Article:

McManus, Kevin and Marsden, Emma Josephine orcid.org/0000-0003-4086-5765 (2018) Online and offline effects of L1 practice in L2 grammar learning : a partial replication. Studies in Second Language Acquisition. pp. 459-475. ISSN 1470-1545

https://doi.org/10.1017/S0272263117000171

Online and offline effects of L1 practice in L2 grammar learning

A partial replication

Accepted in *Studies in Second Language Acquisition*

Accepted version, 31st January 2017

NOT FOR CIRCULATION AS PENDING MINOR AMENDMENTS

Kevin McManus[1] & Emma Marsden[2]

[1]The Pennsylvania State University, USA, [2]University of York, UK

Address for correspondence:

Kevin McManus
Department of Applied Linguistics
209 Sparks Building
The Pennsylvania State University
University Park, PA 16802
USA

kmcmanus@psu.edu

Online and offline effects of L1 practice in L2 grammar learning. A partial replication

Abstract

This study partially replicates McManus and Marsden (2016), who found that providing L1 explicit information (EI) plus task-essential practice led L2 learners to make more accurate and faster interpretations of French morphosyntax. The current study removed the original study's L1 EI component to examine the extent and nature of its role and that of the remaining L1 practice. This design tested whether providing L1 task-essential practice only (alongside a core treatment of L2 EI plus practice) resulted in similar online and offline learning gains compared to the original study's L1 EI plus L1 practice. We used the same online and offline tests, with a similar population of English-speaking learners of L2 French (n=19). Compared to L2 EI plus L2 practice, the findings suggest that L1 practice did not generally benefit accurate interpretation or speed of online processing, indicating that the original study's L1 EI contributed to the benefits previously observed.

Understanding how explicit information (EI) and practice shape second language development remains an active goal of instructed second language acquisition research. Evidence indicates that, in some cases, providing EI benefitted learning very little (Sanz & Morgan-Short, 2004; Stafford, Bowden, & Sanz, 2012; VanPatten & Oikkenon, 1996). In other cases, pre-practice EI has been found to accelerate the point at which learners begin to interpret the input more accurately, increasing the efficiency of practice (Fernández, 2008; Henry, Culman, & VanPatten, 2009; VanPatten & Borst, 2012). Research has focused on the roles played by EI and practice in L2 learning, especially the effects of different types of EI and/or practice (DeKeyser & Prieto Botana, 2015), such as EI about L1 and L2 properties for crosslingusitically problematic target features (Horst, White, & Bell, 2010; Spada, Lightbown, & White, 2005).

To this end, McManus and Marsden (2016) (henceforth, M&M) examined whether providing instruction about L1 form-meaning connections would benefit the L2 learning of the French *Imparfait*, a late-acquired form by English-speaking learners because of complex L1-L2 differences (Author, XXXX; Izquierdo & Collins, 2008). In a pre-, post-, delayed post-test design over 12 weeks, that study included three groups of L1 English learners of L2 French: two treatments (L2-only and L2+L1) and a test-only control. The L2-only group received a core treatment of EI about the L2 and task-essential practice in the L2. The L2+L1 group received the exact same core L2 treatment (as received by L2-only) as well as EI about the L1 and task-essential practice in the L1 (for a description of the instruction, see M&M, and IRIS for materials). Pre, post- and delayed post-test measures were online (self-paced reading, SPR) and offline (context-sentence matching tests, CMT, in listening and reading). These measured sensitivity to the *Imparfait*'s ongoing and habitual meanings, in which a French stimulus either matched or mismatched the ongoing (or habitual) meaning of a preceding event description in

English. Results showed no significant improvement over time for the Control group in both speed or accuracy of interpreting the *Imparfait*. The L2-only group's gains were mostly in offline tests at Post (not Delayed) in matched trials, with negligible effects in mismatched trials. Improvement up to Delayed was found in CMT-Read for the *Imparfait*'s ongoing meaning (matched trials). Improvement up to Post was found for CMT-Listen (habitual, matched) and SPR (ongoing, matched), but these were lost by Delayed. The L2+L1 group, in contrast, improved in all measures up to Delayed (six weeks after the intervention), in both offline and online outcome measures, in matched and mismatched trials. Critically, SPR results demonstrated sensitivity to aspectual anomalies, as evidenced by slower reading times (RT) in mismatched trials for the L2+L1 group only.

