This is the accepted version of the publication Rogers, J., & Li, P. (2025). The time between tasks in task repetition research: A systematic review. Language Teaching Research, 0(0). Copyright © 2025 The Author(s). DOI: 10.1177/13621688251323047.

Abstract

This article reports on a systematic review of oral task repetition research carried out between 1996 and 2023. This review focuses on the methodological features of these studies, specifically on issues related to how tasks have been spaced and repeated within this body of research. This review starts with an overview of the concept of input spacing and the major methodological paradigms that have been used to investigate it across the psychological sciences. It then discusses task repetition, providing definitions and elaborations of theoretical models, discussing why spacing might influence task performance, and how this informs the synthesized research. The article then presents a synthesis of the methods and results of the 107 studies that have been analyzed as part of the synthesis. The methodological synthesis includes analyses of how spacing has been implemented, the number of times tasks were repeated, and how researchers have justified their methodological decisions. The results of the synthesis highlight the need for greater systematicity and theoretical rationales for choice of spacing intervals, number of task repetitions, and transparency in reporting practices. The methods and results are discussed by identifying trends, exemplifying practices, and recommending solutions.

Task repetition, that is, when learners repeat a task after having completed it, has emerged as a central issue within task-based approaches to second language acquisition (SLA). Within SLA, and, in particular, instructed SLA, the expansive growth of empirical research over the past thirty years on the effects of task repetition is due to its potential to facilitate both task performance and second language (L2) development (Bygate, 2018; DeKeyser, 2018; Ellis et al., 2019). The rationale for the benefits of repeating spoken tasks[1], the focus of the current manuscript, is that speech production is an effortful process for L2 learners. Repeating a spoken task supports speech production processes by reducing a learner's cognitive load. It does so by providing the opportunity for learners to recycle and reuse conceptual and linguistic content from previous task performances (Suzuki et al., 2022).

The amount of time between task repetitions, or spacing, may influence the impact of repetition on task performance and learning outcomes (Bygate, 1996; DeKeyser, 2018; Rogers, 2023). Memory decays over time. Depending on how much time elapses between task performances, information from the first task performance may be forgotten. Thus, the learner may be unable to reuse any content or linguistic constructions from the first performance. In shorter spacing gaps, information, potentially content and linguistic, is available to be retrieved and reused from earlier task performances, thus reducing a learner's cognitive load (Lambert et al., 2017). Despite these putative theoretical benefits, it is only recently that task repetition researchers have begun to directly control the time between task repetitions, i.e., *spacing*, as part of their experimental designs (e.g., Bui et al., 2019; Kobayashi, 2022; Suzuki & Hanzawa, 2023). These studies have provided some preliminary evidence that spacing influences task performance, though this relationship's full extent and nature require further research.

Given this preliminary evidence and the fact that task repetition researchers have not considered spacing as a variable in their experimental designs, some researchers have suggested that spacing may be a confound that accounts for some of the inconsistent findings across previous task repetition literature (DeKeyser, 2018; Fukuta, 2016; Rogers, 2023). Thus, understanding what spacing gaps have been used constitutes a critical step in interpreting previous findings. This article aims to provide a methodological synthesis of the spacing gaps that have been used in oral task repetition literature to motivate and inform future research. This article has two major sections: background and research synthesis. In the background, we first discuss the principles of input spacing and the methodological paradigms used in input spacing research. We then turn to task repetition, discussing why input spacing and the number of task repetitions might influence task performance. The second section provides a discussion of the methods and findings of the current study. Specifically, it provides a synthesis of the methodology and implementation of spacing gaps in task repetition research.

Input Spacing

Input spacing, or *distribution of practice*, concerns the role of repetition in learning, specifically how manipulating the amount of time between repetitions influences performance, learning, and/or long-term retention (Wiseheart et al., 2019). A number of unitary and hybrid theoretical accounts have been proposed to account for the benefits of spacing on learning (Delaney et al., 2010). Although no single theoretical account is considered fully satisfactory (see Wiseheart et al., 2019 for discussion), retrieval mechanisms are generally accepted as a vital cognitive component of current unitary or hybrid models. Given the methodological focus of the current article, a discussion of these theoretical

accounts is beyond the scope of the current article, given its methodological focus. For a more in-depth discussion, readers may consult Rogers (2023) or Delaney et al. (2010).

Methodological Paradigms in Spacing Research

Two major research paradigms have been used to investigate distribution of practice effects in the psychological sciences: within-session and between-session. We discuss these as they are relevant for discussing how spacing intervals have been used in past task repetition studies in SLA and we will argue that these will provide a framework for a more systematic application of spacing intervals for future task repetition research.

The *within-session paradigm* manipulates spacing within a short time frame, typically a single experimental session. There are multiple variations of this paradigm. Most commonly, researchers manipulate the time between repetitions indirectly by having a single list and manipulating the distance that items repeat within the list. This can be done by having some items repeat in close proximity and others repeat after a larger number of intervening items. For example, massed items might be presented in immediate succession with no intervening items (e.g., AABBCC). Spacing can be achieved through the use of intervening items (e.g., ACCA), including but not limited to the use of filler items. In the first example, there are no intervening items between the target item and *A*. In the second example, there are two items between the repeat presentation of the target item *A*. One variant of the within-session paradigm examines *blocked* versus *interleaved* presentations. In these variants, the emphasis is on the sequence rather than on the amount of spacing between presentations (Rogers, 2023). In blocked schedules, target items are presented in a massed sequence, with a complete set followed by another (e.g., AAABBBCCC, etc.). Interleaved presentations shuffle the sets of items (e.g., ABC ABC ABC; Suzuki, 2021).

