POLINGUA
Scientific Journal of Linguistics, Literature and Language Education
available online at: http://www.polingua.org/index.php/polingua

The Impact of Working Memory on Language Gain of Corrective Feedback: A Meta-analysis

Liping Chen[1], Craig Lambert[2]

[1]*Southwest Petroleum University, China*
E-mail: [1]*chenliping@swpu.edu.cn*

[1, 2]*Curtin University, Australia*
Email: [2]*craig.lambert@curtin.edu.au*

***Abstract*— The topic of working memory has attracted researchers in second language acquisition. Working memory is one important individual difference variable affecting the language learning process. Previous research has focused on the relationship between working memory and corrective feedback. However, the effectiveness of measuring working memory and corrective feedback remains unclear. The current study provides a meta-analysis of studies of working memory on the effects of language gains after corrective feedback. This meta-analysis retrieved samples from 12 primary studies involving 489 participants to calculate the average effect sizes of the correlations between working memory and language gains of oral corrective feedback. The results found that working memory has a positive small effect size on language gains of oral corrective feedback (r =.321). The findings also revealed that the methodological factors had an impact on the correlation between working memory and language gains of oral corrective feedback. These findings help researchers and learners better understand working memory's working mechanics and corrective feedback. These findings also give suggestions from the methodological perspective for future research on the study of WM on language outcomes of CF in second language acquisition.**

***Keywords*— Meta-analysis; methodological factors; oral corrective feedback; second language acquisition; working memory.**

Manuscript received 04 August. 2023; Revised 20 September. 2023; Accepted 24 September. 2023

I. INTRODUCTION

Corrective feedback (CF) has been studied widely in the context of second language acquisition (SLA)since the 1960s. CF is defined as "the responses to a learner's non-target-like second language (L2) production" (Li, 2010, p. 309). CF provides learners with positive and negative evidence in language learning through interaction processes. CF provided by teachers helps learners notice the gap between their own incorrect language production and correct language forms from interlocutors (Brown, 2016b; Gass & Mackey, 2007; Li, 2010; Long, 2007; Lyster & Saito, 2010b). Early CF studies emphasised the effects of six different types of CF in language learning: namely, explicit correction, recasts, clarification requests, metalinguistic feedback, elicitation, and repetition. CF can also be differentiated in terms of external and internal factors. However, most corrective feedback studies focused on external factors only (e.g., Kim, Payant and Pearson, 2015; Li, 2013; Lyster, 2004), that is, they considered types of corrective feedback, age of learners, learners' level of language proficiency, and interlocutor differences (such as being a native or non-native speaker), but ignored learners' internal factors (individual differences such as foreign language aptitude and working memory).

Individual differences (IDs) are the cognitive and psychological variables that affect the learning process (Sheen, 2007). IDs result in some learners experiencing successful uptake of feedback from teachers and others experiencing unsuccessful uptake of such feedback (Li, 2017).

One important individual difference that has received considerable attention in the L2 research literature is working memory. There are two types of memory, namely, long-term memory and short-term memory. Working memory (WM), which is a short-term memory, is different from long-term memory. It is defined by Baddeley (2003) as "the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities" (p. 189). Working memory is an important ID variable in language learning relating to learners' variable abilities to control and regulate the cognitive process of learning, including allocating attentional resources, storing information, and retrieving information during second language acquisition (Baddeley, 2012, 2017). WM is also important in learners' ability to direct attention to noticing gaps between their own language production and targeted language forms while interacting in the target language (Gass & Mackey, 2007; Li, 2010).

Two previous meta-analyses have investigated the relationship between working memory and language interaction (Li, 2017; Linck et al., 2014). Linck et al. (2014) reviewed 79 studies on the relationship between working memory and second language comprehension and production. The results of the meta-analysis indicated that working memory had a positive relationship with L2 comprehension and production outcomes, however Linck's study only focused on the overall relationship between WM and L2 outcomes (r=.255), ignored the specific effect sizes of WM on CF outcomes in L2 interaction. Li (2017) combined a meta-analysis with a narrative review to synthesise the results of 24 studies on the effects of language aptitudes (working memory and grammatical analytical ability) on language interaction and production. In Li's study, he also synthesised eight studies of working memory on the effects of corrective feedback and production. Li found that working memory had a weak positive association with the effectives of corrective feedback in L2 interaction (r = .23).

Although several meta-analysis studies have reported positive effects for WM on the interaction of SLA with small effect sizes, to date little is known about overall effect sizes of working memory actually on corrective feedback in L2 interaction. In addition, there is lack of studies to synthesise the impact of methodological variables on average effect size of the correlation between WM and oral corrective feedback in L2 learning. This meta-analysis in this paper attempts to address this issue by investigating the overall effect sizes of working memory on CF. The meta-analysis in this paper also attempts to identify research design variables in the process of CF and the measures of WM in previous research from the perspectives of methodology, which might impact the heterogeneity of research results.