While some previous research has provided EI about the L1 (but without L1 practice), with mixed learning outcomes (González, 2008; Horst et al., 2010; Spada et al., 2005), M&M is understood to be the first to have employed task-essential practice in the L1. It is possible that this L1 task-essential practice contributed to the online and offline benefits observed, a possibility partly supported by the fact that the effects of L1 EI alone remain inconclusive. However, the extent to which L1 practice without L1 EI would result in learning benefits similar to those observed in the original study remains unclear. The present study contributes to that gap by means of a partial replication using the exact same procedure and outcome measures as in the original study. As received by both treatment groups in the original study, our intervention provided L2 EI plus L2 practice. In addition, it included L1 task-essential practice, like M&M's L2+L1 only group, but, unlike that group, it did *not* include any EI about the L1. Critically, this design allowed us to investigate whether providing L1 task-essential practice without L1 EI would result in the same learning gains as found for L1 task-essential practice with L1 EI.

MOTIVATION FOR THE PRESENT STUDY

A small number of studies have provided EI about both L2 and L1, finding some benefits offline for learning L2 morphosyntax. González (2008) compared L2+L1 EI with L2-only EI for lexical terminativity. L1 EI described how terminativity operates in Dutch (L1). Participants used that EI to construct rules about the (un)grammaticality of Dutch sentences. For the L2 EI, participants first described what they knew about Spanish past tenses before receiving EI about terminativity and perfectivity in Spanish (L2). That EI was then used to construct rules about the (un)grammaticality of Spanish sentences. Pre-Post results showed that L2+L1 EI was more beneficial than L2-only EI in a cloze test, but effects of EI type on written composition were inconclusive due to test design limitations. Gains, however, were not fully substantiated because the test (identical at Pre and Post) may have tapped particularly well into the same kind of knowledge that had been taught (largely controlled, offline, explicit) and no Delayed Posttests were administered. Spada et al.'s (2005) EI included a contrastive component that explicitly directed learners' attention to L1-L2 differences and similarities. L2+L1 EI was complemented with whole-class games and activities in the L2 (but not in L1) that required learners to produce the target feature. Data from a range of task types (passage correction, oral production, judgments) indicated an advantage for L2+L1 EI. However, further investigation is required because between-group differences at Pretest may have affected the results and, importantly, L1 EI was not used systematically or to the same extent by all the teachers in the study. In sum, previous research providing L1 EI has indicated potential advantages. Thus, the extent to which the learning gains observed by M&M were due to providing L1 EI or, in fact, to the L1 practice remains unclear.

An additional motivation for the current study relates to the nature of the outcome measures to date (although see Andringa & Curcic, 2015; Author, XXXX; Henry, 2015, but these studies did not incorporate L1 information or practice). That is, research into L1 EI has not measured effects of instruction in online processing. This is of particular interest given that many interventions have aimed to influence the processing of input.