The between-session paradigm examines the effects of spacing over multiple experimental sessions (≥ 2 training sessions). Simple designs comprise two learning sessions plus a testing session, whereas more complex designs may include multiple training sessions. The gaps between training sessions (*intersession intervals*, or ISI) can be equal or different lengths, and the time between training and time between training and testing (*retention interval*, or RI) can also be manipulated to examine the effects of spacing on long-term retention.

Task Repetition

We acknowledge that several definitions of task exist in the L2 literature. As we were conducting a review of the task repetition literature, we adopted a definition that was consistent with early task repetition studies where a *task* "is an activity which requires learners to use language, with emphasis on meaning, to attain an objective" (Bygate et al., 2001, p. 11). We chose Bygate et al.'s definition as it is broad and most L2 researchers would agree with its core characteristics, that is, that tasks are activities with a communicative purpose and a non-linguistic outcome (Ellis & Shintani, 2013; Long, 2015; Mackey, 2020). *Task repetition* refers to the repetition of the "same or slightly altered tasks—whether whole tasks or parts of a task" (Bygate & Samuda, 2005, p. 43). As the above definition suggests, there are a number of variations of how tasks might be repeated. Commonly, studies differentiate between *task repetition* (or *exact* or *same* task repetition) and *procedural repetition*. Same task repetition involves performing the same task, with the same content and same procedure, more than once. Procedural repetition (or *task-type repetition*) involves repeated performance of a task with the same procedure but with new content. An example of task repetition might entail learners performing the same task on multiple occasions, such as an information-exchange task about hosting an American friend.

A contrasting example of procedural repetition would entail learners performing three information-exchange tasks with different content, such as (a) hosting an American friend, (b) describing school events or activities, and (c) discussing mayoral candidates (e.g., Y. Kim & Tracy-Ventura, 2013). A third category of task repetition proposed by Patanasorn (2010) is *content repetition*. Content repetition involves the repetition of a task that draws upon the same content knowledge but with modification to the procedure. In Patanasorn (2010), this involved drawing upon the same background knowledge to complete an array of different task types (story completion, information exchange, voting for a candidate).

The benefits of task repetition are most commonly attributed to Levelt's (1989) speech production model (see also Kormos, 2006), which involves the three stages of conceptualization, formulation, and articulation. Conceptualization involves the creation of a preverbal message; formulation involves the encoding of the preverbal message in linguistic form; and articulation involves the production of the speech utterance. When learners perform tasks for the first time, they devote considerable attentional resources to the conceptualization of the intended message. Subsequent task performances will be likely to require fewer attentional resources at the conceptualization stage, as learners are familiar with the content of the task and their intended message. This allows them to shift their attention resources toward processes of formulation and articulation. Additionally, learners can also reuse linguistic constructions at subsequent task performances (Suzuki et al., 2022), contributing to the proceduralization and automatization of speech processes (see de Jong & Perfetti, 2011 for discussion). The benefits of task repetition may vary depending on the type of task repetition that the learners perform. In exact task repetition, learners carry out the same task with the same content and procedures. In procedural repetition, learners engage in different content but with the same procedural or type of task (e.g., two different information exchange tasks: hosting an American friend and describing school events or activities). In content repetition, learners could perform two tasks that draw upon the same content knowledge but have different procedures (e.g., changing from an information exchange to a decision-making task; see Patanasorn, 2010).

Why might spacing influence task performance?

Intuitively, if there is a short spacing gap between task repetitions, then there will be a reduction in cognitive load during subsequent task performances with regard to conceptualization and formulation (Bygate & Samuda, 2005; Lambert et al., 2017). As memory decays over time, we can expect the memory of the first performance to fade with longer spacing gaps. This would result in increases in cognitive load with longer spacing gaps. With too long spacing gaps, then too much forgetting can take place for the repetition to be effective (DeKeyser, 2018). From a skill acquisition perspective, shorter spacing gaps may also be preferred in that they may further promote proceduralization processes. For discussions, see DeKeyser, 2018; Rogers, 2023.

Task repetition researchers may have an intuitive understanding that the frequency and intensity of task repetitions bear an influence on task performance and learning outcomes. For instance, Bygate (1996), in discussing the implications of his study, highlighted that "varying the amount of time between task repetitions might affect later performance" (p. 145). Similarly, in his (2001) discussion, he noted that "possibly for the fuller effect of task-type practice to emerge, more – or more massed – task exposure might be needed" (p. 43). Despite both spacing and intensity of exposure being identified as a variable of interest at an early stage, researchers have noted the lack of systematic attention paid to these variables in the task repetition literature (e.g., Ahmadian & Tavakoli, 2011), and it is only recently that studies have begun to task repetition researchers have begun to isolate and

examine the impact of spacing on task performance and learning outcomes. The results of these studies have indicated that spacing intervals do influence task performance and possible L2 learning outcomes, though the precise influence is in need of additional research (see Rogers, 2023 for a recent review).

With regard to the number of task repetitions, there are potentially theoretical links between both input spacing and task repetition and the number of times learners are asked to repeat the tasks. First, the magnitude of spacing effects may be linked to the number of repetitions (Rogers, 2023). This may be due to the opportunity for repeated retrieval associated with the additional training episodes. In task repetition studies, learners are typically required to perform tasks two, three or four times in total, although studies such as by Ahmadian (2011) have included up to 11 repetitions. In the literature on task repetition, it is clear that the number of repetitions can influence performance. For example, in a study involving young children, Sample and Michel (2014) found that task repetition improved task performance, although the performances (as measured via CAF measures) fluctuated between the second and third repetitions. Lambert et al. (2017), who asked learners to repeat tasks six times within a single lesson, observed a sharp increase in speech rate across the first three performances, followed by gradual increases from performances four to five and, finally, a leveling off from performances five to six, suggesting diminishing returns across multiple instances of task repetition.

