Learners' reflective practice between the repeated performances of tasks: Effects on second language development

Sima Khezrlou[1]

[1]*Independent Researcher*

Abstract This study attempted to explore the role of reflection in the accurate use of the English regular past tense structure using task repetition. Thirty-one learners were assigned into two conditions: task repetition only (TR) and task repetition with self-reflection (TR+SR). Both groups repeated an oral narrative task two times and then carried out a new task of the same type (i.e., another oral narrative task). However, only the TR+SR learners were engaged in self-reflection through responding to a questionnaire developed for the purpose of this study. Results revealed that learners' reflection on their first task performance helped them notice the gap between their existing and target structure use as attested by their significantly high scores in the repeated tasks as well as the new task. The results therefore indicate the potential of reflective practice as an effective intervention strategy between repeated performances of the same task in terms of accuracy.

1 Introduction

Interest in task-based language teaching (TBLT), which uses meaning-focused activities, also called "tasks", as its core unit of teaching has been mounting since the 1980s. Particularly, classroom-based TBLT research has been interested in how implementing tasks in different ways may influence learners' task performance and second language (L2) development (Khezrlou, 2019b, 2020a, 2020b). One task implementation variable which has received a considerable amount of attention in recent years is task repetition (TR) (Bygate, 2001; Ellis, 2019). It is claimed that repeating the same or a slightly different task is beneficial for learners by freeing up their cognitive resources that are mainly focused on meaning than form during the first performance of a task (Samuda & Bygate, 2008). Although studies have provided evidence in support of TR (see Khezrlou (2021)

for a review), to date, there is a dearth of research on how the provision of intervention between the first and second performance of a task – for instance, by asking learners to reflect on the first performance- affects L2 development.

To reap the benefits of TR as far as L2 development is concerned, it may therefore be vital to demonstrate whether the L2 gains as a result of repeated task performance can be carried over to the new tasks (Ellis, 2009, 2019). To achieve this, as Ellis (2019) argues, some type of form-focused intervention between the first and following performance(s) of a task is needed. As one type of post-task focus on form, reflection has been suggested by Ellis, Skehan, Li, Shintani and Lambert (2020) as an intervention option. Nevertheless, up to now, there have been few attempts to explore the potential of reflection as a form-focused strategy in the TBLT literature in general and TR literature in particular. The present study's objective was therefore to bridge this gap by exploring whether encouraging learners to reflect on their performance of a task could enhance L2 development.

2 Review of the literature

2.1 Task repetition

TR refers to when learners carry out a task more than once and the first performance is viewed as a preparation for subsequent performances (Ellis, 2009). There are diverse types of repetition, such as repetition of the same task procedure with same content (exact task repetition), same procedure with different content (procedural repetition) and same content with different procedure (content repetition) (Patanasorn, 2010). In the present study, TR is operationalized as repeating the same task procedure and content. The utility of TR has been clarified based on Skehan's (1998) trade-off hypothesis which claims that because of the limited capacity of working memory, learners are not able to concurrently focus on both form and meaning and thus they prioritize one over the other. But, as learners repeat tasks, they are likely to divert their attention to form and produce language that is more accurate and successful (Bygate, 2001). Improvements in output as a result of TR may also be accounted for with reference to Levelt's (1989) model of speech production. Speech production, as conceptualized by this model, entails three stages: conceptualization, formulation, and articulation. In the conceptualization, the content of the message is planned. Provided that the content remains unchanged in the repeated task, then less time would be needed to decide what to say when the production is done again. This will thus enable the learner to dedicate more processing resources to the retrieval of the language to encode the message (formulation) and deliver it (articulation), leading to higher linguistic accuracy and complexity.

To date, numerous studies have tested these claims following three main research streams (see Table 1 for a summary of TR studies): the comparison of TR types and

their effectiveness (e.g., Carver & Kim, 2020; Gass et al., 1999; Khezrlou, 2019b; Lynch & MacLean, 2000), the transfer of TR effects to a new task (e.g., Kim & Tracy-Ventura, 2013; Sheppard & Ellis, 2018), and the impacts of intervention or "enhanced repetition" as Lynch (2018) puts it on subsequent task performances (e.g., Hawkes, 2012; Hsu, 2019; Kartchava & Nassaji, 2019; Khezrlou, 2019c, 2020b; Sheppard, 2006 as cited in Ellis, 2009; Sheppard & Ellis, 2018). With respect to the first area, evidence in favor of repeating the same tasks in enhancing fluency and complexity has been furnished, with controversial results concerning accuracy. Regarding the transfer of TR effect, previous research (Bygate, 2001; Gass et al., 1999; Patanasorn, 2010) has demonstrated that the effects of repeating the same task may not be carried over to a new task. Yet, some recent studies (e.g., Khezrlou, 2021; Sheppard & Ellis, 2018) showed the opposite.

