studies on MTFI were included for the review. For a full list of journals in which these papers were published, see Appendix A.

Figure 2

PRISMA flow diagram for including MTFI studies



Coding

In light of the RQs, the three researchers worked together and developed a coding scheme for the selected studies (Table 1). Accordingly, three types of information were extracted: a) study contexts, b) methodological characteristics, and c) primary areas of focus and key findings. From the selected papers, we operationalized study contexts as research places, learners' educational levels, learner types, and learners' L2 proficiency. With regard to methodological characteristics,

two main sub-categories were generated, which included general design features (research types, effect sizes in quasi-experimental studies, research settings, means of communication, number of exposures to models, piloting of models, sample sizes in quasi-experimental studies), and MTFI-related design features (writing modes, writing types, number of model texts, and writers of model texts). As for the last dimension, seven categories were generated: noticed problems in the composing stage, noticed features in the comparing stage, incorporations of features in models in the rewriting stage, relationship between noticing and incorporations, impact of MTFI on text quality, learners' attitude towards MTFI, and others.

Based on the scheme in Table 1, the first and third researchers coded four papers (16%) independently. The agreement rate for all the sub-categories (22 in total) between the two coders was 92.05%, demonstrating a high level of interrater reliability (McHugh, 2012); all disagreements were resolved via in-depth discussion. Following this, the first researcher continued to code the remainder of the studies. Next, the second researcher randomly selected and coded five out of the 21 remaining articles. The proportion of similarities between the first and the second researchers' coding was 98.18%, indicating a strong consensus (McHugh, 2012). All differences between the two coders were discussed until agreements on all aspects were reached. For details of interrater agreements, see Appendix B.

Table 1

Coding scheme

Category	Sub-category	Item
Study contexts	Research places	Africa, America, Asia, Australia, Europe
	Learners' educational levels	primary school, middle school, high school, university
	Learner types	Young learners, adults

	Learners' L2 proficiency	Low, low-intermediate, intermediate, upper-intermediate, advanced, not reported
Methodological characteristics	General design features	Research settings (classroom-based, lab-based, not reported)
		Means of communication (face-to-face, computer-mediated)
		Research types (quasi-experimental/ non-experimental)
		Piloting of models (reported, not reported)
		Effect sizes in quasi-experimental studies (reported/ not reported)
		Number of exposures to MTFI (one, multiple)
		Sample sizes in quasi-experimental studies (Mean, SD, Minimum, Maximum)
	MTFI design features	Writing modes (individual, collaborative, both)
		Writing types (narrative, descriptive, argumentative)
		Number of provided models (one, two)
		Writers of models (class teachers/ researchers/ external writers, not reported) (L1 users of the target language/ L2 users of the target language, not reported)
Primary areas of focus and key findings	Noticed problems in the composing stage	Most noticed (vocabulary, grammar, organization, content)
	Noticed features in the comparing stage	Most noticed (vocabulary, grammar, organization or content)
	Incorporations of features in models in the last stage	Most incorporated (vocabulary, grammar, organization or content)
	Relationship between noticing and incorporations	Positive, negative Stage 1 led to better incorporations than Stage 2 Stage 2 led to better incorporations than Stage 1
	Impact of MTFI on text quality	Scale-based measures, lexical complexity, syntactic complexity, accuracy, fluency (significantly improved, insignificantly improved)
	Attitude towards MTFI	Positive, negative, positive but declining

	Others (List of other foci)	List of key corresponding findings

Analysis

All of the coded data were input into an Excel Spreadsheet before converted to SPSS (Statistical Packages for the Social Sciences) version 28.0. Descriptive analysis was employed to measure the frequency and percentage of the items in each sub-category; however, as for the sub-category of sample sizes in quasi-experimental studies, values of Mean, Standard Deviation (SD), Minimum, and Maximum were used. It should be acknowledged that although the total number of the selected MTFI studies was 25, the sum of the presented values in some tables might exceed this figure. This is because some articles applied to more than one sub-category and/or item.

Results

This section presents the obtained outcomes in response to the three research questions. The results are demonstrated in tables which contain sub-categories and the corresponding statistics.

RQ1. The study contexts

Table 2

The study contexts of the MTFI studies

Sub-category	Frequency	Percentage

Research places		
Asia	9	36%
Europe	16	64%
Learners' educational levels		
Primary school	12	46.2%
Middle school	4	15.4%
High school	1	3.8%
University	9	34.6%
Learner types		
Young learners	16	64%
Adults	9	36%
Learners' L2 proficiency		
Low	6	18.7%
Low-intermediate	11	34.4%
Intermediate	10	31.3%
Upper-intermediate	3	9.4%
Advanced	1	3.1%
Not reported	1	3.1%

Table 2 demonstrates that a majority of the selected studies on MTFI were conducted in Europe (n = 16; 64%) and Asia (n = 9; 36%), solely in the EFL setting. As for learner-related factors, the main participants were young learners (n = 16; 64%) at primary schools (n = 12; 46.2%) whose proficiency levels of English were mostly low-intermediate (n = 11; 34.4%) and intermediate (n = 10; 31.3%). High-school and advanced learners accounted for very small percentages in the selected MTFI research, at under 4% each group (n = 1).

RQ2. The methodological characteristics

In terms of methodological characteristics, two main categories were examined, including general and MTFI-related design features. Details are illustrated in Tables 3 and 4 below.