II. LITERATURE REVIEW

A. Corrective Feedback in SLA and Measurement

Corrective feedback (CF) provided by teachers is defined as the responses that learners receive on their errors in the language learning process (Li, 2018). CF also is one important variable of language interaction to contribute to language gains in L2 (Li, 2017). Lyster and Ranta (1997) classified CF into six types, namely, recasts, explicit correction, clarification requests, repetition, elicitation and metalinguistic feedback. The six types of CF are further divided into two categories: input-providing feedback and output-prompting feedback. Recasts and explicit correction provide the correct form as input-providing feedback as repair is encouraged, while the other four types of CF are output-prompting feedback without positive evidence of target forms as repair is not needed (Li, 2018). The six types of CF are also divided into two categories: implicit and explicit feedback (Li, 2010). Metalinguistic feedback and explicit correction provide explicit learning of target forms. While implicit feedback includes recasts and negotiation (clarification requests, elicitation, and repetition) (Li, 2010, 2019). Recasts as a typical type of implicit feedback are defined as "a reformulation of a non-target-like L2 utterance" (Li, 2013, p. 636), while metalinguistic correction is the typical explicit feedback, which is defined as "comments, information, or questions related to the well-formedness of the student's utterance" (Lyster & Ranta, 1997, p. 47).

The effectiveness of CF on language gains is measured by gains in language process of CF (noticing and uptake) and gains in language production (score gains between pre-test and post-test) in SLA (Li, 2017). To be specific, language process of CF consists of noticing and uptake. Learners' noticing can be considered as one aspect of enhancing L2 development in language interaction. Noticing is defined as the "registration (detection) of the occurrence of a stimulus event in consciousness awareness and subsequent storage in long-term memory" (Schmidt, 1994, p. 179). The gains of noticing are to detect the linguistic stimulus and recognise or understand the rules (Schmidt, 1994). Instruments for measuring noticing include online or offline verbal reports (i.e., learners report what they noticed in the interaction process to researchers in written or oral forms) (Li, 2017). Online verbal methods are used to report what learners notice immediately following the interaction process including think-aloud protocols and immediate recall. In contrast, offline methods require learners to report what they noticed in the interaction process in oral or written form after the interaction including stimulated recall, uptake charts, language-related episodes (LREs), and interviews (Ellis, 1995; Li, 2017; Mackey, 2006). Therefore, noticing is based on learners' reports and difficult in measuring noticing gains accurately.

Uptake is also widely used to measure language gains of target structures in the CF process, which is related to L2 learning by real-time production of the interaction treatment (Egi, 2010; Fu & Li, 2019; Goo, 2012; Kartchava & Ammar,2014; Li, 2013; Li et al., 2019; Mackey, 2006; Mackey & Sachs, 2012; Revesz, 2012; Zhao, 2015). Uptake is defined as "a learner utterance immediately following teacher feedback and constituting a reaction to that feedback" (Mackey, 2006, p. 406).

Score gains from pre-test to post-test after the CF treatment are widely used to measure language gains after CF interaction of target grammar structures. Grammaticality judgment tests (GJTs), elicited imitation tests (EITs), free oral productions tests and controlled production tests are common tests in L2 research to measure learners' knowledge of grammatical structure (Ellis & Roever, 2021).

*B. Working Memory and Measuring Working Memor*y

Working memory has been shown to be a predictor of language comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Goo, 2012; Linck, Osthus, Koeth, & Bunting, 2014). WM is defined as "the kind of memory system(s) that allows us to maintain and manipulate a very small amount of information in our head when we are carrying out some cognitive tasks in daily life, such as language comprehension, arithmetic calculation, reasoning and problem solving" (Wen, 2015, *p*.41). Baddeley (2015) established the multicomponent model of WM, consisting of a central executive, a phonological loop, a visuospatial sketchpad and an episodic buffer. The central executive is domain-general to control and regulate attention, while the phonological loop and the visuospatial sketchpad are domain-specific. The phonological loop is used to store verbal information temporarily, while the visuospatial sketchpad is used to maintain visual or spatial information. The episodic buffer contains a temporary storage system, accesses the long-term memory and integrates information from the phonological loop and the visuospatial sketchpad (e.g., verbal and visual information), and links the integrations with the central executive (Baddeley, 2015). However, because of scarce research no agreed approaches to measurement in current research (Baddeley, 2015; Li, 2017), the episodic buffer was excluded from measures of working memory in the meta-analysis presented in this paper.

Wen (2015) distinguished working memory from phonological short-term memory (PSTM) and executive working memory (EWM). PSTM refers to "the storage capacity only", while EWM contains "the dual task of information manipulation and storage" (Li, 2017, p. 44). PSTM is defined as "the ability to repeat phonological input correctly" (Mackey, Philp, Egi, Fujii, & Tatsumi, 2002, p. 185), which plays an important role in "the efficiency of acquisition of novel phonological forms and the serial-order retention of sequences/chunks of word forms" (Wen, 2015, p. 55). PSTM is also considered as an essential ability to learn from input during the initial stages of SLA (Robinson, 2005). EWM focuses on the efficiency of acquiring novel words forms/vocabulary and retention of sequences of forms/chunks at an early stage of language learning. According to Wen (2015), EWM is associated with the executive and attentional functions and is "mainly involved in monitoring and self-repair aspects of L2 learning" (p. 56). EWM emphasises the language learning process, including efficiency of encoding and retrieval, noticing, monitoring and self-repair by advanced L2 learners at a later stage (Wen, 2015).