"The justification for task repetition as a pedagogic device must lie in whether its effects transfer to a new task (i.e., impacts on development)" (Sheppard & Ellis, p. 190, 2018). To achieve this, some form of intervention to direct learners' attention to language form is needed (Ellis, 2019). As highlighted by Kartchava and Nassaji (2019), when learners are subjected to intervention followed by TR and practice, their attention would be guided towards language that can bring about remarkable enhancement in their subsequent output. The pioneering study by Hawkes (2012) explored the role of direct instruction and follow-up practice of grammatical and pragmatic structures and vocabulary on Japanese learners' repeated task performance. It was found that learners could focus on form in their second enactment, resulting in enhancements in the use of lexis, grammar, and partially, pronunciation. In Sheppard's (2006, as cited in Ellis, 2009) study, Japanese learners received corrective feedback after their first task performance which led to improvements in fluency, complexity, and notably in accuracy. The use of stimulated recall after the first task performance in Sheppard and Ellis's (2018) study, however, was unsuccessful in focusing learners' attention on form. In another study, Khezrlou (2021) concluded that the provision of explicit instruction between performances of the same task resulted in learners' explicit and delayed implicit knowledge development. Lastly, Khezrlou (2019b) looked into the role of task repetition and procedural repetition with input-providing or output-prompting oral corrective feedback in the development of regular and irregular past tense structures. Results underscored the superiority of output-prompting corrective feedback regardless of repetition type or linguistic structure.

Overall, these studies lend support to Lynch's (2018) argument that intervention is "enhanced repetition" referring to "the opportunity to engage in some sort of cognitive activity related to the first run" (p. 196). Nevertheless, this area of research within the TR literature has only recently begun to attract widespread attention from the TBLT researchers. Not until recently have there been calls to explore the role of involving learners' in reflecting on their initial task performance and the resulting effects on subsequent task performances (see Ellis et al., 2020). Hence, the present study provides insights into the added value of learner reflection in repeated task performance.

Table 1 Summary of the task repetition studies

Study	Participants	Measures	Task	Number of repetitions and interval between	Intervention	Results
Bygate (2001)	48 ESL learners	Fluency (unfilled pauses per t-unit), complexity (words per t-unit)	A narrative and an interview task	Twice with a 10-week interval	No	Increase in fluency and complexity in repeated task compared to those in new tasks of the same type
Carver and Kim (2020)	42 EFL learners	Pre-test and post-test (picture description test)	Collaborative writing tasks	Three times with 7–2 day intervals	No	Both content repetition and procedural repetition increased accurate French structure production
Gass et al. (1999)	103 students	Lexical sophistication (number of advanced words per total words), accuracy (correct use of *ser* and *estar*)	Same and slightly altered narrative tasks	Four times with 2–3 and one-week intervals	No	Increase in lexical complexity and partial accuracy in use of *estar* in same but not in new task
Hawkes (2011)	12 EFL learners	English linguistic forms use and correction	Interactive speaking tasks	Twice with no interval	Consciousness-raising activities	Increase in self-correction of lexical and grammatical forms
Hsu (2019)	39 EFL learners	Multiples measures of complexity, accuracy and fluency	An oral narrative task	Three times with a one-week interval	Post-task transcription	Post-task transcription increased both global and structure accuracy and was carried over to a new context
Kim and Tracy-Ventura (2013)	36 EFL learners	Multiples measures of complexity, accuracy and fluency	Information-exchange tasks	Three times with a one-day interval	No	Procedural repetition increased syntactic complexity; both task and procedural repetition increased use of linguistic features, neither group increased fluency
Kartchava and Nassaji (2019)	52 EAP learners	An assessment rubric	A five-minute oral presentation	Twice with a four-week interval	Reflection and self-assessment	Overall L2 performance development
Khezrlou (2019b)	74 EFL learners	Obligatory context analysis	Monologic oral narrative tasks	Three times with a one-day interval	Recast and clarification request	Clarification request increased knowledge of the regular and irregular past tense
Khezrlou (2021)	26 EFL learners	Explicit knowledge (untimed grammaticality judgment test), implicit knowledge (elicited imitation test)	Monologic oral narrative tasks	Three times with a one-day interval	Explicit instruction	Explicit instruction led to immediate and delayed explicit knowledge and delayed implicit knowledge gains
Lynch and Maclean (2000)	14 learners in an ESP context	Fluency (number of words per minute), accuracy (correct English structure use)	Poster carousal task (Dialogic)	Repeating the same question and answer sequence for 6 times with a 3-min interval	No	Increase in accuracy and fluency
Patanasorn (2010)	92 EFL learners	Fluency (number of syllables per minute), accuracy (past simple tense)	Decision-making information-exchange tasks	Three times with a one-day interval	No	Procedural repetition increased accuracy in new task; content repetition increased fluency; exact repetition did not have an impact on new task

Table 1 Summary of the task repetition studies (*cont.*)

Study	Participants	Measures	Task	Number of repetitions and interval between	Intervention	Results
Sheppard (2006)	Undergraduate learners	Fluency (pruned syllables per minute), complexity (grammatical phrases per t-unit and lexical richness), accurate linguistic form use	A monologic oral narrative task	1 repetition after an undetermined period	Input or feedback	Task repetition increased complexity, fluency and minimally accuracy; input between repetition increased CAF; feedback between repetition increased CAF the most
Sheppard and Ellis (2018)	40 EFL learners	Complexity (phrases per t-unit), accuracy (percentage of error free clauses), fluency (pruned words per minute)	A monologic oral narrative task	Three times with no interval between the first and second performances and a two-week interval between the second and third and fourth performances	A stimulated recall session	Both the TR and TR plus stimulated recall groups enhanced fluency, but not in the forth/new performance. Both groups failed to show any improvement regarding complexity and accuracy over time