Table 3

General design features in the MTFI studies

Sub-category	Frequency	Percentage
Research settings		

Lab-based	2	8%
Classroom-based	22	88%
Not reported	1	4%
Means of communication		
Face to face	25	100%
Computer-mediated	0	0%
Research types		
Quasi-experimental	24	96%
Non-experimental	1	4%
Effect sizes in quasi-experimental research		
Reported	5	20.8%
Not reported	19	79.2%
Number of exposures to MTFI		
One	23	92%
Multiple	2	8%
Piloting of model texts		
Reported	6	24%
Not reported	19	76%

Sub-category	**Value**
Sample sizes in quasi-experimental research	
Mean	39.54
SD	20.89
Minimum	12
Maximum	90

Table 3 shows that most MTFI studies took place inside a classroom (n = 22; 88%), and face-to-face was the only employed means of communication between teachers/researchers and learners as well as between learners and learners. Furthermore, 24 out of 25 (96%) of the selected research was quasi-experimental; however, the sample sizes in these studies were relatively small (M = 39.54) with a large variance (SD = 20.89) given that there were at least two groups of students (e.g., control group and treatment group). In addition, although interventions were widely adopted in the selected MTFI articles, only five (20.8%) reported effect sizes of the significant outcomes. It is also notable that the established effect of MTFI so far was mostly short-term as a large portion of studies only provided learners with one-time exposure to model texts (n = 23; 92%).

Additionally, only six studies (24%) reported conducting the piloting of model texts prior to the full-scale research.

Besides general study design features, how MTFI-related choices were made in the selected pool was also explored. These include writing modes, writing types, number of provided models, and writers of models, which was delineated in Table 4.

Table 4

MTFI-related design features

Sub-category	Frequency	Percentage
Writing modes		
Individual writing	13	52%
Collaborative writing	9	36%
Both	3	12%
Writing types		
Narrative	20	76.9%
Argumentative	5	19.2%
Descriptive	1	3.9%
Number of provided model texts		
One	10	40%
Two	15	60%
Writers of model texts		
Class teachers	4	15.4%
Researchers	2	7.7%
External authors	15	57.7%
Not reported	5	19.2%
L1 users of the target language	11	44%
L2 users of the target language	4	16%
Both	2	8%
Not reported	8	32%

As can be seen from Table 4, although individual writing was mostly employed (n = 13; 52%) in the selected research, a large portion of the articles (n = 9; 36%) also adopted the collaborative writing mode. As for writing types, narratives were predominantly chosen (n = 20; 76.9%) while argumentative (n = 5; 19.2%) and descriptive (n = 1, 3.9%) writing did not receive

much attention. With respect to model texts, they were mainly written by external authors (*n* = 15; 57.7%) rather than class teachers (*n* = 4; 15.4%) or researchers (*n* = 2; 7.7%). Furthermore, writers of model texts were commonly L1 speakers of the target language (i.e., English) (*n* = 11, 44%); however, in eight papers (32%), model writers' first language was not clearly stated. Finally, two models (*n* = 15; 60%) were usually administered to students instead of one model (*n* = 10; 40%).

RQ3. Primary areas of focus and key findings

Despite adopting various methodological choices, the selected MTFI studies encompassed some common characteristics regarding research foci as well as findings. Detailed information is demonstrated in Table 5.

Table 5

Primary areas of focus and key findings in the MTFI studies

Areas of focus	Key findings
noticed problems in the composing stage (*n = 13*)	• lexical aspects were noticed the most (*n* = 11) • grammatical features were noticed the most (*n* = 2)
noticed features in the comparing stage (*n* = 13)	• lexical aspects were noticed the most (*n* = 11) • content-related features were noticed the most (*n* = 2)
incorporations of features in the revising/rewriting stage (*n* = 10)	• lexical aspects were incorporated the most (*n* = 9) • grammatical features were incorporated the most (*n* = 1)
relationship between noticing and incorporations (*n* = 9)	• noticing was positively linked to better changes in revised/rewritten texts (*n* = 8) • noticing was not linked to better changes in revised/rewritten texts (*n* = 1)
Impact of MTFI on text quality (*n* = 9)	• Overall grading (using the scale-based approach) (*n* = 7) were significantly improved (*n* = 6) vs insignificantly improved (*n* = 1) • Lexical complexity (*n* = 4) was significantly improved (*n* = 2) vs insignificantly improved (*n* = 2)

	• Syntactic complexity (n = 4) was significantly improved (n = 1) vs insignificantly improved (n = 3) • Accuracy (n = 4) was significantly improved (n = 1) vs insignificantly improved (n = 3) • Fluency (n = 4) was significantly improved (n = 1) vs insignificantly improved (n = 3)
Attitude towards MTFI (n = 8)	• Learners' view on MTFI was positive (n = 4) • Learners' view on MTFI was negative (n = 1) • Learners' view on MTFI was positive but declined slightly from the comparing stage to the revising/rewriting stage (n = 3)

As presented in Table 5, there were six major areas of research on MTFI within the past two decades. The first four pertained to writing as a process: the problems that learners noticed while composing their original draft (n = 13), the features in model texts to which learners paid attention in the comparing stage (n = 13), the features in models that learners chose to implement in the rewriting version (n = 10), and the correlation between noticing in stage one as well as stage two and the rates of incorporations in stage three (n = 9). The fifth strand of research addressed writing as a product: the impact of MTFI on L2 learners' text quality (n = 9) while the last one tapped into learners' attitude towards this instructional approach (n = 8).

Turning to the findings, it is apparent that the results remained mixed in all six areas of focus. However, tentative conclusions could be generated on account of the major outcomes (i.e., those with most frequent reports), which should still be interpreted with caution. First, in the composing (n = 11 out of 13) and comparing (n = 11 out of 13) stages, it was found that learners mainly drew their attention to lexis. Second, vocabulary was consistently reported to be the feature incorporated the most in the rewriting stage (n = 9 out of 10). Third, noticing in stage one and stage two resulted in better changes in the rewritten version (n = 8 out of 9). With regard to text quality, which was measured in terms of overall grades (i.e., based on a given rubric with several criteria such as vocabulary, grammar, organization, content) (n = 7 out of 9), lexical complexity (n = 4 out

of 9), syntactic complexity (*n* = 5 out of 9), accuracy (*n* = 4 out of 9), and fluency (*n* = 4 out of 9), it was found that while MTFI had a significantly positive effect on learners' overall grades (*n* = 6 out of 7), results in lexical complexity, syntactic complexity, accuracy, and fluency (CAF) were inconsistent. Finally, L2 learners' attitude towards MTFI was found to be generally positive (*n* = 4 out of 8); however, it might drop slightly between stages (i.e., from stage two to stage three) (*n* = 3 out of 8).