According to the different constructs of WM, ways to measure WM are divided into simple tasks and complex tasks (Wen, 2015). A span test requires learners to recall a string of information such as words, digits, or images. Simple span tasks are used to test phonological short-term memory (PSTM), which refers to the storage component requiring learners to memorise unrelated words or digits. Simple span tasks require participants to "recall a string of nonrelated letters, words, digits, or visual objects after the presentation" (Linck et al., 2014, p. 863), such as used in a digit span test, a non-word span test and a forward digit span test, and a non-word repetition span test (Baddeley, 2003; Gathercole & Baddeley, 1993; Li, 2017). Complex span tasks are widely used to gauge both the storage and processing components required by learners in memorising an element of semantic, syntactic, or mathematical processing (Li, 2015). Complex span tasks require participants "to actively process input while remembering a string of letters, words, digits, or objects" (Linck et al., 2014, p. 863), such as using a listening span test, a reading span test, an operation span test and backward digit span test (Daneman & Merikle, 1980, 1996; Engle, Cantor, & Carullo, 1992; Li, 2017; Mackey et al., 2002; Waters & Caplan, 1996).

C. The Impact of Working Memory on Corrective Feedback

The noticing process of CF is a process of executive attention based on the attention-control mechanism of WM (Conway et al., 2001; Engle, 2002; Engle, Kane, et al., 1999; Kane et al., 2001; Kane et al., 2007; Kane & Engle, 2003). Corrective feedback provides positive and negative evidence regarding target language use to learners and might draw their attention to specific linguistic forms during interaction. During interaction of CF, learners might use their working memory resources to notice target structures, compare the received target structures with their previous resources with positive or negative evidence, then integrate their received information in their subsequent performances (Li, 2017). Several studies found that WM helps learners notice and uptake salient or non-salient linguistic forms in CF's meaning-primary interaction (e.g., Goo,2012; Mackey et al, 2002; Zhao, 2015).

D. Potential Effects of Design Variables to Affect the Correlation between WM and Language Gains of CF

CF types are one important design variable to affect the correlation between WM and language gains of CF. In the previous studies, different types of CF were used in the previous research including implicit CF in the study (e.g., Mackey and Sachs, 2012; Revesz, 2012; Trofimovich, et al., 2007), or explicit CF in the study (e.g., Li, 2010, 2013, 2014; Li, Ellis & Shu, 2016; Li, et al, 2019; Goo, 2012;).

Measurement of CF's language gains is another important design variable to affect the correlation between WM and CF, consisting of noticing, uptake and score gains between pre-test and post-test in SLA (Li, 2017). Criteria of language gains' measurement of CF were inconsistent used in previous studies. For example, uptake of CF as an important measure of language gains was different and inconsistent in previous research (Fu & Li, 2019; Goo, 2012; Kartchava & Ammar,2014; Li, 2013; Li et al., 2019; Mackey, 2006; Mackey & Sachs, 2012; Revesz, 2012; Zhao, 2015). Mackey (2006) employed a qualitative approach to measure uptake by using learning journals, oral stimulated recall protocols, written answers to focused questions, and questionnaires. In contrast, Egi (2010) used a quantitative approach in which instances of uptake were coded as no opportunity, uptake/no uptake, repair/needs repair and modified output/unmodified output and calculate the total score of uptake after teaching interaction. Similarly, Kartchava and Ammar (2014) coded four types of evidence of noticing in immediate recalled episodes, including uptake (detection of corrective feedback with awareness), exact repetition (repeat teachers' reformulation without awareness), noticing of help (explicit mention that the teacher was trying to help), and no noticing.

Similarly, measures of WM are also an important design variable to affect the correlation between WM and CF. PSTM, EWM or combined WM are measured in previous research. For example, some studies have compared the difference of overall language gains of CF between two types of WM (EWM and PSTM) (Lado, 2008; Mackey & Sachs, 2012; Revesz, 2012; Trofimovich et al., 2007; Zhao, 2015). In contrast, other studies compared the differences of language gains of one type of CF between two types of WM (EMW and PSTM) (Goo, 2012; Li, 2013; Li et al., 2019; Yang, et al., 2019). Thus, because the different types of CF and WM were measured in previous research, the findings of the correlations between WM and language gains of CF were inconsistent.

In addition, other research design variables may also result in the heterogeneous results of previous studies. Those variables include timing of post-test (immediate post-tests or delayed post-tests), research setting (classroom or laboratory) and academic status of participants (secondary school, university, or language program). Therefore, the heterogeneity of methodologies to measure language gains of the effectiveness of CF and WM affects the validity and reliability of research results.

III. RESEARCH QUESTIONS

The research purpose of the meta-analysis presented in this paper is that reviewed 12 empirical studies attempts to obtain the magnitude of the overall effect sizes between working memory and language gains of corrective feedback in L2. In addition, moderator analyses were carried out to examine the interaction of the relationship and to identify variables that helps explain variability in the relationship between working memory and outcomes of corrective feedback (Hedges & Pigott, 2004). Through the moderate analyses, the study intended to examine the measurement as a means to provide methodological references and suggestions. The independent variable was working memory, comprising EWM and PSTM, whereby dependent variable was language gains of corrective feedback.

The following research question (RQ) are addressed with this meta-analysis:
RQ1: How does working memory relate to language gains of corrective feedback?
RQ2: What are more important variables to moderate the effect of working memory on language gains of corrective feedback?