2.2 Reflection

In reflective learning, reflection refers to a process that is beyond thinking and signifies a critical consideration of an issue which entails increased consciousness (Kellenberg et al., 2017). It is commonly viewed as a precious attribute likely to foster learning through deliberate and perceptive thinking about former experiences (Moon, 2004). These accounts of reflection, then, highlight it as the key to learning from experience that can be used in new contexts. Hence, reflection is generally linked to experiential learning which considers the experience as the nucleus of the learning process (Jarmon et al., 2009). The experience-based nature of reflection is represented in Kolb's (1984, 2014) four-phase model (see Figure 1). This model is built on a concrete experience (phase 1) followed by the perception of and reflection on this experience (phase 2) which leads to a more distinctive understanding of the situation (phase 3) and subsequently helps the individual to put knowledge into practice in new contexts. Applied into the context of L2 learning, it is argued that since the reflective learning practice promotes learners' constant reflection on their past task performance and language use, this consecutively conscious reflection could augment the task performance and eventually advance learners' attention to form (Dao et al., 2020).

The significant role of reflection in different educational contexts is well established (see Farrell, 2011), but its positive effects in the TBLT literature have not been underlined until recently. Based on the results of two case studies, Lam (2018), for instance, underscores the importance of self-reflection in Showcase Portfolio Approach in helping learners understand "where they are, where they want to go and what is next in their writing development" (p. 230). Put differently, reflection upon the showpiece dossiers enables learners to become aware of their writing standards and attempt to bridge the
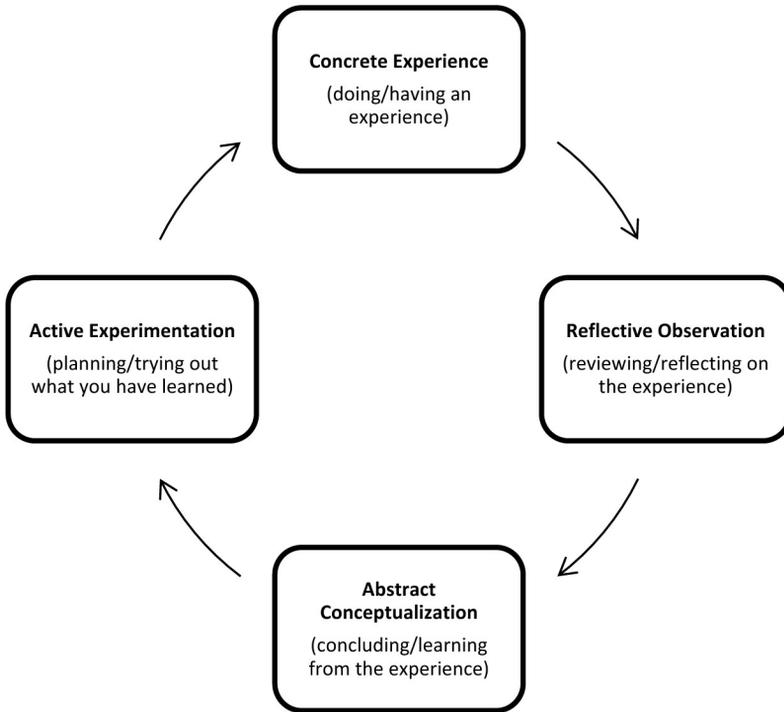
Figure 1 Kolb's (1984) reflection model

gaps between the existing and desired task performances. Self-reflection through keeping a diary, according to Svalberg (2012), can also figure in consciousness-raising tasks based on which learners evaluate the task as a means to learning and engagement with language. Involving learners in reflective activities has also been suggested by Ellis et al. (2020) as one type of post-task methodological options that is built on the main task. Ellis et al. (2020) classify reflection activities into two types: reflective accounts and transcription. They state that learners' reflective accounts include self-reports about their task performance, the knowledge that they gain from their performance, their assessment of task design features, and their views on how to upgrade the experience. Transcription as the other reflective choice proposed by Ellis et al. refers to learners' provision of a transcript of their own or a peer's task performance.

Although the second category, namely transcription, has been previously explored by Hsu (2019) and Kartchava and Nassaji (2019), the use of a questionnaire to obtain learners' opinions about their task performance has not been investigated so far. Hsu (2019) examined whether asking learners to transcribe their first oral narrative task performance could enhance the complexity, accuracy and fluency of their repeated and a new task of the same type performances. Results indicated that compared to the TR only group, the post-task transcribing condition resulted in more accurate production of

clauses in the repeated task that was also transferred to the new task. However, it did not benefit complexity or fluency. Kartchava and Nassaji's (2019) study revealed that feedback on technology-based oral presentation and reflection on performance brought about overall effective task performance. More recently, Dao, Nguyen and Chi (2020) examined whether self-reflection could promote 68 adolescent Vietnamese EFL learners' attention to form during peer interaction. The analysis of learners' produced language-related episodes revealed that this experience enabled them to self-correct their language errors and engage in metalinguistic talks. It was also found that the use of learned skills through practice depended on learners' proficiency and perceptions of their partner's performance.