Apart from the primary foci already reported, we also investigated other lines of research in the selected sources to provide a more comprehensive picture of MTFI up to the first of April, 2023. A synthesis of other areas of focus and their findings is reported in Table 6.

Table 6

Other areas of focus and findings in the MTFI studies

Areas of focus	**Key findings**
Noticing quality (*n* = *1*)	• Learners being exposed to model texts as a feedback instrument accepted noticed features with more reasoning, leading to higher-quality noticing, than those receiving peer feedback.
Impact of proficiency levels (*n* = 5)	• More proficient learners noticed and incorporated more features in model texts than less proficient partners (*n* = 3) • No significant relationships were found between proficiency levels and rates of noticing as well as incorporations (*n* = 1) • Less proficient learners were more likely to notice surface-level features in model texts (*n* = 1)
Comparison between models and reformulations (*n* = 1)	• Model texts helped address overt and covert problems almost equally while reformulations mainly coped with overt problems in the composing stage
Guided vs unguided noticing (*n* = 2)	• No significant differences in noticing and incorporations were found between guided and unguided noticing groups (*n* = 2)

Learners' trajectories in processing model texts (n = 1)	• Learners followed various trajectories of noticing (i.e., reported/unreported problems in the first stage), and incorporations (i.e., unsolved/partially solved/fully solved problems) in their writing tasks. • Learners being exposed to model texts tended to follow a trajectory deemed to be more beneficial than those without model-text exposure.
The role of working memory (n = 1)	• Working memory was a strong predictor of learners' text quality
Learners' acquisition of third person possessive pronouns (n = 1)	• Long-term exposure to models helped learners to significantly enhance the acquisition of third person possessive pronouns
Learners' lexical development (n = 1)	• Exposure to argumentative models helped enhance L2 lexical development, measured via vocabulary tests

Table 6 demonstrates eight other areas of focus in the selected MTFI research, encompassing noticing quality (n = 1), trajectories of processing feedback (n = 1), comparison between models and reformulations (n = 1), lexical (n = 1) and grammatical (n =1) development, and the role of proficiency levels (n = 5), guides on noticing (n = 2) as well as working memory (n = 1) on learners' subsequent writing. In general, it seemed that MTFI could lead to positive results in terms of noticing quality, writing trajectories, and development of lexis and grammar (i.e., third person possessive pronouns). Moreover, working memory and proficiency levels might be good indicators of more or better noticing and incorporations whereas detailed guides on noticing did not play a significant role. Nonetheless, these outcomes should be read with caution as reported findings were still very limited.

Discussion

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

The present study aims to systematically review empirical L2 writing research using model texts as a feedback facilitator, published in journals indexed in Web of Science (SSCI and ESCI) up to the first of April, 2023 (n = 25). Particularly, it focuses on three main aspects: study contexts, methodological features as well as key areas of focus and findings.

RQ1: The study contexts in the selected MTFI research

One important pattern in the previous MTFI studies was the central focus on young learners (n = 16; 64%) at primary school (n = 12; 46.2%). This trend diverges from other fields in L2 writing research such as WCF (for a review, see Crosthwaite et al., 2022) and computer-mediated collaborative writing (CMCW) (for a review, see Zhang et al., 2021), which mostly concentrated on adult learners in higher education. It could be indicated that most of the selected articles on MTFI have tapped into the population that is often underrepresented in L2 research. However, learners of other age groups such as middle (n = 4; 15.4%) and highschool (n = 1; 3.8%) students were still not well reflected in previous MTFI studies, which may limit the generalizability of their findings to these populations.

Another tendency was that all of the selected MTFI studies (n = 25) took place in the EFL setting. This is also echoed in other L2 writing domains (e.g., WCF, Crosthwaite, 2022; CMCW, Zhang et al., 2021) given that English is the global language (Rose et al., 2021). Therefore, practitioners and learners in the English as a second language (ESL) and non-English contexts may encounter dissimilar outcomes when applying this instructional approach in their classrooms. Additionally, a majority of the MTFI works recruited participants whose English proficiency ranged from low-intermediate to intermediate levels. One possible explanation for this choice is that lower-level learners, whose interlanguage is not well developed yet, may benefit from models

as a facilitator. This tendency makes it hard for educators who work with advanced EFL learners to decide on whether MTFI should be implemented in their classes or not.

These outcomes indicate that the role of learner-related factors such as ages, L2 proficiency levels, and educational backgrounds in MTFI has not been thoroughly investigated yet. As group analyses may mask the understanding of individual work (Chan, 2017; Gánem-Gutiérrez & Gilmore, 2018), it is recommended that MTFI-related findings should be read with care.

RQ2: Methodological characteristics in the selected MTFI research

General design features

It was found that a large majority of the selected MTFI studies adopted a quasi-experimental design ($n = 24$; 96%), taking place mainly in a classroom-based setting ($n = 22$; 88%). This trend matched Manchón's (2011) call for (quasi)experimental studies to be conducted in a real classroom with real students and teachers, which strengthens the research's ecological validity. A plausible explanation for a wide adoption of the quasi-experimental design is that MTFI is an emerging research domain, which heightens the need to explore its efficacy. However, the experiments were mostly one-shot, meaning that learners were only exposed to model texts once ($n = 23$; 92%). It can thus be indicated that the effectiveness of MTFI to date was largely short-term.