IV. RESEARCH METHOD

A. *Data Collection: Identification of Analysable Research*

In conducting the meta-analysis, the first step was to search for related studies with key words in the Web of Science for English language studies published between January 2000 to March 2022. The following key words and combinations of key words were used to search: working memory, EWM, PSTM, listening span test, non-word span test, operational span test, corrective feedback, second language acquisition/learning, feedback, negative evidence, negotiation, foreign language learning/education, grammar, grammatical structure. The search for working memory and corrective feedback in the fields of linguistics, education, educational research, language and linguistics in databases (the Web of Science, ProQuest, PsycINFO, and ERIC) and the Google Scholar search engine yielded a total of 58 articles as March 29, 2022. The second step was to search for articles and book chapters listed in literature reviews and meta-analyses on corrective feedback. Articles of renowned researchers in the field were also checked and carefully examined. The third step was to search ProQuest Dissertation for unpublished Master theses and PhD dissertations to minimise publication bias or the "file drawer problem". The fourth step was browsing and checking references of existing meta-analyses on similar or related topics (Li, 2017; Linck et al., 2014) and searching physical books of the bookshelves in the library for related book chapters on this theme.

Our literature search was designed to include all available sources that assessed the relationship between WM and outcomes of CF. Studies in the present meta-analysis were limited to experimental studies that examined the relationship between working memory and outcomes of oral corrective feedback with enough statistics (*r, β*, Mean, SD) to convert one effect size index. In addition, the experimental studies included at least one experimental group with CF and one control group. If the study did not have one control group, the study provided score gains of pre-test and post-tests without the control group. Descriptive studies were excluded without providing specific data. Studies that measured more than at least one type of WM (EWM, PSTM or other) were included. The experimental studies provided at least one type of CF (implicit, explicit or both) which is mediated by face-to-face or computer with learners' language-related responses to their errors of target structures. Finally, the study quantitatively measured score gains of corrective feedback by pre-tests and post-tests.

Followed the above criteria, 12 studies on the correlation of WM and CF were reviewed in the present meta-analysis and identified with an asterisk (*) in the References list (see Table 1 below). The final data set includes 12 published studies, one book chapter, and two dissertations. To be specific, the final data set provides effect sizes from 42 independent samples involving 489 participants. Table 1 contains the following information for 12 studies: dependent variable, sub-groups of dependent variables, measure of CF score gains (GJT, Oral production, GJT & EIT, GJT & Written production, and GJT, Oral production & Written production), means and standard deviations (SD) of experimental groups, correlation coefficient (*r*), regression coefficient (*ß*), and sample size.

Table I provides a summary of analysed characteristics, which were coded for six variables. Out of the 12 studies, 10 were published articles in peer reviewed journals, one was Ph.D. dissertations and one was a book chapter. In terms of CF outcomes, three types of score gains were used in the included studies. GJT were used in 12 included studies separately or combined. Out of 12, four studies only GJT to measure language gains of CF, while two studies used oral productions. Combined language gains were used in six included studies, consisting three types- GJT & EIT, GJT & Written production, GJT & Written production & Oral production). Regarding to measure of WM, combinations of EWM and PSTM were used in six of the included studies. EWM was adopted in five of the studies, while PSTM was used in one study. Studies on feedback types included seven studies on implicit corrective feedback, two studies on explicit CF, and three on implicit and explicit corrective feedback. As regards the timing of post-tests, five studies focused on immediate post-tests, while eight studies used delayed post-tests.

TABLE I
CHARACTERISTICS OF THE RETRIEVED STUDIES

Data for calculation of average effect sizes		kc
	Mean and SD	1
	regression coefficient (ß)	6
	correlation coefficient (r)	5

Published types		
	Published studies	10
	Book chapter	1
	Dissertations	1
Measure of Variables		
Measures of CF outcomes (Independent variables)		
	GJTa	4
	Oral production Combined (GJT & EITb, GJT & Written production, and GJT & Oral production & Written production)	2 6
Measure of Working memory (Dependent variables)		
	EWM (Listening span, reading span, operation span, running span, letter-digit series, backward digit	6
	PSTM (Non-word recall, forward digit span))	1
	Combined (EWM & PSTM)	5
Methodological Features of CF		
CF types		
	Implicit	7
	Explicit	2
	Implicit & Explicit	3
Timing of post-tests		
	Immediate	4
	Delayed	1
	Immediate + Delayed	7
Research setting	Classroom Laboratory	9 3
Academic status	Secondary school University Language program	2 7 3

Notes:

a. GJT=Grammaticality judgment test
b. EIT= Elicited imitation test
c. Number of studies in the analysis (*k*)

B. Variables and Coding Procedure

The 12 selected studies were identified several variables that might impact the score of average effect sized and coded in terms of different tests for score gains of corrective feedback as the independent variable (GJT, Oral production, GJT & EIT, GJT & Written production, or GJT & Oral production & Written production) and WM as dependent variables (PSTM, EWM, or PSTM & EWM) (see above Table I). Some clarification on corrective feedback studies is also listed in above Table I. T

Effect sizes were calculated for the independent variables of types of WM (EWM, PSTM and Combined) he different types of corrective feedback were coded as implicit and explicit with those studies to investigate whether working memory functions differently in implicit or explicit conditions. Implicit feedback does not provide rule explanation, including recasts in this meta-analysis, while explicit feedback directly provides rule explains or correct answers, including metalinguistic feedback and explicit correction in this meta-analysis (Lyster, 1998; Li, 2010). In addition, the difference of timing of post-test on effect of corrective feedback were measured by immediate post-tests and delayed post-tests (Li, 2017).