It is likely that encouraging learners to reflect on their first task performance with respect to the strengths, weaknesses, and plans for progress, may boost the following iteration (Kartchava & Nassaji, 2019). In addition, reflection may have the advantage of directing learners' attention to the three stages of speech production and therefore cultivate their accurate linguistic production. The potential of reflective practice for L2 development and the scarcity of research in this area point to the need for further research. In light of this, the current study was conducted to investigate the following research question:

> What is the effect of reflective practice after the first performance of an oral narrative task on learners' linguistic accuracy in the repeated task and a new task of the same type?

3 Method

3.1 Design

A between-subjects design was adopted to explore the impacts of self-reflection between repeated task performances on L2 development. The independent variable is TR condition: (1) task repetition only (TR, control) and (2) task repetition with self-reflection (TR+SR). Learners in each group produced oral language elicited by means of two picture-cued oral narrative tasks. Their oral task performances were analyzed in terms of accurate use of the target structure. This study was based on Kolb's (2014) four levels of reflective practice: concrete experience (participants first performed the oral narrative task), reflective observation (participants reflected on and interpreted their performance based on recalling the experiences by filling out the reflection questionnaire), abstract conceptualization (participants connected pieces of the experiences and tried to learn from them) and active experimentation (participants were encouraged to put what they had learned into the performance of the subsequent tasks). The study design is exhibited in Figure 2.

	TR (N = 17)	TR + SR (N = 14)
Week 1, session 1 (Saturday)	Key English Test	
Week 2, session 2 (Saturday)	The winner	The winner
Week 2, session 3 (Monday)	The winner	Self-reflection + The winner
Week 2, session 4 (Wednesday)	The winner	The winner
Week 3, session 5 (Saturday)	Landslide	Landslide

Figure 2 Study design

3.2 Participants

A total of 31 learners from two general English classes in a higher education college in Iran voluntarily agreed to participate in the present study. This course is instructed over one semester and consists of around thirteen teaching weeks of 1.5 hour sessions twice a week. Admission to this college is based on the nation-wide entrance exam performance. Learners in each class were randomly assigned into one condition: TR (N = 17) and TR+SR (N = 14). Participants comprised 12 male and 19 female learners ranging in age from 16 to 21. Their first language was Turkish and/or Kurdish and they also spoke in Farsi as the official language in Iran. All of them had been studying English for at least 5 years starting from the junior high school. Their proficiency level was A2 level of the Common European Framework of Reference for Languages (CEFR) as determined by a standardized test (Cambridge Key English Test or KET). The results of an independent samples t-test indicated no significant differences between the groups (p = .342) regarding their level of English proficiency. Since English is a foreign language in Iran, participants did not have any opportunity to use the language outside the classroom and none had ever been to or lived in any English-speaking countries prior to the study. All participants signed informed consent forms prior to data collection.

3.3 Target structure

This study attempted to measure the effect of TR condition on learners' development of English regular past tense structure. Although the past tense structure is introduced to the learners in quite early stages of learning, it still poses challenges to them even at high levels (Ellis et al., 2006). This difficulty is partially related to the phonological challenge triggered by the consonant clusters with final [t] or [d] for Asian learners of English (Ellis et al., 2006).

3.4 Task

The tasks used in this study were adopted from Heaton (1975). The first task, 'The winner' was implemented as the main task and the repeated tasks, and the second task, 'Landslide', served as the new task of the same type but with different content. They were both narrative tasks that entailed six wordless pictures presented chronologically. Participants were asked to look at the pictures and tell the story to the researcher who only used backchannels such as I see, OK, hmm during the narration. They had to narrate the story in an understandable manner that even those who have not seen the pictures could understand it. They were permitted to look at the pictures during the story telling. The results of a pilot study indicated that learners completed the task in no more than 20 min. There were three major reasons behind opting for this task. First, the majority of previous TR studies have used this type of task. Second, this task directs more attention to meaning than form which is the crucial requisite of focus on form (Park, 2010). Additionally, this task requires imagination and analysis by the learner making it a cognitively demanding task and thereby limiting focus on form (Ellis & Yuan, 2004). All the narrative productions of participants were audio recorded via a digital voice recorder.

3.5 Reflection questionnaire

Learners were encouraged to reflect on their first task performance by filling out a questionnaire which was developed for the purpose of this study (see Appendix A). The questionnaire included 27 items, to which the participants responded on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The content of the questionnaire was arrived at by adopting Ellis et al.'s (2020) proposal that learners' reflective accounts as self-reports need to be about five themes:

- what they think they learned during the task,
- their evaluation of their task performance,
- their perceptions of the design features of the task including its objective, nature and difficulty,

– their attitudes towards the task, and
– their opinions about how to improve it (p. 235).