Another finding is that sample sizes in these quasi-experimental studies were relatively small with a large variance ($M = 39.54$, $SD = 20.89$). Given that there were at least two groups of students in such a study, the average number of participants in each group was just under 20, which may render the findings not generalizable to other research contexts (Cohen et al., 2017). Additionally, there was a lack of reports of effect sizes in most of the quasi-experimental studies

on MTFI (n = 19; 79.2%). This could be a limitation because without effect sizes, it would be difficult to understand the practical significance of an intervention (Larson-Hall & Plonsky, 2015).

In terms of piloting of models, only six studies (24%) on MTFI reported performing this step prior to the full-scale research. For example, Luquin and García Mayo (2021) had the model text piloted to check if it was comprehensible and suitable for the target learners. It is salient to pilot model texts before the official research because inappropriate models (i.e., too difficult for learners to comprehend) would lead to low motivation (García Mayo & Loidi Labandibar, 2017) or low incorporations of model features into subsequent writing (Hanaoka, 2007).

Overall, small sample sizes as well as a lack of model-text piloting, effect-size reporting in most of the selected papers may limit our understanding of MTFI and the generalizability of the obtained outcomes. Therefore, interpretations related to this instructional approach should be made with caution.

MTFI-related design features

With regard to choices related to MTFI, we found that narratives were primarily employed (n = 20; 76.9%), far more than argumentative (n = 5; 19.2%) and descriptive writing (n = 1; 3.9%). The predominant choice of this writing type could be due to a heavy focus on young learners (n = 16; 64%) who often work with stories rather than academic compositions. However, MTFI has been demonstrated to be also effective in promoting gains in L2 argumentative writing (Kang, 2020, 2022; Authors, XXXX). This finding coupled with the claim that different writing genres require different language use and may lead to dissimilar outcomes (Lu, 2011; Yoon & Polio, 2017) highlights that further investigation into the argumentative genre is needed, especially for adult learners at tertiary levels (Roohani & Taheri, 2015; Yang et al., 2023).

Another pattern is that both individual (n = 13; 52%) and collaborative (n = 9; 36%) writing were employed in the selected MTFI research. It can be inferred that MTFI can be applicable to either of the writing modes. The choice is largely dependent on the research purposes and conditions. On the one hand, collaborative writing was adopted probably because writing together could facilitate L2 learning and writing skills (Coyle & Roca de Larios, 2014; Li & Storch, 2017; Storch, 2013). On the other hand, individual writing might be preferred due to the background noise of students' conversations (García Mayo & Loidi Labandibar, 2017), making it hard to record their speaking for later analysis.

As for the number of model texts administered to students, two models (60%) were employed more often than one model (40%), probably because the former would offer more solutions and thereby increase the possibility of students' incorporations of these elements into their subsequent writing (Hanaoka, 2007; García Mayo & Loidi Labandibar, 2017). However, the use of one model has the benefit of reducing cognitive burdens for learners as well as saving time for other activities in the classroom. Therefore, the choice of one or two models should be taken into careful consideration depending on the purpose and conditions of MTFI research.

Concerning the authors of models, a large proportion of the selected MTFI studies (n = 15; 57.7%) invited external writers (i.e., neither the class teachers nor the researchers nor the teachers at the research site) to compose the texts. Furthermore, the writers of the models were mostly L1 users of English (n = 11; 44%), with only 16% (n = 4) being L2 users. Up to 32% (n = 8) of the selected papers did not clearly report the writers' mother tongue. Although external L1 writers were able to compose objective and authentic texts, they might not fully understand the participants' proficiency levels and background knowledge as well as class teachers. Therefore, in-charge teachers should be invited to write model texts for their learners, which helps strengthen the

ecological validity. As long as models meet certain requirements (e.g., having good quality, being reviewed by other experienced teachers or researchers), the writers' first language should not matter. This is particularly helpful, especially in contexts where it is difficult to have access to L1 users of English with sufficient experience and expertise in English language teaching.

RQ3: Primary areas of focus and key findings in the selected MTFI research

Key areas of focus

During the past two decades, research on MTFI has revolved around six key areas: noticed problems in the composing stage (n = 13), noticed features in the comparing stage (n = 13), incorporations of features in models into subsequent texts (n = 9), relationship between the noticing and incorporations (n = 10), text quality (n = 9), and learners' attitude towards MTFI (n = 8). It can be indicated that a large portion of the selected MTFI studies followed the process-oriented approach by focusing on learners' cognition via a multi-stage writing task (e.g., Coyle et al., 2018; Hanaoka, 2007). This trend is in line with the call for more research on writing processes by Machón and Leow (2020), which provides us with insights into L2 writing in general and MTFI in particular. Recent research on MTFI has also employed the product-oriented approach by analyzing learners' texts in terms of complexity, accuracy, fluency, and overall grades (e.g., Lázaro-Ibarrola, 2021; Tieu & Baker, 2022). Furthermore, the exploration of learners' attitude sheds light on how they were emotionally engaged with MTFI tasks. A firm grasp of learners' affective engagement is important as it may in part affect their behavior (Authors, XXXX; Oga–Baldwin & Nakata, 2017). In brief, the focus on these six areas could advance our understanding of how MTFI might help promote learners' L2 form-meaning development.

Other foci include how L2 proficiency levels (n = 5) as well as working memory (n = 1) may predict learners' writing performance and the trajectories that learners follow (n = 1) when

processing models. It can be inferred that MTFI researchers also attempted to examine the role of individual differences in writing as both a process and a product. These areas of focus, though still very limited, are promising because they tap into the learner-related factors which have been demonstrated to play a significant part in language learning (Dörnyei, 2005; Ellis, 2008; Pawlak & Kruk, 2022).