In terms of potential learning mechanisms at play, M&M argued that the L1 EI may have helped to establish new knowledge (or make existing knowledge explicit) that two concepts can be expressed by one morpheme in the L1 system (ed -> perfective and habitual), and that one concept (past ongoingness) is expressed in the L1 by one set of morphemic features (BE_past + ing) yet is expressed in the L2 system by a morpheme that also expresses another concept (ait -> habituality and ongoingness). This clearer conceptual distinction, along with its mappings to oral and written morphemic distinctions, may have been proceduralised/automatised during L1 practice. This knowledge (or rehearsed processing routine) may have helped the L2+L1 group to parse and/or anticipate cues in the linguistic context needed to disambiguate the *Imparfait*'s habitual and ongoing functions more accurately and faster than the L2-only group. (For evidence of L1 activation during L2 processing and production, see Runnqvist, Gollan, Costa, and Ferreira, 2013; Sanoudaki and Thierry, 2014). Given the above, the possibly different roles played by the EI and practice are unclear. For example, L1 practice alone may have been sufficient to establish such knowledge or processing routines, or, alternatively, it may not have provided any additional benefit.

In sum, to examine the role of L1 practice in L2 learning, the present replication examines the effect of providing L1 practice without L1 EI. This allowed us to ascertain the

extent to which providing L1 EI with practice or L1 practice only - both alongside a core L2 treatment of L2 EI and practice - resulted in similar learning gains, online and offline.

In terms of predicted outcomes, providing only L1 task-essential practice could result in similar learning gains as for L1 EI plus L1 practice i.e. results would pattern with M&M's L2+L1 group. This outcome would be accounted for by, for example: a) learners being able to *induce* L1 EI from the L1 practice (see DeKeyser, 2007, AUTHOR XXX, Sanz & Morgan-Short, 2004), i.e. this account assumes that EI about the L1 can benefit L2 learning; or b) EI about the L1 does not, in fact, affect or interact with L2 processing mechanisms, but practice in interpreting the L1 does.

On the other hand, as no previous research has examined L1 practice or online effects, it remains possible that providing M&M's L1 EI contributed to the observed benefits. For example, some research suggests that EI can make the processing of (more complex) language features more efficient (e.g., Henry et al., 2009; VanPatten & Borst, 2012) by helping learners correctly interpret sentences with fewer practice items than those not receiving EI. Due to the nature of the cross-linguistic learning problem under investigation here, it is possible that EI about the *L1* could also help improve the effectiveness of practice. Also, as noted above, the EI about the L1 could help conceptualize ongoingness *versus* habituality and their morphological mappings. In short:

1. Results that pattern like those from M&M's L2+L1 group would suggest a less important role for L1 EI, indicating that the benefits observed for the L2+L1 group were not (uniquely) attributable to the intervention's L1 EI component. This would be compatible to some extent with previous findings that providing EI (albeit about the

L2) was not necessary (Author, XXX, Sanz and Morgan-Short, 2004) or beneficial for online processing (Andringa and Curcic, 2015).

2. Results that pattern more closely to those from M&M's L2-only group would suggest a more important role for L1 EI, indicating that the benefits observed for the L2+L1 group were, at least in part, attributable to the intervention's L1 EI component. This would corroborate findings by González (2008), Horst et al. (2010), Spada et al. (2005) that instruction about the L1 can be beneficial. It would also be compatible to some extent with the findings of Henry et al. (2009) and VanPatten and Borst (2012) that EI (albeit about L2) made practice more efficient.

METHOD

Participants

Participants were 19 English-speaking learners of French as a foreign language in semester two of a four-year BA Honours degree in French, recruited from two large universities in England. Following M&M, every participant was a native speaker of English, had completed A-level French (CEFR B2), and had not spent more than six weeks in a French speaking country. Participants' mean age was 19, and the mean time spent in a French speaking country was two weeks.

Design

The study design, timescale, procedure and outcome measures were exactly as reported by M&M. The design included a Pretest (week 1), Posttest (week 5), and Delayed Posttest (week 12), and the treatment was delivered in four 45-minute sessions over three weeks, totalling 3.5

hours. Sessions 1 and 2 were delivered in Week 2, Session 3 in Week 3, and Session 4 in Week 4. As in the original study, none of the participants received any instruction between posttests.

Instructional treatment

The treatment included L2 EI and task-essential form-meaning mapping practice in interpreting the French *Imparfait*. It also included task-essential form-meaning mapping practice in interpreting equivalent English forms for ongoingness and habituality. These treatment components are briefly described below (see M&M for descriptions and examples).