Additionally, nine English teachers all with PhDs in TESOL were interviewed to provide general feedback about the questionnaire regarding whether the items could appropriately induce learners to reflect on their task performance experience in general and the target structure in particular. Teachers' feedback led to a better evaluation of the face and content validity of the questionnaire items. Based on these guidelines, 32 statements in English were composed to measure the five constructs. The questionnaire was then translated into Farsi. A classroom-based pilot study was conducted to trial the questionnaire with 108 Iranian learners of English. To find out the construct validity of the questionnaire, an exploratory factor analysis was run using varimax rotation. A five-factor solution was found which aligned closely with Ellis et al.'s suggested dimensions (Learned Knowledge, Task Performance, Task Design, Attitude, and Improvement). Yet, because nine items loaded very weakly, they were eliminated and four new items related to the five factors were developed. Thus, the revised questionnaire included 27 items. Questions 1–4 concern the knowledge that learners gained as a result of doing the task. Questions 5–10 relate to their task performance. Questions 11–16 examine the task design features influencing their task experience. And, while questions 17–21 focus on learners' beliefs and attitudes towards the task, Questions 22–27 emphasize their opinions about the improvement of their task performance experience. The reliability of the questionnaire was also estimated using Cronbach's alpha (α = .93). Participants were given at most 15 minutes to complete the questionnaire-as identified in the pilot study.

3.6 Procedure

This study was carried out in a quiet room after learners' regular class time. The treatment lasted over a 3-week period. The researcher met with each learner five times and each task completion session took approximately 20 to 40 minutes. The procedure instructions were in Farsi. One week before the study, participants were screened using the KET. Session two began in the next week. Participants in both groups were provided with a five-minute preparation time and then performed 'The winner' task for the first time. In session three in the same week, participants repeated the same task. However, although the TR+SR experimental group learners were asked to fill out the self-reflection questionnaire before the second task performance, the control group learners repeated the task without any self-reflection opportunities. Two days later, participants repeated the same task once again. Lastly, they enacted the 'Landslide' task as the new task with the same procedure but with different content. It should be mentioned that no counterbalancing was used in this study.

3.7 Data analysis

The recorded oral narratives were transcribed to identify the obligatory contexts. An obligatory context consisted of a sentence in which learners had to use an exemplar of the target linguistic form to make it grammatically correct. Subsequently, the erroneous and non- or overuse of the target structure were determined as incorrect. The overuse referred to the oversupply of an exemplar of the target form in a context where the exemplar was ungrammatical. The percentage of correct use was attained using the following formula which demonstrates the sum of obligatory contexts and overuse.

$$\frac{\text{Number of forms supplied correctly}}{\text{Number of obligatory contexts + number of overused forms}} \times 100$$

To ascertain the coding consistency, a second rater who was an Associate Professor in a public university in Iran and the researcher scored the data. Cohen's Kappa was used to assess inter-rater reliability with the resulting value of .98. All the discrepancies were resolved through discussion.

Statistical analyses were conducted using the Statistical Package for Social Sciences (SPSS) version 22.0. The alpha level was set at .05 for between-group analyses. To examine which type of TR condition resulted in the accurate use of the target structure, a repeated measures ANOVA was conducted with time (Time 1, Time 2, Time 3, and Time 4) as a within-subject variable, self-reflection (TR+SR, TR) as a between-subject variable, and participant performance regarding accurate structure use as the dependent variable. Independent samples t-tests were used to compare the two groups' performance. Partial eta-squared ($\eta p2$) and d values were used to estimate effect sizes. Based on the discipline-specific benchmarks proposed by Plonsky and Oswald (2014), for $\eta p2$, an effect size of .40 was considered small, .70 was interpreted as medium, and 1.00 was taken as a large effect. According to Plonsky and Oswald's criteria for interpreting the magnitude of d, .60 was considered small, 1.00 medium, and 1.40 large. The Kolmogorov-Smirnov test indicated that the data were all normally distributed ($p > .05$). Furthermore, Levene's tests verified the homogeneity of variance among all sets of group means compared in the analyses ($p > .05$). Additionally, the assumption of Sphericity using Mauchly's test was also met, $\chi2(5) = 21.61$, $p = .27$. Finally, to evaluate each group's development over time, a number of paired sample t-tests were used for within-group comparisons. To avoid Type I error because of numerous comparisons, the significance level was adjusted to .0125, and values of $p \leq .01$ were accepted to be significant.

Table 2 L2 Development across groups over time

Time	Groups	
	TR (N = 17) Mean (SD)	TR+SR (N = 14) Mean (SD)
Time 1	6.23 (2.27)	5.78 (1.67)
Time 2	6.94 (1.78)	8.50 (1.28)
Time 3	6.52 (1.80)	8.14 (1.23)
Time 4	6.29 (1.68)	7.92 (1.32)

Note: Time 1 = main task, Time 2 = first repetition, Time 3 = second repetition, Time 4 = new task

4 Results

Table 2 presents the group means and standard deviations for each TR condition's accurate use of the target structure over time, and Figure 3 exhibits the group means graphically.