Key findings

The review of 25 selected papers on MTFI reveals that L2 learners paid most of their attention to lexical aspects, in both composing and comparing stages. For example, Hanaoka (2007) reported that of all learners' noticed features, 92.4% (composing stage) and 65.4% (comparing stage) were related to lexis. These findings were supported by subsequent studies (e.g., Baleghizadeh & Arab, 2011; Montealegre Ramón, 2019) which displayed similar outcomes. The predominant focus on lexis could be because this aspect was more apparent and easily reported than grammatical features (Hanaoka, 2007). Another plausible factor was that vocabulary carries the most content that is needed for learners to convey their intended message. These noticed features were shown to affect what L2 learners chose to incorporate in their subsequent texts, which was predominantly lexis (e.g., Hanaoka, 2007; García Mayo & Loidi Labandibar, 2017). These outcomes align with the reports in the review by Cao and Mao (2022), lending support to the Output Hypothesis (Swain, 1985) and the Noticing Hypothesis (Schmidt, 1990, 2001). Swain (1985) stated that learners tended to notice the problems they had when composing their original draft (output), which motivated them to look for solutions (presented in model texts) to these issues. Such noticing might facilitate their learning gains (Schmidt, 1985, 2001) as those who noticed the most could benefit the most (Schmidt & Frota, 1986). Thus, it can be indicated that MTFI has the potential to promote L2 learners' cognitive and lexical development.

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

Another key finding that MTFI could be effective in enhancing L2 learners' text quality, mainly in terms of overall grading. For instance, Kang (2022) found that Korean EFL undergraduates, after exposure to model texts, had their scores in four criteria (i.e., vocabulary, grammar, content, organization) considerably enhanced, outperforming those without access to models. Regarding the CAF (complexity, accuracy, and fluency) measures, results remained mixed, probably due to the choices of different indices. For example, findings regarding fluency were inconsistent, with three papers reporting insignificant impact (Lázaro-Ibarrola, 2021; Roothooft et al., 2022; Villarreal & Lázaro-Ibarrola, 2022) and one study showing positive progress (Criado et al., 2022). While Criado et al. (2022) employed ten product-process indices (e.g., words per minute, characters per minute), Lázaro-Ibarrola (2021) used only three indices (i.e., number of words, T-units, and clauses per text). The adoption of various indices of CAF makes it difficult to compare results across studies, which limits their generalizability.

With respect to learners' attitude, MTFI was generally perceived as a useful tool that helped improve text quality. However, it was reported that learners' motivation dropped from the comparing stage to the rewriting stage (Lázaro-Ibarrola & Villarreal, 2021; Roothooft et al., 2022). The motivation drop could be because the models were difficult for learners to comprehend (García Mayo & Loidi Labandibar, 2017) or might not contain suitable solutions to learners' noticed problems in the composing stage (Hanaoka, 2007). Another factor could be due to the state that students thought they had to compose something as good as the models (Roothooft et al., 2022), which put them under pressure. However, the MTFI studies on learners' attitude mainly used a thermometer (i.e., a single scale from 1.0 to 10.0 that indicates learners' motivation with 1.0 being the lowest and 10.0 being the highest) (e.g., Lázaro-Ibarrola, 2021; Roothooft et al., 2022), in-

depth interviews and surveys have not been widely employed to further examine learners' perceptions yet. Consequently it is advisable that these results should be read with care.

The selected MTFI research also demonstrates that proficiency levels and working memory may significantly affect L2 learners' writing. Specifically, it was found that while more proficient learners noticed and incorporated more features (Cánovas Guirao et al., 2015; Hanaoka, 2007), less proficient ones were more likely to focus on the surface-level features in model texts (Coyle & Roca de Larios, 2020). This finding corroborates the MTFI synthesis made by Cao and Mao (2022), which states that "L2 proficiency ... could impact students' processing of models and their motivation to use this feedback method to improve writing" (p. 4). As for working memory, Kang (2022) reported that this variable was a strong predictor of learners' text quality. This outcome is well echoed in other fields of language learning such as L2 vocabulary (Teng, 2023) and L2 writing (Vasylets & Marín, 2021). Such findings indicate that individual differences are of paramount importance in advancing our understanding of learners' writing in the MTFI instructional approach, especially when research on this line is still scarce.

Directions for future research

Despite the gallant endeavors made to explore the potential of MTFI in previous studies, there are still many aspects that need to be further explored, mainly in terms of individual differences, methodological design, MTFI-related choices, and L2 learners' text quality. The first dimension, individual differences, should receive more attention because it has been demonstrated to play an integral part in L2 learning and writing (Dörnyei, 2005; Ellis, 2008; Pawlak & Kruk, 2022). The selected MTFI studies attempted to explore the effect of L2 proficiency levels (e.g., Coyle & Roca de Larios, 2020; Hanaoka, 2007), working memory (Kang, 2022), and motivation on learners' writing (Lázaro-Ibarrola & Villarreal; Roothooft et al., 2022), but such works are still

very limited. Moreover, as most MTFI research has focused mainly on young learners at primary schools in the European context, other populations such as adolescents and adult learners in other contexts are still underrepresented. Thus, it is recommended that future studies concentrate on these learner-related factors to shed light on how each individual learns from and processes model texts as a feedback instrument.

Regarding methodological design, more longitudinal studies with a sufficient sample size are needed to extend our understanding of the long-term effect of MTFI as the majority of past research on this topic adopted the one-shot experiment with a relatively small sample size. In addition, effect sizes should be reported because without such figures, the magnitude of the effectiveness of a treatment or a condition would be obscured (Larson-Hall & Plonsky, 2015). Moreover, effect sizes would make it easier for researchers to perform 'exact cross-study comparisons' (Villarreal & Lázaro-Ibarrola, 2022, p. 13). Another notable point is that model texts need to be piloted prior to the full-scale study. This is because if a model does not contain suitable solutions to learners' noticed problems (Hanaoka, 2007) or if it is difficult for learners to comprehend (García Mayo & Loidi Labandibar, 2017), learning gains might not occur (Hanaoka, 2007), and motivation might suffer (García Mayo & Loidi Labandibar, 2017).