EI about the L2 was provided pre-practice (approximately 5 minutes) and during the task-essential practice activities following incorrect answers. The amount of L2 EI received during-practice was very similar across treatment groups (see Supplementary Materials). Identical to M&M, the pre-practice EI depicted conceptual-semantic information via a short video, image or sound file of events. Then the appropriate aural and written forms were presented, and information given about how to interpret their meaning. No L1 EI was provided, either pre- or during practice.

Immediately following the L2 pre-practice EI, participants undertook task-essential form-meaning mapping practice in which learners chose the form's meaning from fixed options (as in, e.g., VanPatten, 2002). Sample sentences are shown in (1) for past ongoingness and (2) for past habituality. English practice items (k=112) were interweaved with French practice items (k=384). The French items drew on 48 lexical verb types: each type occurred eight times with *Imparfait*: four for reading (two habitual, two ongoing); four for listening (two habitual, two ongoing). The English items drew on 14 lexical verb types, which were also balanced across listening/reading and habitual/ongoing.

(1) Elle… mangeait un sandwich quand la cloche a sonné
'She…was eating a sandwich when the bell rang'

(2) Il… buvait une bière quand il sortait avec sa femme
'He… drank/used to drink a beer when he went out with his wife'

Outcome measures

Context-matching tests (listening and reading). All participants took two context-matching tests: first listening (CMT-Listen), then reading (CMT-Read), each with 24 target and 8 filler trials. Each trial consisted of: (i) a written English context: two sentences describing either a habitual or an ongoing activity and (ii) a French stimulus: a two-clause French sentence that either matched (k=12) or mismatched (k=12) the meaning of the English context. Two versions of the test were counter-balanced among participants across time of test.

In both CMTs the English context appeared on screen for 10 seconds. Then, the French stimulus appeared orally (CMT-Listen) or as text (CMT-Read). Participants rated how good the match was between the meaning of the French stimulus and English context by pressing a number on the keyboard from 1 ("very good"), 2 ("good"), 3 ("neither good nor bad"), 4 ("poor"), 5 ("very poor"), with a separate option for "I don't know" (9). The written French stimulus remained on screen until a number was pressed, and then participants could not change their answer. The task was untimed and took approximately 20-25 minutes.

Self-paced reading test. The self-paced reading (SPR) test was administered after the CMTs and used 16 items from the CMT-Listen, with eight context-stimulus matches and eight mismatches. Half the items were followed by Yes/No comprehension questions. For each trial, the English context appeared for 10 seconds before an X appeared in the centre of the screen. A spacebar press brought up the first and then each subsequent word of the French stimuli in the

centre of the screen. After the last word, the next screen displayed “END”. Participants were instructed to read as quickly as possible and understand. Reaction times were collected from each word.

Data scoring and analysis

CMT responses were coded as follows: 5 points per correct response (i.e. 1 or 2 for matched, and 4 or 5 for mismatched trials); 3 points per midway responses (3 for matched and mismatched trials); 1 point per incorrect response (4 or 5 for matched, and 1 or 2 for mismatched trials). Cronbach’s alphas were: CMT-Listen version A ($\alpha = .81$), version B ($\alpha = .86$), CMT-Read version A ($\alpha = .74$), version B ($\alpha = .79$).

In the SPR, a ‘critical word’ in each French stimulus disambiguated ongoing from habitual meaning of the *Imparfait* and thus determined whether the sentence matched or mismatched the English context. Reading times (RT) for the critical word were calculated from the onset of the critical word to the onset of the next word. Whole sentence RTs are the time taken to read from the onset of the first word to the onset of the ‘END’ screen. We analysed the raw RT data, removing critical word RTs less than 150ms and greater than 2000ms.

As none of the datasets were normally distributed (according to Shapiro-Wilks tests, all datasets $p < .05$), we present the results of nonparametric tests.