The research synthesis

The purpose of this research synthesis is to provide a review and critique of the methodology related to spacing intervals used in previous oral task repetition SLA research. This includes spacing intervals, the number of task repetitions, and the rationale given by researchers for their methodological decisions related to these variables. This is done with the aim of presenting the current state of the literature, i.e., the status quo, to help facilitate the interpretation of existing literature and raise awareness of these aspects of methodological design to inform future research.

The synthesis is guided by the following research questions:

RQ1: What spacing paradigms, i.e., within-session, between-session, or a combination of within and between session (*within-between*), have been investigated in task repetition research?

RQ2: What spacing gaps (intersession intervals) have been investigated in task repetition studies?

> RQ2a: Do these spacing gaps differ by the type of task repetition: same task repetition, procedural task repetition, or content task repetition?
>
> RQ2b: What justification do studies provide for their choice of spacing gaps? (e.g., what is the rationale given, if any, for having 1-day between task repetitions)?

RQ3: How many times do learners repeat tasks in task repetition studies?

> RQ3a: Does the number of task repetitions differ by spacing pattern?
>
> RQ3b: What justification do studies provide for the number of task repetitions? (e.g., what is the rationale given, if any, for having learners repeat tasks three times?).

Method

To identify the relevant literature, we carried out a systematic search following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses method (*PRISMA*; Moher et al., 2009).

Identification

In the identification stage, there are three types of search work in this study: database and web engine search, journal search, and reference search.

The search for studies began with a database search involving five academic databases, including LLBA, MLA, PsycINFO, ERIC via EbscoHost, and ERIC. This was carried out using "task repetition" as a keyword search, following the search method used in Johnson and Tabarai's (2022) systematic review and meta-analysis on task planning and oral L2 production. The Google Scholar search terms were "second language acquisition" AND "task repetition" AND "oral task performance" OR "spoken task performance" and "1996-2023." The database and web engine search were finished on 30 January 2023, with 1,458 studies identified. We then hand-searched the table of contents of eight relevant journals: Applied Linguistics, Language Learning, Studies in Second Language Acquisition, TESOL Quarterly, The Modern Language Journal, Language Teaching Research, Language Teaching, and System. This search was done from 1996 to 2023. It yielded a total of 83 potential studies to be included in the systematic review. We then checked the references of articles that contained comprehensive reviews on L2 perception and production (e.g., Bui & Yu, 2022; chapters included in Bygate (ed.), 2018; Lambert et al., 2017; Li & Rogers, 2021). This resulted in 268 additional studies.

Inclusion Criteria

The full justification of our inclusion criteria is as follows:

1. Task repetition must be the independent variable of interest.
2. We included studies published from 1996 to 2023. 1996 was chosen as a starting point based on Bygate's seminal (1996) study, which was arguably a "watershed moment" that sparked the discussion in the field regarding the potential of task repetition for oral task performance and L2 development. The endpoint reflects the time at which the search for the systematic review was conducted.
3. The report was a published study (academic journal or book chapter) written in English. It would be ideal to include studies reported in languages other than English, but the idea was discounted as it was infeasible for the researchers to include all languages equally.
4. Learners in the study performed tasks at least twice (performed the task and then repeated them at least 1x);
5. Learners repeated the same or similar tasks. This was limited to same task repetition, procedural task repetition, content task repetition, or a combination thereof.
6. The study measured spoken task performance. Tasks performed in other modalities (e.g., written task performance) are also subject to spacing, but arguably subject to different theoretical mechanisms (see, e.g., Manchón, 2014). Because of these theoretical differences, and also to maintain focus of analysis, we elected to include only studies that measured spoken task performance.

Exclusion criteria

We excluded

1. Any study where participants did not repeat a task at least twice. This requirement superseded all other requirements of the study. Numerous studies in the TBLT literature only require their participants to perform the tasks once. This is not a design flaw but can be seen to reflect their research questions. However, these studies are not task repetition studies as a task repetition necessitates repetition of task. Studies that only included a single task performance were thus excluded.
2. Studies where task repetition was not the independent variable of interest. For example, some studies in the corrective feedback literature have included multiple training sessions where tasks are used as part of the training sessions (e.g., Sato & Lyster, 2012). This also included TBLT studies focused on issues related to pre-task planning and task complexity in language development (e.g., Revesz, 2009).
3. Fugitive literature, such as conference proceedings and dissertations (e.g., Patanasorn (2010). In systematic reviews and meta-analyses, whether or not to include unpublished studies is a decision that researchers face. Including unpublished studies may make the findings more comprehensive and robust. However, excluding the findings may aid in future replication attempts (Oswald & Plonsky, 2010). For reasons of improving data quality and also partly based on our knowledge of the literature, where we were confident that our search would likely yield a high number of published empirical studies, we elected to exclude unpublished sources.
4. We excluded studies that did not investigate oral task repetition. This included, for example, written task performance (e.g., Lázaro-Ibarrola & Hidalgo, 2021) and computer-mediated task-based language teaching (e.g., Hsu, 2018)
5. We also excluded from our analyses some studies that could not be classified as either same-task, procedural, or content repetition based on our operationalizations (e.g., Matsumura et al., 2008). This is because these studies manipulated both the content and the procedure of the tasks between repetitions (see our discussion below). In other words, according to our operationalizations, these studies did not fall into any category of task repetition; they performed a new, unrelated task. This represents a key issue as to what constitutes task repetition and how to differentiate between task repetition and when learners perform unrelated tasks.

A total of 1809 studies were then screened for inclusion in the analyses.

We conducted two rounds of screening. The first round of screening was by title and abstract only. This first round focused on the removal of duplicated publications and items that clearly did not meet inclusion criteria (e.g., dissertations, clearly irrelevant topics, e.g., publications related to frequency effects in incidental vocabulary acquisition). All of the items were double-screened by two members of the research team, with 92% agreement. Discrepancies were resolved after discussion. The end of the first round of screening left 358 possible studies for inclusion. The second screen involved inspection of full texts to ensure that they met the inclusion criteria. The final number of eligible studies included for coding was 107.