C. Analysis

The analyses were conducted using Comprehensive Meta-Analysis (CMA) (Borenstein, Hedges, Higgins, & Rothstein, 2005) 3.0 software. The process of consisted of the following steps:

1. Where necessary, *Q*-tests were performed to detect between-group differences. Random-Effects models are used in this meta-analysis. A moderator analysis was measured by a Q_b test to investigate the different predicative validities of the effect of WM on CF;
2. If a study had a control group and experiment group, effect sizes were calculated by comparing experiment groups with the control group;
3. Effect size analysis and moderator analysis were not performed for variables with less than three effect sizes.

The following steps were used to calculate results:

1. If a study had correlation *r* value, effect sizes were calculated based on r value.
2. If a study did not provide r value, data was transformed to correlation *r* value. In terms of Beta coefficients, β was transformed to correlation *r* value by $r=\beta\times0.98+0.05(\beta\geq0)$ or $\beta\times0.98-0.05(\beta<0)$ (Peterson & Brown, 2005). Subsequently, raw data were automatically transformed to correlation *r* value by CMA 3.0 software;
3. If a study did provide mean and SD of the experiment groups and the control group, effect sizes were calculated by *t* value comparing experiment groups with the control group, data were automatically transformed to correlation *r* value by CMA 3.0 software;
4. Fisher Z was used to auto transform *r* as the unit of analysis for interpretation by Complex meta-analysis 3.0 software (Hunter & Schmidt, 2004). Cohen's *d* value and 95% Confidence intervals (CI) were calculated for reliability of effect sizes. CI were calculated to determine whether the mean effect sizes were significantly different from zero;
5. Effect size was measured by a random-effects, because the variables involved did not meet homogeneity of variance assumptions ($p=.002 < .05$) (Hunter & Schmidt, 2004).
6. Publication bias was assessed via a funnel plot to display the distribution of effect sizes.
7. For interpretation of effect size, the criteria of Plonsky and Osward (2014) were considered in this study, namely: small (*r*=.25), medium (*r*=.40) and large(*r*=.60) in SLA research field. .

and for the dependent variables of corrective feedback:

1. Measures of score gains of CF (GJT, Oral production, and combined)
2. Score gains of post-tests (immediate, delayed, and combined)
3. Types of corrective feedback (implicit and explicit)
4. Research setting (classroom and laboratory)
5. Academic status of participants (secondary school, university, and language program)

D. The research synthesis

A total of 12 studies published between 2000 and 2022 were analysed (data collection was finished in March 2022). The whole dataset included 41 samples and 489 learners (see Appendix A). Among 12 studies, 10 are peer reviewed published articles, one was Ph. D dissertations, and one was book chapter from 2007 to 2020. Since 2012, the number of studies on the correlation between WM and CF has been increased in SLA research field.

A funnel plot was created to calculate more accurate effect sizes, a funnel plot was created (see Figure 1). According to Figure 1, the funnel plot of this study shows the following patterns there is no availability bias because the data the studies were almost symmetrically distributed around the mean. According to the results of classic fail-safe N, there was no publication bias (P =0.00 <0.05).

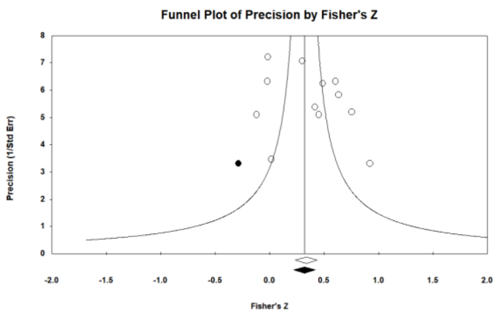

Figure 1. Forest plot of precision by effect sizes

V. RESULT

The correlation coefficients (*r* values) represent the associations between working memory and corrective feedback. The criteria of correlation coefficient followed the criteria of Plonsky and Osward (2014): small (*r*=.25), medium (*r*=.40) and large(*r*=.60) in SLA research field.

RQ 1: How does working memory relate to language gains of CF?

The answer to RQ1 is that WM was approximately significantly correlated with the score gains of CF in post-tests with the approximately medium effect size (*r* = 321 <.40) (Plonsky & Osward, 2014). As shown in Table 2, the overall random effect size between working memory and corrective feedback was *r*=.321, *p* <.05.

RQ 2: What are more important variables to moderate the effect of working memory on language gains of corrective feedback?

In terms of features of methodology, measures of CF outcomes and academic status of learners were more important to moderate the effect of WM on language gains of CF than other factors, especially types of WM score gains measured by GJT (*r=.41*> .25) and academical status of learners (*r*> .25) (see Table 2).

In sum, WM was negatively correlated with the language gains of CF with the small effect size. Measures of CF outcomes and academic status were two important moderating variables to affect the effect of WM.