First, Friedman tests were used to compare Pre, Post, and Delayed scores. If a significant result was found, within-subject comparisons were made using Wilcoxon Signed-Rank tests with Bonferroni corrected alpha levels between pairs of test results: Pre-Post, Pre-Delayed, and Post-Delayed. Second, we tested for parity (Kruskall-Wallis H tests) on all measures at Pretest between all four groups: the current group, L2+L1prac, and the three groups from M&M, L2+L1

(henceforth, L2+L1prac+EI), L2-only, and Control. Having established baseline parity, we then compared the current study's results with the patterns of results found by M&M, calculating between-group effect sizes (ES) and their confidence intervals (CI). Finally, we compared the current group's (L2+L1prac) performance on matched versus mismatched trials in the SPR, using Wilcoxon Signed-Rank tests. This analysis mirrored that of M&M.

In line with the original study, the alpha was set at 0.10 for the Kruskall-Wallis H and Friedman tests. The Bonferroni adjustment for the post-hoc Wilcoxon Signed-Rank tests revised the alpha to 0.10/3 = .033. For interpreting magnitudes of change, we present Cohen's *d* ES and CIs (95%) for ES. CIs for *d* that do not pass through zero can be considered reliable indicators of meaningful change (see Cumming, 2016; Field, 2013). We provide ES for comparisons with each of M&M's groups using the mean and standard deviation of the relevant group from M&M as the 'comparison/control' group. We also present between-group ES adjusted for differences at Pret. Following Plonsky and Oswald (2014), field-specific benchmarks are used for interpretation of Cohen's *d*: within subject: $d \geq .40 < .70$ (small), $\geq .70 < 1.00$ (medium) and ≥ 1.00 (large); between subject: $d \geq .60 < 1.00$ (small), $\geq 1.00 < 1.40$ (medium) and ≥ 1.40 (large).

RESULTS

HABITUAL TRIALS

Descriptive statistics for accuracy (CMTs) and reaction times (SPR) for the L2+L1prac group are shown in Table 1. All participants scored 100% on the SPR comprehension questions.

[TABLE 1 HERE]

Matched trials

Friedman tests showed no statistically significant improvement over time on any measure. All ES were negligible and CIs for *d* passed through zero.

- CMT-Read: $X^2(2)=1.80$, $p=.41$ (pre-post, *d* [CI] = .14 [-.50, .78]; pre-delayed, *d* [CI] = .19 [-.45, .82]; post-delayed, *d* [CI] = .03 [-.61, .66])
- CMT-Listen: $X^2(2)=2.44$, $p=.29$ (pre-post, *d* [CI] = .34 [-.30, .98]; pre-delayed, *d* [CI] = .04 [-.60, .67; post-delayed, *d* [CI] = -.24 [-.87, .41])
- SPR critical word: $X^2(2)=.74$, $p=.69$ (pre-post, *d* [CI] = -.25 [-.89, .39]; pre-delayed, *d* [CI] = -.13 [-.77, .50]; post-delayed, *d* [CI] = .14 [-.50, .78])
- SPR whole sentence: $X^2(2)=2.95$, $p=.23$ (pre-post, *d* [CI] = -.35 [-.98, .30]; pre-delayed, *d* [CI] = -.18 [-.82, .46]; post-delayed, *d* [CI] = .25 [-.39, .89])

Comparisons with the groups from M&M are presented in Table 2.

[TABLE 2 HERE]

Baseline parity was found between the four groups on all measures (CMT-Read, $X^2(3)=.081$, $p=.99$; CMT-Listen, $X^2(3)=1.21$, $p=.75$; SPR critical word, $X^2(3)=.45$, $p=.93$; SPR whole sentence, $X^2(3)=.29$, $p=.96$).