68/107 (63.5%) studies included sufficient information about spacing details to be included. As noted, this varied from no information provided to vague reporting. We attempted to contact the authors of these studies by email. 14/40 responded. These 14 studies were included in analyses on spacing. The remaining studies were not included in the analyses on spacing gaps (e.g., RQ1 & RQ2) but were included in other analyses where possible. For example, all studies included information on the number of task repetitions (RQ3), so we included their information

in these analyses. Our overarching philosophy was to preserve as much of the data from these 107 studies as possible to address the RQs where possible.

Coding

The coding scheme was developed in an iterative process of coding and revision in a pilot study. The final coding scheme included items related to demographics, identifying task repetition type, spacing (within-session, between-session, or combination), and total number of task repetitions. We also coded for whether researchers provided justifications for their spacing gaps and number of task repetitions and, if so, what their justification was.

1. Operationalization of Task Repetition Types

As noted, studies typically differentiate between same-task repetition (performing the same task with the same content and procedure more than once) and procedural repetition, that is, performing a similar task of the same procedure but with new content. Some also differentiate a third category: content repetition. As noted, content repetition is traditionally operationalized to involve drawing upon the same content knowledge but with modification to the procedure (Patanasorn, 2010). In the case of Patanasorn (2010), this involved using content knowledge to carry out completely different task types (story completion, information exchange, voting for a candidate). In other words, it was an absolute change of the procedure that involved completely new tasks.

In our initial survey of the literature, we found inconsistencies in how task repetition types were defined, operationalized, and implemented in the field. For example, some studies operationalized the same repetition with exactly the same task and exactly the same procedure, whereas other studies included some changes to the procedure. These included, for example, changing the partner (Kim & Payant, 2014), changing the language from L1 to L2 (Pinter, 2007), and changing the venue (Kartchava & Nassaji, 2019). Content repetition, by contrast, included drawing upon the same background knowledge but with changes to the procedure. However, this involves an absolute change to the procedure in the form of a completely new task. To reconcile this, we adopted a new conceptualization of content repetition from previous literature. This involved a nuanced operationalization where content repetition can represent any scale of change to the procedure, whereas same task repetition involves an exact repetition of both the task and procedure. We do this on theoretical grounds in that having distinct operationalizations of different task types without overlap may provide a more sound basis for testing theoretical frameworks for L2 speech production in task repetition research.

2. Operationalization of Spacing Framework

For this study, we needed a systematic framework for describing the spacing intervals in task repetition research. We drew upon existing methodological terminology from the input spacing literature to achieve this (see Table 1 for an overview of the three basic task repetition methodological designs).

- *Within-session spacing*: Refers to situations when spacing is manipulated within a single training session or class period (see description above). For example, a learner performs a spot-the-difference task and immediately performs the same task again. There may also be some time delay between the tasks, but importantly the tasks take place within the same training session or class period.

- *Between-session spacing*: Refers to situations where spacing is manipulated across training sessions or class periods. For example, the learner performs a spot-the-difference task and returns to the laboratory to repeat the same task a week later. Alternatively, the learner might repeat the same task in class a week later, for example.

The difference between within designs and between designs is temporal. Within designs take place within a single session, whereas between designs are spread out over multiple sessions. More complex designs combine both within and between designs. Based on our knowledge of the literature, we were aware that some task repetition studies repeat tasks both within a single training session (within-session spacing) and across multiple training sessions (between-session) spacing).

- *Within-between spacing*: Spacing is manipulated both within a single training session (or class period) and across training sessions (or class periods). For example, a learner might repeat a narrative task three times within a single training session and then return to the lab on a separate day to repeat the narrative task three more times (e.g., de Jong & Perfetti, 2011).

Table 1. Examples of spacing paradigms.

Task Repetition Design	Spacing Designs
Within-session design	A(immediate)A A(10 minutes)B
Between-session design	A-(3 days)-A A-(7 days)-B
Within-between session design	AAA-(3 days)-AAA ABCD-(7 days)-ABCD

We did not code for the exact intervals in within-session spacing. Based on our initial piloting, it was not feasible to code for within-spacing studies due to reporting issues. We focus our reporting of spacing intervals on between-session intervals. Where feasible, we discuss within-spacing intervals in within-between studies.

Approximately 25% of the studies were double-coded in entirety by members of the research team (28/107). Interrater agreement was 96.4%. Discrepancies were resolved by discussion. The remaining studies were coded individually by a member of the research team.

Findings

RQ1: What spacing paradigms, i.e., within-session, between-session, or a combination of within and between session (*within-between*), have been investigated in task repetition research?

In terms of spacing designs, task repetition studies have used relatively similar within-session, between-session designs, and within-between designs. There have also been a number of idiosyncratic designs. We present examples to illustrate below (Table 2).

A number of task repetition studies have used within-session spacing, reflecting data collection in a single laboratory or classroom setting. In these studies, the amount of time between task repetition varied. Some studies reported that subsequent repetitions were carried out immediately, suggesting no break between repetitions (e.g., Bei, 2013). Other studies reported short breaks between repetitions (e.g., Sun & Revesz, 2021). Others reported longer amounts of time occupied by filler tasks (e.g., Strachan et al., 2019).

Table 2. Examples of Task Repetition studies that used within spacing designs

Study	Spacing Designs	Time Between Repetitions
Bei (2013)	AAAA	Immediate
Sun & Revesz (2021)	AAA	5 minutes
Strachan et al. (2019)	AB	15mins+

Similarly, task repetition studies have also utilized between-session spacing designs, in which they report at least a one-day gap between the initial task performance and subsequent task performances. Examples are presented in Table 3. These ranged from a single gap design where the task was performed twice in total (1 repetition) to multiple gap designs where the task was performed multiple times. In designs where tasks were performed multiple times, the general trend was that the length of the spacing gaps was uniform within a study (e.g., if there was a 1-day gap between the first and second performance, then generally a 1-day gap would also be used between the other task performances). However, there were studies where the gaps were not held constant, though it was often the case that this change was not explicitly justified (see RQ2b).