TABLE II
EFFECTS OF WORKING MEMORY ON CORRECTIVE FEEDBACK

Variables	**subset**	***k***[a]	***r***[b]	**95%CI**		P[c]
				Upper	**Lower**	
Full data		12	0.321	0.463	0.162	.000*
Measure of WM						
	PSTMd	8	-0.049	0.765	-0.667	0.893
	EWMe	8	0.049	0.667	-0.765	0.893
Type of CF						
	Explicit	9	-0.072	0.474	-0.618	0.796
	Implicit	9	0.072	0.618	-0.474	0.796
Measures of CF outcomes						
	EITf	6	-0.188	0.465	-0.84	0.573
	GJTg	10	-0.413	0.162	-0.988	0.159
	Oral production	3	-0.172	0.492	-0.835	0.612
Timing of post-tests						
	Immediate	11	-0.062	0.859	-0.735	0.878
	delayed	8	0.062	0.735	-0.859	0.878
Research settings						
	Classroom	9	-0.018	0.407	-0.442	0.935
	Laboratory	3	0.018	0.442	-0.407	0.935
Academic status						
	Secondary school	2	0.307	0.829	-0.215	0.249
	University	7	-0.344	0.113	-0.801	0.249
	Language program	3	0.307	0.215	-0.829	0.856

Notes:

a. Correlation coefficient (*r*)
b. Number of studies in the analysis (*k*)
c. * p<0.05

d. PSTM=Phonological short-term memory
e. EWM=Executive working memory
f. EIT=Elicited imitation test
g. GJT= Grammaticality judgment test

VI. DISCUSSION

This meta-analysis investigated the role of working memory in the language gains of corrective feedback and how this relationship might be moderated other factors that vary in research design and instruction. 12 studies on the effect of WM on CF were analysed to provide a quantitative synthesis of findings on the correlation between WM and the effectiveness of CF as well as an examination of common variables in research and instructional design that may moderate this relationship. The results revealed that WM is likely to be a predictor variable in the effectiveness of CF, and that some factors in instructional and research design moderate this relationship either increasing or easing demands on working memory capacity.

RQ 1 asked whether WM would be correlated with the effectiveness of CF in terms of language gains. This question can be answered affirmatively WM capacity was correlated with language gains of CF overall with a small effect size (*r*=.072). Furthermore, both EWM and PSTM were positively correlated with outcomes of CF with small effect sizes, which is consistent with previous research (Li 2017; Linck et al.). In addition, the correction between PSTM and CF was same with the correlation between EMW and CF. This is inconsistent with the findings of William's (1999) and Wen's (2015) studies on these variables.

RQ 2 then asked whether working memory would be more important for language gains of corrective feedback in some instructional and research conditions than in others. This question can also be answered affirmatively. The results of the present meta-analysis indicate that the role of WM in the effectiveness of correctively is thus likely to be at least partially impacted by different forms of instruction, making variation in learners' working memory more critical in some instructional conditions than in other instructional conditions.

In particular, WM appears to be more important when implicit types of corrective feedback are provided than when explicit types of corrective feedback are provided. Although working memory was weakly positively correlated with both implicit corrective feedback (*r* =.072), and negatively correlated with explicit feedback (*r*=.72), the effect size of the correlation was same in the case of explicit corrective feedback than it was for implicit corrective feedback. To be specifically, WM was negatively correlated with corrective recasts as explicit CF with a small effect size (*r*= -.372), while negatively correlated with recasts as implicit CF (*r*=.149). These findings consistent with previous research by Li (2017), Lado (2008) and Norris and Ortega (2000). The most frequently investigated implicit feedback in the analysed studies were recasts. In contrast, the most common forms of explicit CF investigated in the studies analysed were metalinguistic feedback and corrective recasts. The reason for the increased demands on working memory in language gains of explicit correct feedback might thus be related to the fact that explicit corrective feedback tends to interrupt learners focus of attention on meaning when using language and requires them to focus their attention specifically on form. The difference in effect size in the correlations may be because implicit corrective feedback does not always require learners to divide their attention by switching their focus from meaning to form. Li (2010) also found that explicit feedback made increased working memory demands on learners' by attracting that attention and language resources rather than meaning. The present study thus provides some support for different working memory demands of implicit and explicit corrective feedback, suggesting that that learners with high working memory capacity may best be able to benefit from the affordances of implicit corrective feedback, whereas learners with average levels of working memory capacity may be disadvantaged in situations where implicit CF only is provided. In fact, previous researchers have tried to combine implicit and explicit feedback to mitigate the demands of these two forms of feedback. Ahmadian (2020), for example, combined recasts with explicit corrective feedback to mitigate the differential demands of explicit and implicit feedback.

In respect to the outcome measures of the transfer of CF, WM was more important when language gains of CF were measured by GJT than by EIT, oral production and written production. Language gains by GJT are direct measures of the effect of CF with accurate in data than the three measures. the effect size of the correlation was larger in the case of GJT with a medium effect size (*r*=- .41) than the other three ways with small effect sizes (*r* <.25). These findings consistent with previous research by. GJT were widely used in the SLA field (Ellis & Roever, 2021). Li (2017) also compared the effect of EIT and narrative production.

As for two types of post-tests, working memory was negatively correlated with immediate post-test (*r*=.062), working memory was positively correlated with the delayed post-test (*r*=0.62, *p* <.05), although both effects were of a small magnitude according to Plonsky & Oswald's (2014). That means that the present study thus provides some support for different working memory demands of immediate post-tests and delayed post-tests, suggesting that that learners with higher WM capacity had better performance in delayed post-test than learners with lower capacity, which is consistent with findings of previous studies (Li et al, 2016).