Turning to MTFI-related choices, class teachers, irrespective of their first language, should be invited to compose model texts. This is because they work directly with their students, which provides them with a better understanding of these learners' L2 levels, needs, and background knowledge. Furthermore, although a definition of model texts was given (Cánovas Guirao et al., 2015; Hamano-Bunce, 2022), it remains unclear what a good model text should be. Thus, future studies need to devise certain criteria for a good model text, which would help the choices of models become more comparable across research. Additionally, argumentative writing, which is

still underexplored, deserves more attention from MTFI researchers as this genre plays an integral part in higher education (Yang et al., 2023).

Another dimension that should be further investigated is the effect of MTFI on L2 learners' text quality in terms of CAF. This is because findings related to CAF are still limited and controversial. One possible factor lies in the different choices of indices, which probably led to opposing outcomes. Another variable is related to writing genres, particularly narrative and argumentative writing. Indeed, it was found that whereas narrative writing helped improve lexical diversity, argumentative writing led to higher scores in syntactic complexity and lexical sophistication (Lu, 2011; Yoon & Polio, 2017). Thus, it is recommended that future research should take these points into consideration when comparing results across studies.

Conclusion

This study reviewed a total of 25 studies on model texts as a feedback instrument in L2 writing. By examining the selected papers' study contexts, methodological characteristics as well as primary areas of focus, this review illuminates key trends and findings, identifies major shortcomings, and makes empirically founded suggestions for future works in this domain. However, there are two limitations that need to be acknowledged. First, our focus was placed on research exclusively published in journals indexed in Web of Science (i.e., SSCI and ESCI), which might entail a representativeness limitation (Hiver et al., 2021). Second, as the data collection in our study was limited to papers published until the first of April 2023, other MTFI publications after this time point have not been analyzed yet.

This review presents several pedagogical, theoretical, and methodological implications. From a pedagogical perspective, this study's findings suggest that model texts could be a potential means of feedback to promote L2 writing gains and thereby should be implemented into the

classroom context. However, MTFI should be considered as an alternative to other feedback instruments rather than an absolute replacement (Cao & Mao, 2022). This is because models seem to have selective effects on text quality (i.e., overall grading rather than CAF) (Lázaro-Ibarrola, 2021; Roothooft et al., 2022). A combination of both MTFI (implicit feedback) and WCF (explicit feedback) might be a potential strategy to maximize L2 writing development. Furthermore, models should be carefully composed and piloted because they dictate the success of the writing process. Besides this, teachers should take learner-related factors, especially motivation, into careful consideration since students seemed to feel demotivated after comparing their draft to model texts as a result of inferiority (i.e., thinking that their writing was poor) (Roothooft et al., 2022). To address this issue, students should be made aware of the task's objective (i.e., using models as a tool for learning rather than as final products that students must achieve). Additionally, the choice of one or two models for the writing task has its own merits and demerits, so teachers can determine the quantity of provided model texts depending on their classroom conditions and teaching plans. Theoretically, this review could help advance our understanding of the role of model texts as a feedback instrument in L2 writing in a systematic way. Moreover, it lends support to the combination of three important theories in SLA: the Output Hypothesis (Swain, 1985), the Noticing Hypothesis (Schmidt, 1990), and the Input Hypothesis (Krashen, 1982) in L2 writing. Methodologically, it provides suggestions on potential research designs that may help alleviate the limitations in the selected MTFI studies. Once these shortcomings are addressed, the role of model texts as a feedback instrument would be understood in a more comprehensive way, which could lead to even better learning gains.

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

References

Baleghizadeh, S., & Arab, F. (2011). Comparing native models and peer feedback in promoting noticing through written output. *Innovation in Language Learning and Teaching*, *5*(1), 63-79. https://doi.org/10.1080/17501229.2010.517535

Cánovas Guirao, J., Roca de Larios, J., & Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System*, *52*, 63-77. https://doi.org/10.1016/j.system.2015.04.002

Cao, Z., & Mao, Z. (2022). Studies on positive evidence feedback in second language writing: Status quo and implications. *RELC Journal*, 00336882211072905. https://doi.org/10.1177/00336882211072905

Chan, S. (2017). Using keystroke logging to understand writers' processes on a reading-into-writing test. *Language Testing in Asia*, *10*(7), 1–27. https://doi.org/10.1186/s40468-017-0040-5

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education*. Routledge.

Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition*, *36*(3), 451-485. https://doi.org/10.1017/S0272263113000612

Coyle, Y., & Roca de Larios, J. (2020). Exploring young learners' engagement with models as a written corrective technique in EFL and CLIL settings. *System*, *95*, 102374. https://doi.org/10.1016/j.system.2020.102374

Coyle, Y., Cánovas Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts.

Journal of Second Language Writing, 42, 25-43. https://doi.org/10.1016/j.jslw.2018.09.002

Criado, R., Garcés-Manzanera, A., & Plonsky, L. (2022). Models as written corrective feedback: Effects on young L2 learners' fluency in digital writing from product and process perspectives. *Studies in Second Language Learning and Teaching*, *12*(4), 697–719. https://doi.org/10.14746/ssllt.2022.12.4.8

Crosthwaite, P., Ningrum, S., & Lee, I. (2022). Research trends in L2 written corrective feedback: A bibliometric analysis of three decades of Scopus-indexed research on L2 WCF. *Journal of Second Language Writing*, *58*, 100934. https://doi.org/10.1016/j.jslw.2022.100934

DeKeyser, R.M. (2007), *Practice in a second language: perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Dörnyei, Z. (2005). The psychology of the language learner: Individual differences in second language acquisition. Lawrence Erlbaum.