Table 3. Examples of Task Repetition studies that used between spacing designs

Study	Spacing Designs	Time Between Repetitions
Khezrlou (2019)	AAA ABC	1 day
Ahmadian (2011)	AAAAAAAAAAAB	2 weeks
Azkarai et al. (2020)	AA	3 months

Other studies utilized what we have termed within-between designs, repeating tasks multiple times within a single experimental or classroom session and then repeating this across multiple training episodes. These have included studies examining blocked versus interleaved task repetition (e.g., Suzuki, 2021), or more commonly for fluency development (e.g., de Jong & Tillman, 2018; Tran & Saito, 2021). Table 4 presents some examples of these designs.

Table 4. Examples of Task Repetition studies that used within-between spacing designs

Study	Spacing Designs	Between-spacing
Suzuki (2021)	AAA-BBB-CCC ABC-ABC-ABC	1 day
Tran & Saito (2021)	AAA-BBB-CCC	3 days
Bygate (2001)	AB-BC-CD	2 weeks

RQ2a: Do these spacing gaps differ by the type of task repetition: same task repetition, procedural task repetition, or content task repetition?

As noted, we only report on between-session spacing gaps. Studies that used within-session spacing (whether as a standalone within-session design or within-between design) often reported "immediate" repetition or did not reliably report the amount of time between repetitions.

The range of spacing gaps in between-session designs ranged from immediate repetition to 7 months in length (Table 5). To calculate these, each spacing gap was calculated independently.

Table 5. Between-session spacing by type of task repetition

	Task Repetition	Procedural Repetition	Content Repetition
Total groups (N)	50	29	16
Number of groups that did not report spacing information	3	11	4
Total Number (N) of groups for Analysis	47	19	12
Spacing range	immediate ~ 3 months	30 seconds ~ 7 months	immediate ~ 4 weeks
Mean (SD)	13.53 (20.55) days	23.58 (46.59) days	3.38 (5.29) days
Median	7 days	7 days	1 day

Overall, the spacing gaps among task repetition studies within a single training session vary from "immediate" to "7 months". Procedural repetition showed the largest mean of all task repetition types. However, the high standard deviation and the fact that the median value is the same as the task repetition group suggests that outlying values may be the cause of the higher mean value in this cause. Content repetition groups show significantly lower spacing gaps both in terms of average mean and median relative to task repetition and procedural repetition.

Table 6. Within-Between spacing by type of task repetition

	Within-between Task Repetition		Within-between Procedural Repetition		Within-between Content Repetition	
Total groups (N)	11		20		3	
Number of groups that did not report spacing information	3	7	4	16	2	3
Total Number (N) of groups for Analysis	6	2	13	1	1	0
ISI	between ISI	within ISI	between ISI	within ISI	between ISI	within ISI
Spacing range	1 day ~ 2 weeks	immediate	immediate ~ 2 weeks	7 days	30 minutes	/
Mean (SD)	4.09 (3.95) days	immediate (0)	5.67 (4.10) days	7 (0) days	30 (0) minutes	/
Median	3 days	immediate	7 days	7 days	30 minutes	/

With regard to within-between designs (Table 6), the results showed that task repetition studies conducted over several training sessions have similar spacing gaps to procedural repetition studies conducted over only one training session, which was from "immediate" to "2 weeks". There were not enough content repetition groups for a meaningful comparison.

RQ2b: What justification do studies provide for their choice of spacing gaps? (e.g., what is the rationale given, if any, for having 1-day between task repetitions).

Thirty-four studies provided some form of an a priori justification of their spacing intervals, with some studies giving multiple justifications (Table 8). The rationale given by fifteen of these studies related to reasons of fitting with class timing or curriculum. This can perhaps be attributed to the quasi-experimental nature of many of these studies, as they collected data in authentic classroom settings and needed to work within the constraints of the curriculum. For instance, researchers collected data from classes that had regular class meetings, regularly scheduled exams, etc, all of which the researchers had to plan and negotiate around. Nineteen studies cited previous research as part of their justification of spacing schedules. These studies varied in terms of the comprehensiveness of their justification. At one end of the spectrum, studies adopt the same spacing intervals as previous studies without making explicit their motivation for doing so. Most commonly, these studies state "Following…, as in "Following Bygate (2001)", or "similar to previous literature" or "in line with previous literature". Other researchers highlighted the lack the consistency in the literature and the difficulty in identifying meaningful patterns. For example, some researchers made clear that it was not possible to justify spacing intervals based on previous literature. Ahmadian and Tavakoli (2011) made explicit that no regular spacing intervals could be identified based on their survey of the literature. At the other end of the spectrum, a few studies discuss the spacing intervals of relevant previous research and justify their spacing schedules in light of theoretical models (see, e.g., Khezrlou, 2019; Wang, 2014).

Table 7.

Justifications for spacing intervals n=34 total	
n	Justification
15	Class schedule / fit into curriculum
18	Link to previous literature
4	Theory
1	Based on pilot study results
1	Practicality: to give participants a break
1	Pedagogical relevance

RQ3: How many times do learners repeat tasks in task repetition studies (number of task iterations)?

The third research question explores the number of times learners repeat tasks in task repetition studies. For task repetition, procedural repetition, and content repetition, the total number of repetitions range from 2 to 11 overall (Table 9). For studies operating in the within-between paradigm, the range is from three to fifty repetitions, though appears to be skewed higher by outlying values, with the typical range of three to nine repetitions.