WM was also correlated with language status of learners' academic status (*r* >.25). Because different types of students have different levels of previous language knowledge. University students had more previous knowledge on grammatical structures than learners from language program and secondary schools. For example, in this meta-analysis, secondary school students were provided by new grammar structures, thus depended more on WM than the other structures with previous knowledge (Li et al., 2016, Fu & Li, 2019).

Finally, although there was no obvious difference in the effect size on language setting (classroom and laboratory), WM was positively correlated with a laboratory setting and negatively correlated with a classroom setting. The one reason may that a laboratory setting can make target structures more salient and easier to detect for learners. In terms of the classroom setting, three samples of the classroom studies were used corrective recasts, a typical explicit CF, focused on one target structure to mitigate the side effect of classroom.

VII. CONCLUSION

In general, WM was negatively correlated with the language gains of CF. In addition, two important moderating variables of WM were methods to measure CF outcomes and academic status of students. It should be kept in mind that many of the variables meta-analysed in the present were based on a limited number of previous studies. More studies are needed to further investigate the effect of WM on CF. In addition, more consistent measurements could be used to measure WM and outputs of corrective feedback to make the results more comparable among different studies. Finally, more studies could be conducted in second language settings to provide a full view of the influence of the language context, not just foreign language settings.

It should be kept in mind that many of the variables meta-analysed in the present were based on a limited number of previous studies. More studies are needed to further investigate the effect of WM on CF. In addition, more consistent measurements could be used to measure WM and outputs of corrective feedback to make the results more comparable among different studies. Finally, more studies could be conducted in second language settings to provide a full view of the influence of the language context, not just foreign language settings.

REFERENCES

*Ahmadian, M. J. (2020). Explicit and implicit instruction of refusal strategies: Does working memory capacity play a role? *Language Teaching Research, 24*(2), 163-188. doi:10.1177/1362168818783215

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nat Rev Neurosci, 4*(10), 829-839. doi:10.1038/nrn1201

Baddeley, A. D. (2012). Working Memory: Theories, Models, and Controversies. *Annu Rev Psychol, 63*(1), 1-29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A. D. (2015). Working memory in second language learning. . In Z. Wen, M. Mota, & A. McNeil (Eds.), *Working memory in second language acquisition and processing* (pp. 17-28). Bristol: Multilingual Matters.

Baddeley, A. D. (2017). Modularity, working memory and language acquisition. *Second language research, 33*(3), 299-311. doi:10.1177/0267658317709852

Brown, D. (2016a). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research, 20*(4), 436-458. doi:10.1177/1362168814563200

Brown, D. (2016b). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language teaching research : LTR, 20*(4), 436-458. doi:10.1177/1362168814563200

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences ED 2nd ed*. Pennsylvania: Philadelphia, Pennsylvania, USA: Lawrence Erlbaum Associates.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior, 19*(4), 450-466. doi:10.1016/S0022-5371(80)90312-6

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

Egi, T. (2010). Uptake, Modified Output, and Learner Perceptions of Recasts: Learner Responses as Language Awareness. *Modern Language Journal, 94*(1), 1-21. doi:10.1111/j.1540-4781.2009.00980.x

Ellis, R. (1995). Uptake as language awareness. *Language Awareness, 4*(3), 147-160. doi:10.1080/09658416.1995.9959877

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual Differences in Working Memory and Comprehension: A Test of Four Hypotheses. *J Exp Psychol Learn Mem Cogn, 18*(5), 972-992. doi:10.1037/0278-7393.18.5.972.

Ellis, R., & Roever, C. (2021). The measurement of implicit and explicit knowledge. The Language Learning Journal,49(2), 160-175.

*Fu, M., & Li, S. (2019). The associations between individual differences in working memory and the effectiveness of immediate and delayed corrective feedback. *Joural of Second language studies, 2*(2), 233-257. doi:10.1075/jsls.19002.fu

Gass, S., & Mackey, A. (2007). Input, interaction and output in second langauge acquisition. In B. VanPatten & J. N. Williams (Eds.), *Theories in second language acquisition* (pp. 175-199). New York/London: Routledge.

Gathercole, S. E., & Baddeley, A. (1993). *Working memory and language* Hove, UK: L. Erlbaum Associates.

*Goo, J. (2012). Corrective feeback and working memory capacity in interaction-driven L2 learning *Studies in second language acquisition, 34*(3), 445-474. doi:10.1017/S0272263112000149

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological methods*, 9(4), 426.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks: Thousand Oaks: SAGE Publications Inc. (US).

Kartchava, E., & Ammar, A. (2014). The noticeability and effectiveness of corrective feedback in relation to target type. *Language Teaching Research, 18*(4), 428-452. doi:10.1177/1362168813519373

*Lado, B. (2008). *The role of bilingualism, type of feedback, and congitive capacity in the acquisiton of non-primary langauges :a computer-based study*. Unpublished Ph.D thesis. Georgetown University. Washington, DC.

Li, S. (2010). The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis. *Language learning, 60*(2), 309-365. doi:10.1111/j.1467-9922.2010.00561.x

*Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal, 97*(3), 634-654. doi:10.1111/j.1540-4781.2013.12030.x.

Li, S. (2014). The interface between feedback type, L2 proficiency, and the nature of the linguistic target. *Language Teaching Research*, 18(3), 373-396.

Li, S. (2015). Cognitive differences and ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition*. New York: Routledge.