Dyer, B. (1996). L1 and L2 Composition theories: Hillocks' 'environmental mode' and task-based language teaching. *ELT Journal* 50 (4), 312-317. https://doi.org/10.1093/elt/50.4.312

Ellis, N. C., & Wulff, S. (2015). *Second language acquisition*. De Gruyter Mouton.

Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly, 29*, 87-105. https://doi.org/10.2307/3587806

Ellis, R. (2008). Learner beliefs and language learning. *Asian EFL Journal*, *10*(4), 7-25.

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, *1*(1), 3-18. https://doi.org/10.5070/l2.v1i1.9054

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition, 32*(2), 181-201. https://doi.org/10.1017/S0272263109990490

Gánem-Gutiérrez, G.A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, *68*(2), 469–506. https://doi.org/10.1111/lang.12280

García Mayo, M. P., & Loidi Labandibar, U. (2017). The use of models as written corrective feedback in English as a foreign language (EFL) writing. *Annual Review of Applied Linguistics*, *37*, 110-127. https://doi.org/10.1017/S0267190517000071

Hamano-Bunce, D. (2022). The effects of direct written corrective feedback and comparator texts on the complexity and accuracy of revisions and new pieces of writing. *Language Teaching Research*, 13621688221127643. https://doi.org/10.1177/13621688221127643

Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, *11*(4), 459–479. https://doi.org/10.1177/1362168807080963

Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, *21*(4), 332–347. https://doi.org/10.1016/j.jslw.2012.09.008

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2021). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language teaching research*, 13621688211001289. https://doi.org/10.1177/13621688211001289

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in second language acquisition*, *24*(4), 541-577. https://doi.org/10.1017/S0272263102004023

Izumi, S. (2013). Noticing and L2 development: Theoretical, empirical, and pedagogical issues. In J. M. Bergleithner; S. N. Frota, & J. Kei, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 25-38). University of Hawai'i at Manoa. National Foreign language Resource Center.

Kang, E. Y. (2020). Using model texts as a form of feedback in L2 writing. *System*, *89*, 102196. https://doi.org/10.1016/j.system.2019.102196

Kang, E. Y. (2022). The role of working memory in the effects of models as a written corrective strategy. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2021-0207

Krashen, S. (1982). *Second language acquisition and second language learning*. Pergamon Press.

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, *65*(S1), 127-159. https://doi.org/10.1111/lang.12115

Lázaro-Ibarrola, A. (2021). Model texts in collaborative and individual writing among EFL children: noticing, incorporations, and draft quality. *International Review of Applied Linguistics in Language Teaching*, *61*(2), 329-359. https://doi.org/10.1515/iral-2020-0160

Lázaro-Ibarrola, A., & Villarreal, I. (2021). Are EFL writers motivated or demotivated by model texts and task repetition? Evidence from young collaborative writers. *International Journal of English Studies*, *21*(2), 29-55. https://doi.org/10.6018/ijes.466401

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning, 31*(8), 882–904. https://doi.org/10.1080/09588221.2018.1465981

Li, M., & Storch, N. (2017). Second language writing in the age of CMC: Affordances, multimodality, and collaboration. *Journal of Second Language Writing*, *36*, 1–5. http://dx.doi.org/10.1016/j.jslw.2017.05.012

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Luquin, M., & García Mayo, M. P. (2020). Collaborative writing and feedback: An exploratory study of the potential of models in primary EFL students' writing performance. *Language Teaching for Young Learners, 2*(1), 73-100. https://doi.org/10.1075/ltyl.19007.luq

Luquin, M., & García Mayo, M. P. (2021). Exploring the use of models as a written corrective feedback technique among EFL children. *System*, *98*, 102465. https://doi.org/10.1016/j.system.2021.102465

Luquin, M., & García Mayo, M. P. (2023). The impact of textual enhancement on the acquisition of third person possessive pronouns by child EFL learners. *International Review of Applied Linguistics in Language Teaching.* https://doi.org/10.1515/iral-2022-0176.

Lynch, T. (2009). Responding to learners' perceptions of feedback: The use of comparators in second language speaking courses. *International Journal of Innovation in Language Learning and Teaching*, *3*(2), 191–203. https://doi.org/10.1080/17501220802379109

Manchón, R. M. (2011). The language learning potential of writing in foreign language contexts: Lessons from research. In M. Reichelt, & T. Chimasko (Eds.), *Foreign language writing: Research insights* (pp. 44e64). Parlor Press.

Manchón, R.M., & Leow, R.P. (2020). An ISLA perspective on L2 learning through writing. Implications for future research agendas. In R.M. Manchón (Ed.), *Writing and language learning. Advancing research agendas* (pp. 335–355). John Benjamins.

Martínez Esteban, N., & Roca de Larios, J. (2010). The Use of Models as a Form of Written Feedback to Secondary School Pupils of English. *International Journal of English Studies*, *10*(2), 143–170. https://doi.org/10.6018/ijes/2010/2/119241

McHugh, M.L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276–282.

Montealegre Ramón, F. (2019). Exploring the learning potential of models with secondary school EFL learners. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, *32*(1), 155-184. https://doi.org/10.1075/resla.00013.mon

Muranoi, H. (2007). Output practice in the L2 classroom. In R. M. DeKeyser (Ed.), *Practice in a second language: perspectives from applied linguistics and cognitive psychology*, (pp. 51–84). Cambridge University Press.

Oga–Baldwin, W. Q., & Nakata, Y. (2017). Engagement, gender, and motivation: A predictive model for Japanese young language learners. *System*, *65*, 151–163. https://doi.org/10.1016/j.system.2017.01.011

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting*

systematic reviews. International Journal of Surgery, 88, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Pawlak, M., & Kruk, M. (2022). *Individual differences in computer assisted language learning research*. Taylor & Francis.