RQ3a: Does the number of task iterations differ by spacing pattern

Table 8. Number of task repetitions by task repetition type

	Task Repetition	Procedural Repetition	Content Repetition	Within-between Task Repetition	Within-between Procedural Repetition	Within-between Content Repetition
Total Number (N) of groups for Analysis	50	29	16	11	20	3
Types of Numbers of Task Performances	2,3,4,6,11	2,3,4,5,6	2,3,5,6,10	3,4,6,8,9	4,6,7,8,9,10,12,24,50	4,6,9
Mean (SD)	2.61 (1.43)	3.07(1.03)	3.75 (2.08)	7.5 (2.37)	10.45(10.23)	6.33 (2.52)
Median	2	3	3	8.5	8	6

Data do not suggest differences in the number of task repetitions across task repetitions, procedural repetition, or content repetition in terms of mean or median values. Similarly, the median values for the within-between task repetition and procedural repetition groups were also similar, suggesting that the number of repetitions performed by the learners within these groups was similar.

RQ3b: What justification do studies provide for the number of task repetitions? (e.g., what is the rationale given, if any, for having learners repeat tasks three times?).

Six studies provided justifications for the number of task repetitions. All six of these justified the number of task repetitions in terms of previous studies.

Three of these studies did so by referring by making reference to the number of repetitions carried out in previous research. At one end of the spectrum, studies adopt the same spacing intervals as previous studies without making explicit their motivation for doing so. Most commonly, these studies state "Following…, as in "Following Bygate (2001)". Ahmadian and Tavakoli (2011) reviewed the literature and noted that there appeared no regular pattern in terms of spacing or number of repetitions. The remaining two studies cited this study in part as their justification.

Discussion

The discussion is structured around reporting issues, methodological issues, research findings, and suggestions for future research practices. Reporting issues are discussed by presenting the status quo, identifying pitfalls, recommending solutions and exemplifying practices, and suggesting future directions. The same is repeated for methodological issues. These discussions are followed by suggestions with a view to inform future researchers.

Reporting issues

The most substantial issue impacting the results of this study is reporting. A key issue in this regard is that information related to intersession intervals (ISIs), that is, the amount of time between task repetitions, has often gone unreported, or has been reported in vague terms. With regard to the current paper, 4/16 (25%) of the studies coded as content repetition did not make available their spacing information, 11/29 (38%) of procedural repetition studies, compared to 3/51 (6%) of the exact repetition studies. This poses a potential survivorship bias that potentially influences the results. It is also worth noting that these results are only for the between-spacing gaps; our initial survey of the within-session spacing was far less positive; it is simply not reported. As an illustration, for 34 within-between studies in Table 7, only 3 included sufficient information regarding within-session spacing to be included for analysis. This lack of reporting is perhaps not surprising, given that it is only recently that task repetition researchers have more consistently begun to take spacing intervals into account as part of their experimental designs. Similar reporting issues were seen for the other variables of interest coded for in this study, such as the number of task repetitions.

Research Findings

Any discussion of the research findings must be viewed in light of the reporting issues discussed above. With this in mind, task repetition studies have adopted a number of different spacing patterns as part of their methodological designs. These include both within-session spacing, that is, repeating tasks multiple times within a single experimental session; between-session spacing, that is, repeating tasks across multiple training or classroom sessions; and

within-between spacing, that is, a combination of within- and between-spacing, where tasks are repeated both within a single session and across multiple sessions. Different types of task types (Task repetition, procedural repetition, and content repetition) have been used in combination with each of these patterns of spacing. Descriptive statistics suggest that no differences exist between these types of spacing. However, it is important to highlight that no inferential statistics were carried out, so it is not possible to say whether statistical significance exists between these types. Future researchers may consider meta-analytic work to investigate these issues further, though reporting issues (as discussed above) would need to be taken into consideration.

With regard to the number of repetitions, for between-spacing designs, the results indicated that task repetition studies have investigated a wide range of number of repetitions, from two to 11 in number. The median value across these task types indicated that studies commonly repeated tasks two or three times. For within-between designs, the results showed a much wider range, from three repetitions to 50, though, as noted, these results were likely skewed by an outlying value. The median values suggest that the most common number of repetitions were around eight repetitions in total. Like the descriptive data above regarding spacing intervals between training sessions, the data here do not suggest differences in the number of repetitions between different task repetition types. However, as mentioned above, no inferential statistics are available to support this claim, so this interpretation must be taken with caution.

What appears to be a much-replicated pattern in the within-between paradigm is one utilized within task repetition studies examining the impact of task repetition on fluency development (e.g., de Jong & Perfetti, 2011). In this paradigm, tasks are repeated three times within individual training sessions (within session repetition) and then across three training sessions (between session repetition), and may include additional tasks operationalized as pretests and posttests. Variants may include repeating the same task within individual training sessions, then performing a different task of the same task type in subsequent training sessions (AAA-BBB-CCC), or mixing task types within individual training sessions and repeating this procedure across subsequent training sessions (ABC-ABC-ABC).

Methodological Considerations

We have a few methodological recommendations for future research. The first is the exact methodological reporting of spacing intervals. As noted, in the task repetition literature there is tentative evidence that intersession spacing may influence task performance (Rogers, 2023). From our findings, it appears that the task repetition literature is inconsistent in the reporting of such intervals. Vague reporting, such as stating that tasks were completed three times within two weeks, should be avoided. If the situation occurs where tasks were completed at irregular spacing intervals, then this can be reported as such. Secondly, task repetition researchers should consider more explicit justifications for spacing intervals and number of repetitions. As indicated by our review, spacing intervals are commonly justified on grounds of following previous research (e.g., “following Bygate, 1996”). At this juncture, the practice of basing spacing intervals on previous research without firm justification (e.g., “following Bygate, 2001”) seems unlikely to move the field forward. Based on our review, theory-driven approaches appear rare (see, e.g., Khezrlou, 2019; Wang, 2014, for exceptions). This is unfortunate given that spacing and repetition may be linked with theoretical frameworks for task repetition (e.g., Levelt, 1989). Thus, it has the potential for developing testing theoretical frameworks and the development of new research questions to be tested empirically.