Li, S. (2017). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski-Weiss (Ed.), *Expanding Individual Difference Research in the Interaction Approach. Investigating learners, instructors, and other interlocutors* (pp. 41-70): John Benjamins Publishing Company.

Li, S. (2018). Data collection in the research on the effectiveness of corrective feedback. In A. Gudmestad & A. Edmonds (Eds.), *Critical reflections on data in second langauge acquisition* (pp. 33-66). Amsterdam:: John Benjamins publising company.

*Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied Psycholinguistics, 40*(3), 693-722. doi:10.1017/S0142716418000796.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychon Bull Rev, 21*(4), 861-883. doi:10.3758/s13423-013-0565-2

Long, M. (2007). Recasts in SLA: the story so far. *Problems in SLA* 75-116.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in second language acquisition*, 26(3), 399-432.

Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language learning*, 59(2), 453-498.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language teaching*, 46(1), 1-40.

Lyster, R., & Saito, K. (2010a). Oral feedback in classroom SLA. *Studies in Second Language Acquisition, 32*(2), 265-302. doi:10.1017/S0272263109990520

Lyster, R., & Saito, K. (2010b). Oral feedback in classroom SLA: A Meta-Analysis. *Studies in second language acquisition, 32*(2), 265-302. doi:10.1017/S0272263109990520.

Mackey, & Goo, J. (2007). Interaction research in SLA: a meta-analysis and research synthesis In A. Mackey (Ed.), *Conversational interaction in SLA: A collection of empirical studies* (pp. 407-452). New York: Oxford university press.

Mackey, A. (2006). Feedback, Noticing and Instructed Second Language Learning. *Applied linguistics, 27*(3), 405-430. doi:10.1093/applin/ami051

Mackey, A., & Philp, J. (1998). Conversational Interaction and Second Language Development: Recasts, Responses, and Red Herrings? *Modern Language Journal, 82*(3), 338-356. doi:10.1111/j.1540-4781.1998.tb01211.x

*Mackey, A., & Sachs, R. (2012). Older Learners in SLA Research: A First Look at Working Memory, Feedback, and L2 Development: Older Learners in SLA Research. *Language learning, 62*(3), 704-740. doi:10.1111/j.1467-9922.2011.00649.x

*Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the Relationship Between Modified Output and Working Memory Capacity. *Language learning, 60*(3), 501-533. doi:10.1111/j.1467-9922.2010.00565.x

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Amsterdam: Benjamins.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language learning, 50*(3), 417-528. doi:10.1111/0023-8333.00136

Peterson, R. A., & Brown, S. P. (2005). On the Use of Beta Coefficients in Meta-Analysis. *J Appl Psychol, 90*(1), 175-181. doi:10.1037/0021-9010.90.1.175

Plonsky, L., & Brown, D. (2014). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second language research*, *31*(2), 267-278. doi:10.1177/0267658314536436

Plonsky, L., & Oswald, F. L. (2012). How to do a meta-analysis. Wiley-Blackwell.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language learning*, 64(4), 878-912.

*Revesz, A. (2012). Working Memory and the Observed Effectiveness of Recasts on Different L2 Outcome Measures. *Language learning*, *62*(1), 93-132. doi:10.1111/j.1467-9922.2011.00690.x

Robinson, P. (2005). APTITUDE AND SECOND LANGUAGE ACQUISITION. *Ann Rev Appl Linguist, 25*, 46-73. doi:10.1017/S0267190505000036

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second langauge acquisition: A meta-analysis of the research In J. Norris & L. Ortega (Eds.), *Synthesizing research on langauge learning and teaching* (pp. 131-164). Amsterdam: Benjamins.

*Sachs, R. (2010). *Individual differences and the effectiveness of visual feedback on reflexive binding in Japanese (Unpublished doctoral dissertation).* Georgetown University. Washington, DC.

Schmidt, R. (1994). Implict learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Explicit and implicit learning of langauge* (pp. 165-209). London: Academic press.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301-322). Oxford: Oxford University Press.

*Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytic ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford: Oxford University Press.

Waters, G. S., & Caplan, D. (1996). The Measurement of Verbal Working Memory Capacity and Its Relation to Reading Comprehension. *The Quarterly Journal of Experimental Psychology Section A: Working Memory. Guest Editors: Hitch and Logie*, *49*(1), 51-79. doi:10.1080/713755607

Wen, Z. (2015). Working memory in second language acquisition and processing: The Phonological/Executive model. In Z. Wen, M. Mota, & A. MacNeil (Eds.), *Working memory in second language acquisition and processing* (pp. 41-62). Bristol: Multilingual Matters.

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in second language acquisition*, 32(2), 235-263.

*Yang, Y. L., Zhang, L. J., & Chen, L. (2019). Effects of recasts and form-focused instruction on the acquisition of novel vocabulary. *Applied Linguistics Review*, *10*(4), 603-624. doi:10.1515/applirev-2017-0084

*Zhao, Y. (2015). *The effects of explicit and implicit recasts on the acquisition of two grammatical structures and the mediating role of working memory (Unpublished doctoral dissertation).*. University of Auckland. New Zealand.

*Zhao, Y., & Ellis, R. (2020). The relative effects of implicit and explicit corrective feedback on the acquisition of 3rd person-s by Chinese university students: A classroom-based study. *Language Teaching Research*, 1362168820903343.