The ANOVA of the accuracy scores yielded a significant group effect, $F_{(1, 29)} = 4.25$, $p = .04$, $\eta p2 = .12$, time effect, $F_{(1, 29)} = 9.18$, $p = .005$, $\eta p2 = .24$, and a significant group × time interaction effect, $F_{(1, 29)} = 10.72$, $p = .003$, $\eta p2 = .27$, suggesting that the groups' linguistic accuracy changed over time. Further between-group comparisons showed that both groups had comparable performance at Time 1, $F_{(1, 29)} = 1.99$, $p = .54$. Nevertheless, the TR+SR outperformed the TR at Time 2, $F_{(1, 29)} = 1.09$, $p = .011$, $d = 1.00$, Time 3, $F_{(1, 29)} = 1.92$, $p = .008$, $d = 1.05$, and Time 4, $F_{(1, 29)} = 3.55$, $p = .006$, $d = 1.07$. The results of paired samples t-tests for within-group comparisons depicted that the TR+SR significantly improved their accurate L2 use from Time 1 to Time 2 ($p = .0005$, $d = 1.82$), Time 3 ($p = .0005$, $d = 1.60$), and Time 4 ($p = .001$, $d = 1.42$), all with large effect sizes. The TR+SR was also successful in carrying over their gains from Time 2 to Time 3 ($p = .055$, $d = .28$) and Time 4 ($p = .13$, $d = .44$). Participants' high performance at Time 3 was also maintained at Time 4 ($p = .55$, $d = .17$). These within-group results, then, suggest that scores in the TR+SR remained high during subsequent performances. The TR group's performance, on the other hand, did not change over time ($p > .05$). In sum, as Figure 3 clearly depicts, whereas TR learners did not improve their accurate use of the past tense throughout the experiment, the TR+SR were successful in developing their grammatical knowledge after exposure to self-reflection and maintaining the enhanced knowledge in subsequent performances.
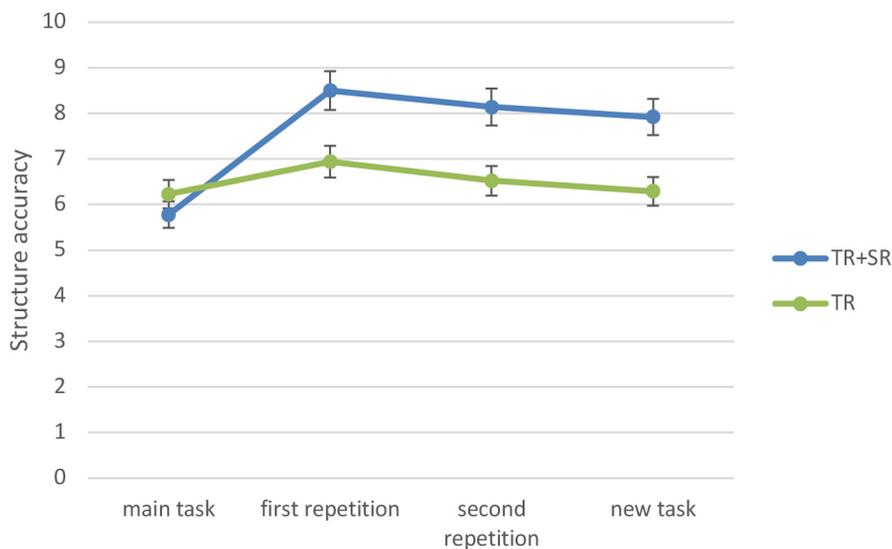
Figure 3 Groups' L2 development over time

Table 3 Questionnaire results

	Strongly agree frequency (percentage)	Agree frequency (percentage)	Neutral frequency (percentage)	Disagree frequency (percentage)	Strongly disagree frequency (percentage)
Learned knowledge	13 (23.2)	20 (35.7)	13 (23.2)	5 (8.9)	5 (8.9)
Task performance	26 (31.0)	24 (28.6)	13 (15.5)	11 (13.1)	10 (11.9)
Task design	31 (36.9)	24 (28.6)	17 (20.2)	7 (8.3)	5 (6.07)
Attitude	34 (48.6)	18 (25.7)	7 (10.0)	5 (7.1)	6 (8.6)
Improvement	17 (20.2)	29 (34.5)	17 (20.2)	13 (15.5)	8 (9.5)

Additionally, participants' responses to the reflection questionnaire in terms of frequency and percentage are reported in Table 3.

As Table 3 reports, over half of TR+SR learners agreed that they learned knowledge from their task performance and could effectively perform the task. Likewise, the majority were content with the design of the task and held positive attitudes toward the task. And, they mostly welcomed the suggested options for the improvement of task performance.

5 Discussion

This study investigated the effects of learner self-reflection carried out between performances of the same oral narrative task on accurate L2 use. The results of the statistical analysis revealed a significant increase in the TR+SR group's performance scores from the main task to both the first and second repeated task performances. What makes the findings of this study particularly valuable is that the participants' enhanced accuracy was transferred to the new task at Time 4. This finding lends credence to the facilitative role of encouraging learners to reflect on their overall performance experience prior to their second task enactment in focusing on linguistic features (e.g., Dao et al., 2020; Hawkes, 2012; Kartchava & Nassaji, 2019; Lynch, 2018). These studies have all indicated that through any type of reflective practice, learners' attention to language structures could be fostered. This is because self-reflection requires mental effort that occurs and is focused on a particular context and pursues a specific structure (Clegg, 2003). Put differently, for reflective practice to be effective, it needs to be carried out with a particular objective and be purposely structured so that certain learning outcomes can be achieved. Focused and structured analysis of past performances facilitate learners' deeper engagement with the learning process and therefore enables them to enhance future performances (Kolb, 1984, 2014). This was evidenced in the present study regarding participants' high performances from Time 2 onwards.