As pointed out by an anonymous reviewer, although we did not code for task type as a methodological feature, it appears that most of the studies included in our sample examined monologic narrative tasks. To speculate to causes of this apparent pattern, this may be partly due to readily available validated materials (e.g., Heaton, 1975), which facilitates data collection. A second reason may also be again where studies have justified their methodological decisions (materials and procedures) based on previous studies (e.g., "following Ahmadian & Tavakoli, 2011"), though, again, this is only speculation on our part. It is unclear to what degree that this skewed in task type is constrained to our sample, however a wider variety of task types may be required to substantiate the generalizability of the current findings.

More interestingly, as the sample does appear to be skewed, this raises an interesting discussion question of whether task type may bear influence on the choice of spacing intervals or, potentially, on even whether to repeat the task. For example, as noted, when repeating monologic narrative tasks under short spacing conditions, there is potential for the content in subsequent repetitions to be similar in subsequent repetitions. This may lead to pedagogical and/or methodological decisions to avoid immediate task repetition and to opt for relatively longer (ISI) intervals. In dialogic tasks, such as discussion tasks, this may not necessarily be the case, in particular when discussion partners change between tasks, as the content of the conversations may change. When using such tasks, educators and researchers might not necessarily feel the need to avoid immediate task repetition or task repetition with short ISI intervals. Studies that have examined teachers' perceptions of task repetition have noted that teachers have reported fears of implementing task repetition due to worries of tasks appearing contrived, despite acknowledging potential benefits of repeating tasks for students' language development (Foster & Hunter, 2016). In this case, the shorter ISI intervals may lead to increased fears on the part of the teachers towards the unrealistic nature of task repetition. The impact of task type and its potential interaction with spacing intervals was beyond the scope of our current review. However, this interaction appears to be an open empirical question for future research to explore.

Conclusion

This synthetic review examined the spacing intervals, number of task interactions, and justifications thereof in task repetition research. Despite the fact that there has been extensive empirical research, the findings were limited by reporting issues. Studies to date have adopted a mix of within, between, and within-between spacing methodological designs. The within-between methodological designs, in particular, have been utilized extensively for research examining the use of tasks for fluency development. The findings also illustrate that task repetition studies have examined a wide range of repetition numbers. In terms of justifications, justifications of spacing intervals were rare but tended to be linked to the nature of quasi-experimental data collection, in particular in linking data collection to class schedules or other considerations related to curriculum. Theoretical justifications were rare.

Overall, the findings suggest that, despite the lack of attention to spacing exhibited by some studies, the studies that have utilized a systematic approach to spacing intervals have achieved significant results. This is best illustrated by studies utilizing within-between methodologies to examine the impact of L2 fluency (e.g., de Jong & Perfetti, 2018). Despite the limited research, such findings are theoretically promising and open doors for future research to systematically examine the interaction between spacing and task performance.

References

Ahmadian, M. J. (2011). The effect of 'massed' task repetitions on complexity, accuracy and fluency: Does it transfer to a new task? *The Language Learning Journal*, *39*(3), 269–280. https://doi.org/10.1080/09571736.2010.545239

Ahmadian, M. J., Mansouri, S. A., & Ghominejad, S. (2017). Language learners' and teachers' perceptions of task repetition. *ELT Journal*, *71*(4), 467-477.

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, *15*(1), 35-59. https://doi.org/10.1177/1362168810383329

Azkarai, A., García Mayo, M. P., & Oliver, R. (2020). The effect of task repetition on the patterns of interaction of ESL children. ITL - International Journal of Applied Linguistics, 171(1), 90–112. https://doi. org/10.1075/itl.17034.azk

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, *4*(5), 316–321. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Bei, X.G. (2013). Effects of immediate repetition in L2 speaking task: A focused study. English Language Teaching, 6, 11–19. http://dx.doi.org/10.5539/elt.v6n1p11

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Macmillan Heinemann

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Pearson Longman. https://doi.org/10.4324/9781315838267

Bygate, M. (2018). Introduction. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 1–15). John Benjamins.

Bygate, M., & Samuda, V. (2005). Integrative planning and the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–74). John Benjamins. https://doi.org/10.1075/lllt.11

Bygate, M., Skehan, P., & Swain, M. (Eds.). (2001). Researching pedagogic tasks: Second language learning, teaching, and testing. Longman.

Bui, G., Ahmadian, M. J., & Hunter, A.-M. (2019). Spacing effects on repeated L2 task performance. *System*, *81*, 1–13. https://doi.org/10.1016/j.system.2018.12.006

Bui, G., & Yu, R. (2021). Differentiating task repetition from task rehearsal. In S.N.P. Rao & L. Mukhopadhyay (Eds.), Task-based language teaching and assessment: Contemporary reflections from across the world (pp. 119–137). Springer.

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 53, pp. 63–147). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(10)53003-2

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. Language learning, 61(2), 533-568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

DeKeyser, R. (2018). Task repetition for language learning: A perspective from skill acquisition theory. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 27–41). John Benjamins.

Ellis, R., & Shintani, N. (2013). Exploring language pedagogy through second language acquisition research. Routledge.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). Task-based language teaching: Theory and practice. Cambridge University Press. https://doi.org/10.1017/9781108643689

Foster, P., & Hunter, A.M. (2016). When it's not what you do but the way that you do it. In: B. Tomlinson (Ed.), *SLA research and materials development for language learning* (pp. 280–298). London: Routledge.

Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, *20*(3), 321–340. https://doi.org/10.1177/1362168815570142

Hanzawa, K., & Suzuki, Y. (2023). How do learners perceive task repetition? Distributed practice effects on engagement and metacognitive judgment. *The Modern Language Journal*, *107*(2), 451-478. https://doi.org/10.1111/modl.12843

Heaton, J.B. (1975). Beginning composition through pictures. Longman.

Hsu, H. C. (2017). The effect of task planning on L2 performance and L2 development in text-based synchronous computer-mediated communication. Applied Linguistics, 38(3), 359-385. https://doi.org/10.1093/applin/amv032

Johnson, M. D., & Abdi Tabari, M. (2022). Task planning and oral L2 production: A research synthesis and meta-analysis. *Applied Linguistics*, *43*(6), 1143-1164. https://doi.org/10.1093/applin/amac026

Kartchava, E., & Nassaji, H. (2019). The role of task repetition and learner self-assessment in technology-mediated task performance. Itl - International Journal of Applied Linguistics, 170(2), 180–203. https://doi.org/10.1075/itl.19013.kar

Khezrlou, S. (2019). Task repetition and corrective feedback: The role of feedback types and structure saliency. English Teaching and Learning, 43(2), 213–233. https://doi.org/10.1007/s42321-01900025-2

Kim, Y., & Payant, C. (2014). A pedagogical proposal for task sequencing: An exploration of task repetition and task complexity on learning opportunities. In M. Baralt, M. R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language* learning (pp. 151–177). London: Bloomsbury Academic.

Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? System, 41(3), 829–840. https://doi.org/10.1016/j.system.2013.08.005

Kobayashi, M. (2022). The distributed practice effects of speaking task repetition. *International Journal of Applied Linguistics*, *32*(1), 142–157. https://doi.org/10.1111/ijal.12409

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum.

Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, *39*(1), 167–196. https://doi.org/10.1017/S0272263116000085

Lázaro-Ibarrola, A., & Hidalgo, M. Á. (2021). Give me a second chance: Task repetition and collaborative writing with child EFL learners. Language Teaching for Young Learners,3(2), 275-299. https://doi.org/10.1075/ltyl.20009.laz

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

Li, S., Hiver, P., & Papi, M. (2022). Individual differences in second language acquisition: Theory, research, and practice. In S. Li, P. Hiver, &M. Papi (Eds.), The Routledge handbook of SLA and individual differences (pp. 3–34). Routledge.

Li, P., & Rogers, J. (2021). *Task repetition in language learning: A bibliography*. OSF Preprints. https://doi.org/10.31219/osf.io/rk7ag

Long, M. H. (2015). Second language acquisition and task-based language teaching. Wiley-Blackwell.

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, *30*, 85-110. doi:10.1017/S0267190510000115

Mackey, A. (2020). Interaction, feedback and task research in second language learning: Methods and design. Cambridge University Press.

Manchón, R. M. (2014). The distinctive nature of task repetition in writing. Implications for theory, research, and pedagogy. *ELIA, 14, 13-41*. https://doi.org/10.12795/elia.2014.i14.02

Matsumura, M., Kawamura, K., & Affricano, A. (2008). Narrative task-type repetition and changes in second language use in a classroom environment: A case study. *横浜国立大学教育人間科学部紀要. I, 教育科学= Journal of the Faculty of Education and Human Sciences, Yokohama National University. The educational sciences*, *10*, 125-145.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Patanasorn, C. (2010). Effects of procedural, content, and task repetition on accuracy and fluency in an EFL context. Doctoral dissertation, Northern Arizona University, Flagstaff, AZ, USA. (ERIC document reproduction service No. ED519953.)

Pinter, A. (2007). Some benefits of peer–peer interaction: 10-year-old children practising with a communication task. *Language teaching research*, *11*(2), 189-207. https://doi.org/10.1177/1362168807074604

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, *31*(3), 437–470. doi:10.1017/S0272263109090366

Rogers, J. (2023). Spacing effects in task repetition research. *Language Learning*, *73*(2), 445–474. https://doi.org/10.1111/lang.12526

Sample, E., & Michel, M. (2014). An exploratory study into trade-off effects of complexity, accuracy, and fluency on young learners' oral task repetition. TESL Canada Journal, 31, 23–46. https://doi.org/10.18806/tesl.v31i0.1185

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, *34*(4), 591-626. doi:10.1017/S0272263112000356

Strachan, L., Kennedy, S., & Trofimovich, P. (2019). Second language speakers' awareness of their own comprehensibility: Examining task repetition and self-assessment. Journal of Second Language Pronunciation, 5(3), 347-373. https://doi.org/10.1075/jslp.18008.str

Sun, B., & Révész, A. (2021). The effects of task repetition on child EFL learners' oral performance. Canadian Journal of Applied Linguistics, 24(2), 30-47. https://doi.org/10.37213/cjal.2021.31382

Suzuki, Y., Eguchi, M., & De Jong, N. (2022). Does the reuse of constructions promote fluency development in task repetition? A usage-based perspective. *TESOL Quarterly*, *56*(4), 1290-1319. https://doi.org/10.1002/tesq.3103

Suzuki, Y., & Hanzawa, K. (2022). Massed task repetition is a double-edged sword for fluency development. *Studies in Second Language Acquisition*, *44*(2), 536–561. https://doi.org/10.1017/s0272263121000358

Tran, M. N., & Saito, K. (2021). Effects of the 4/3/2 activity revisited: Extending Boers (2014) and Thai & Boers (2016). Language Teaching Research, Early View. https://doi.org/10.1177/1362168821994136

Wang, Z. (2014). Online time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), Processing perspectives on task performance (pp. 27–62). Amsterdam, The Netherlands: John Benjamins.

Wiseheart, M., Kim, A. S. N., Kapler, I. V., Foot-Seymour, V., & Kupper-Tetzel, C. E. (2019). Enhancing the quality of student learning using distributed practice. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 550–584). Cambridge University Press.