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of second language writing*, *10*(4), 277-303. https://doi.org/10.1016/S1060-3743(01)00046-7

Robinson, P. (1995). Attention, memory and the "noticing" hypothesis. *Language Learning, 45*, 283–331. https://doi.org/10.1111/j.1467-1770.1995.tb00441.x

Roca de Larios, J., Coyle, Y., & García, V. (2022). The effects of using cognitive discourse functions to instruct 4th-year children on report writing in a CLIL science class. *Studies in Second Language Learning and Teaching*, *12*(4), 597-622. https://doi.org/10.14746/ssllt.2022.12.4.4

Roothooft, H., Lázaro-Ibarrola, A., & Bulté, B. (2022). Task repetition and corrective feedback via models and direct corrections among young EFL writers: Draft quality and task motivation. *Language Teaching Research*, 13621688221082041. https://doi.org/10.1177/13621688221082041

Rose, H., McKinley, J., & Galloway, N. (2021). Global Englishes and language teaching: A review of pedagogical research. *Language Teaching*, *54*(2), 157-189. https://doi.org/10.1017/S0261444820000518

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, *29*, 67–100. https://doi.org/10.1017/S0272263107070039

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158. https://doi.org/10.1093/applin/11.2.129

Schmidt, R. (1995). *Attention and Awareness in Foreign Language Learning.* University of Hawaii Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Newbury House.

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, *15*, 165–179.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 495–508). Routledge.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The modern language journal*, *82*(3), 320-337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Teng, M. F. (2023). Effectiveness of captioned videos for incidental vocabulary learning and retention: The role of working memory. *Computer Assisted Language Learning*, 1-29. https://doi.org/10.1080/09588221.2023.2173613

Tieu, L., & Baker, J. R. (2022). Using model essays in conjunction with noticing as a feedback instrument in IELTS writing preparation. *Innovation in Language Learning and Teaching*, 1–13. https://doi.org/10.1080/17501229.2022.2049799

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183–203. https://doi.org/10.1017/S0272263100012870

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327-369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second language research*, *14*(2), 103-135. https://doi.org/10.1191/026765898674803209

Vasylets, O., & Marín, J. (2021). The effects of working memory and L2 proficiency on L2 writing. *Journal of Second Language Writing*, *52*, 100786. https://doi.org/10.1016/j.jslw.2020.100786

Villarreal, I., & Lázaro-Ibarrola, A. (2022). Models in collaborative writing among CLIL learners of English in primary school: Linguistic outcomes and motivation matters. *System*, *110*, 102922. https://doi.org/10.1016/j.system.2022.102922

Yang, L., & Zhang, L. (2010). Exploring the role of reformulations and a model text in EFL students' writing performance. *Language Teaching Research, 14*, 464-484. https://doi.org/10.1177/1362168810375369

Yang, Y., Yap, N. T., & Ali, A. M. (2023). Predicting EFL expository writing quality with measures of lexical richness. *Assessing Writing*, *57*, 100762. https://doi.org/10.1016/j.asw.2023.100762

Yoon, H.-J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly, 51*(2), 275–301. https://doi.org/10.1002/tesq.296

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication*, *35*(1), 32-57. https://doi.org/10.1177/0741088317735836

Zhang, M., Gibbons, J., & Li, M. (2021). Computer-mediated collaborative writing in L2 classrooms: A systematic review. *Journal of Second Language Writing*, *54*, 100854. https://doi.org/10.1016/j.jslw.2021.100854

Zhang, X. (2020). A bibliometric analysis of second language acquisition between 1997 and 2018. *Studies in Second Language Acquisition*, *42*(1), 199-222. https://doi.org/10.1017/S0272263119000573

Zhang, X., & Wang, M. (2022). The continuation task and the model-as-feedback writing task in L2 vocabulary learning: Timing of model texts. *Language Teaching Research*, 13621688221119077. https://doi.org/10.1177/13621688221119077

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

Appendix A

List of journals

Journal	Number of MTFI publications
Annual Review of Applied Linguistics	1
Innovation in Language Learning and Teaching	2
International Journal of English Studies	2
International Review of Applied Linguistics in Language Teaching	3
Journal of Second Language Writing	2
Language Teaching for Young Learners	1
Language Teaching Research	5
Language Testing in Asia	1
Spanish Journal of Applied Linguistics	1
Studies in Second Language Acquisition	1
Studies in Second Language Learning and Teaching	1
System	5

How to cite this article: Nguyen, L. Q., Dao, P., & Nguyen, B. T. T. (2024). Model texts as a feedback instrument in second language writing: A systematic review. *Language Teaching Research*, https://doi.org/10.1177/13621688241291336

Appendix B

Percentages of agreement between coders

Category	Sub-category	Item	First and third researchers	First and second researchers
Study contexts	Research places		100%	100%
	Educational levels		100%	100%
	Learner types		100%	100%
	L2 proficiency		75%	100%
Methodological characteristics	General study design	Research settings	75%	100%
		Means of communication	100%	100%
		Research types	100%	100%
		Piloting of models	100%	100%
		Effect sizes in quasi-experimental studies	100%	100%
		Number of exposures to model texts	75%	100%
		Sample sizes	100%	100%
	MTFI-related design features	Writing modes	100%	100%
		Writing types	100%	100%
		Number of provided models	100%	100%
		Writers of models	100%	100%
Primary areas of focus and key findings	Noticed problems in the composing stage	Most noticed	75%	80%
	Noticed features in the comparing stage	Most noticed	100%	100%

	Incorporations of features in models in the last stage	Most incorporated	75%	100%
	Relationship between noticing and incorporations	Positive or negative	100%	100%
	Impact of MTFI on text quality	Scale-based measures, lexical complexity, syntactic complexity, accuracy, fluency	100%	100%
	Attitude towards MTFI	Positive, negative, positive but declining	75%	100%
	Others (List of other foci)	List of key corresponding findings	75%	80%
TOTAL			92.05%	98.18%