These findings also underscore the findings of previous research (e.g., Ahmadian & Tavakoli, 2011; Gass et al., 1999; Patanasorn, 2010; Sheppard & Ellis, 2018) showing that repeating an exact task has a limited impact on accurate language use. According to Levelt's speech production model, TR improves both conceptualization and formulation enabling the low proficiency learners such as those in the present study to retrieve the language to deliver their meaning more promptly yet with lesser accuracy. The TR learners' non-significant development in their oral narrative scores underlines the need for some intervention to help learners achieve accuracy not just in the repeated task but also in a new task.

In sum, encouraging learners to reflect on their performance is argued to cultivate their following ratings (Khezrlou, 2019a; Winke, 2014) and bears significant implications from the TBLT viewpoint. This is particularly crucial in the case of task repetition because for TR to result in acquisition, awareness about the first performance is needed to reinforce the subsequent ones (Ellis, 2019). In this study, having reflected on their first performance and made plans for improvement, the participants were arguably better able to monitor and improve their performances. More particularly, learners' reflection on their initial performance stimulated them to set realistic plans for the enhancement of their task performance by noticing and then narrowing the gaps between their existing and target abilities. From a different perspective, apart from bridging the gap in learners' cognitive understanding between what is and what should be, self-reflection can also be an affective motivator for learners to stay on and engage with the task particularly in

repeated performances regardless of possible challenges (Dörnyei & Ryan, 2015). In this way, learners who reflect on their first performance can be perceived as proactive about their learning since having specified the areas to amend, they can then use strategies to foster their L2 in future performances (Dörnyei & Ryan, 2015). To recap, self-reflection empowers learners to "take charge of their own learning, determine their objectives, select methods and techniques and evaluate what has been acquired" (Littlewood, 1999, p. 75).

6 Conclusion

This study has displayed that encouraging learners to self-reflect on their initial task performance facilitated their accurate L2 use in repeated as well as new task contexts. Accordingly, teachers are suggested to benefit from reflective practice in promoting learners' focus on form. They can attain this by following the cyclical reflective learning model to guide learners towards reflection on their task performance. That is, reflective practice, based on the reflective learning model, could be potentially beneficial for training learners to reflect on their task performance and language production for the sake of L2 learning. Therefore, teachers are encouraged to get learners thinking as they engage in reflection and provide them with repeated task performance opportunities through which they can employ their plans for improvement. Moreover, through self-reflection, teachers can understand their learners' capacity for change, needs, interests, and capabilities, based on which, they can alter the instructional materials and approaches they use (Bachman & Palmer, 1996).

   Despite these benefits, the study has some limitations that should be borne in mind. First, this study explored the lower-proficiency learners' accurate use of the regular past tense structure following the treatment. Hence, there is a need for more research to examine fluency and complexity with beginning level learners in comparable conditions. Additionally, more research is needed with different level learners, target features, and learning contexts. Furthermore, the lack of counterbalancing in this study might have led to order and topic effects which should be taken into consideration in future studies. Lastly, some items in the questionnaire (e.g., Item 2 and Item 15) were double-barreled and in spite of the measures taken to ensure the validity of the questionnaire, some items might be improved in future studies to target the five constructs more directly.

References

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

Bachman, L.F., & Palmer, A.S. (1996). *Language testing in practice*. Oxford University Press.

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 Publication.

Carver, J., & Kim, Y. (2020). French learners' past-tense development through collaborative writing tasks: The role of procedural and content repetition. *The Canadian Modern Language Review, 76*(2), 114–138. https://doi.org/10.3138/cmlr-2018-0231

Clegg, K. (2003). *Improving learning through reflection, Part 1 and Part II*. Higher Education Academy.

Dao, P., Nguyen, M.X.N.C., & Chi, D-N. (2020). Reflective learning practice for promoting adolescent EFL learners' attention to form. *Innovation in Language Learning and Teaching*. https://doi.org/10.1080/17501229.2020.1766467

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. (2019). Task preparedness. In Z.D. Wen, & M.J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honor of Peter Skehan* (pp. 15–38). John Benjamins.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*(2), 339–368. https://doi.org/10.1017/S0272263106060141

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*(1), 59–84. https://doi.org/10.1017/S0272263104261034

Farrell, T.S.C. (2011). Exploring the professional role identities of experienced ESL teachers through reflective practice. *System, 39*(1), 54–62. https://doi.org/10.1016/j.system.2011.01.012

Gass, S., Mackey, A., Alvarez-Torres, M.J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*(4), 549–581. https://doi.org/10.1111/0023-8333.00102

Hawkes, M.L. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal, 66*(3), 327–336. https://doi.org/10.1093/elt/ccr059

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

Hsu, H-C. (2019). The combined effect of task repetition and post-task transcribing on L2 speaking complexity, accuracy, and fluency. *The Language Learning Journal, 47*(2), 172–187. https://doi.org/10.1080/09571736.2016.1255773

Jarmon, L., Traphagan, T., Mayrath, M., & Trivedi, A. (2009). Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in second life. *Computers & Education, 53*(1), 169–182. https://doi.org/10.1016/j.compedu.2009.01.010

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

Kellenberg, F., Schmidt, J., & Werner, C. (2017). The adult learner: Self-determined, self-regulated, and reflective. *Journal of Pedagogy and Psychology, 9*(1), 23–29. https://doi.org/10.1515/sigtem -2017-0001

Khezrlou, S. (2019a). Form-focused instruction in CALL: What do learners think? *RELC*, *50*(2), 235–251. https://doi.org/10.1177/0033688217738820

Khezrlou, S. (2019b). 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 -019-00025-2

Khezrlou, S. (2019c). Effects of timing and availability of isolated FFI on learners' written accuracy and fluency through task repetition. *The Language Learning Journal.* https://doi.org/10.1080/ 09571736.2019.1656765

Khezrlou, S. (2020a). Training planning in second language narrative writing. *ELT Journal*, *74*(1), 49–62. https://doi.org/10.1093/elt/ccz050

Khezrlou, S. (2020b). The role of task repetition with direct written corrective feedback in L2 writing complexity, accuracy and fluency. *Journal of Second Language Studies*, *3*(1), 31–54. https://doi.org/10.1075/jsls.19025.khe

Khezrlou, S. (2021). Explicit instruction through task repetition: Effects on explicit and implicit knowledge development. *Language Awareness, 30*(1), 62–83. https://doi.org/10.1080/09658416 .2020.1866590

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

Kolb, D.A. (1984). *Experiential learning: Experience as the source of learning and development.* Prentice Hall.

Kolb, D.A. (2014). *Experiential learning experience as the source of learning and development.* New Jersey FT Press.

Lam, R. (2018). Promoting self-reflection in writing: A showcase portfolio approach. In A. Burns, & J. Siegel (Eds.), *International perspectives on teaching skills in ELT* (pp. 219–231). Palgrave MacMillan.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* MIT Press.

Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, *20*(1), 71–94. https://doi.org/10.1093/applin/20.1.71

Lynch, T. (2018). Enhancing task repetition in second language speaking classes. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 193–222). John Benjamins.

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research, 4*(3), 221–250. https://doi.org/10 .1177/13621688000400303

Moon, J. (2004). *A handbook of reflective and experiential learning.* Routledge.

Park, S. (2010). The influence of pretask instructions and pretask planning on focus on form during

Korean EFL task-based interaction. *Language Teaching Research, 14*(1), 9–26. https://doi.org/10.1177/1362168809346491

Patanasorn, C. (2010). *Effect of procedural content and task repetition on accuracy and fluency in an EFL context.* Unpublished doctoral thesis, Northern Arizona University.

Plonsky, L., & Oswald, F.L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning.* Palgrave Macmillan.

Sheppard, C. (2006). *The effects of instruction directed at the gaps second language learners noticed in their oral production.* Unpublished doctoral thesis, University of Auckland.

Sheppard, C., & Ellis, R. (2018). The effects of awareness-raising through stimulated recall on the repeated performance of the same task and on a new task of the same type. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 177–199). John Benjamins.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford University Press.

Svalberg, A. (2012). Language awareness in language learning and teaching: A research agenda. *Language Teaching, 45*(3), 376–388. https://doi.org/10.1017/S0261444812000079

Winke, P. (2014). Formative, task-based oral assessments in an advanced Chinese-language class. In M. Gonzalez-Lloret, & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 263–293). John Benjamins.

Appendix A: Reflection questionnaire

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Learned Knowledge					
(1) I learned how to tell the story based on the pictures on a one-by-one basis.					
(2) I learned new language (vocabulary and grammar) as a result of doing this task.					
(3) I learned to use appropriate words to tell the story.					
(4) I learned to narrate the story in the past.					
Task Performance					
(5) I could effectively organize my ideas developed from the pictures to have a clear story plot.					
(6) I could communicate the small details.					
(7) I used the verbs in the past tense correctly by adding -ed to the end of regular verbs.					
(8) I thought of more ideas beyond the pictures.					
(9) I could cover the content fully.					
(10) I am satisfied with my task performance.					
Task Design					
(11) I think this task was easy.					
(12) I felt relaxed performing this task.					
(13) The task was clearly formulated.					
(14) This task gave me an opportunity to use the past tense structure.					
(15) I liked the pictures which I think were structured enough.					
(16) I sometimes forgot that this task was meant to be for language learning.					

(*cont.*)

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree

Attitude

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
(17) This task was very frustrating for me.					
(18) The task topic was interesting for me.					
(19) The task was related to my language course.					
(20) I can imagine situations from this task happening in real life.					
(21) I would like to do more tasks like this.					

Improvement

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
(22) I want to be given option to choose the task myself.					
(23) I think tasks can be developed by students themselves.					
(24) I think the teacher should have given us information about the past tense structure before the task to refresh our knowledge.					
(25) Tasks should be closely related to my personal interests.					
(26) The teacher should intervene in case of problems.					
(27) The teacher should use collaborative tasks.					