Mismatched Trials

Friedman tests showed no statistically significant improvement over time on any measure with only negligible ES and CIs that passed through zero:

- CMT-Read: $X^2(2)=2.20$, $p=.33$ (pre-post, *d* [CI] = .26 [-.39, .89]; pre-delayed, *d* [CI] = .31 [-.34, .94]; post-delayed, *d* [CI] = .04 [-.59, .68])
- CMT-Listen: $X^2(2)=.35$, $p=.84$ (pre-post, *d* [CI] = .05 [-.59, .69]; pre-delayed, *d* [CI] = .18 [-.46, .82]; post-delayed, *d* [CI] = .12 [-.52, .75])
- SPR critical word: $X^2(2)=.105$, $p=.95$ (pre-post, *d* [CI] = -.04 [-.67, .60]; pre-delayed, *d* [CI] = -.08 [-.71, .56]; post-delayed, *d* [CI] = -.05 [-.69, .59])
- SPR whole sentence: $X^2(2)=2.203$, $p=.33$ (pre-post, *d* [CI] = -.34 [-.97, .31]; pre-delayed, *d* [CI] = -.20 [-.83, .45]; post-delayed, *d* [CI] = .02 [-.46, .81])

Comparisons with the groups from M&M are presented in Table 2, with between-group trends summarizes in Table 3.

[TABLE 3 HERE]

RTs in matched versus mismatched trials

Wilcoxon Signed Rank tests showed no statistically significant between-trial (matched *versus* mismatched) differences for either SPR critical word or whole sentence RTs, with negligible ES and CIs for *d* that passed through zero:

- SPR critical word
    - Pre: Z=-.12, $p=.90$, *d* [CI] = -.05 [-.69, .59]
    - Post: Z=-.97, $p=.33$, *d* [CI] = -.29 [-.92, .36]
    - Delayed: Z=-.64, $p=.52$, *d* [CI] = -.11 [-.74, .53]
- SPR whole sentence
    - Pre: Z=-.12, $p=.90$, *d* [CI] = .21 [-.43. .85]
    - Post: Z=-.40, $p=.69$, *d* [CI] = .11 [-.53, .74]

- Delayed: Z=-.64, *p*=.52, *d* [CI] = -.25 [-.39, .89]

In contrast, RTs were statistically significantly slower in mismatched compared to matched trials for M&M's L2+L1prac+EI at Post and Delayed. There were negligible, non-statistically significant between-trial differences for whole sentence processing at Post and Delayed, suggesting the differences observed at the disambiguating critical word region were specific to that region, rather than due to generally faster processing.

ONGOING TRIALS

Descriptive statistics for accuracy (CMTs) and reaction times (SPR) are shown in Table 4. All participants scored 100% on the SPR comprehension questions.

[TABLE 4 HERE]

Matched trials

There was statistically significant change over time in one measure only, CMT-Read, with gains pre-post and pre-delayed, both with large ES. For all other measures, ES were negligible with CIs for *d* that passed through zero:

- CMT-Read: $X^2(2)$=26.16, *p*=.00 (pre-post, Z=-3.34, *p*=.00, *d* [CI] = 1.65 [.88, 2.35]; pre-delayed, Z=-3.83, *p*=.00, *d* [CI] = 2.30 [1.44, 3.06]; post-delayed, Z=-1.23, *p*=.22, *d* [CI] = .23 [-.41, .86])
- CMT-Listen: $X^2(2)$=1.97, *p*=.37 (pre-post, *d* [CI] = .55 [-.11, 1.18]; pre-delayed, *d* [CI] = .46 [-.20, 1.09]; post-delayed, *d* [CI] = -.06 [-.70, .58])

- SPR critical word: $X^2(2)=1.37$, $p=.50$ (pre-post, *d* [CI] = -.41 [-1.04, .24]; pre-delayed, *d* [CI] = -.22 [-.85, .42]; post-delayed, *d* [CI] = .15 [-.49, .78])
- SPR whole sentence: $X^2(2)=.32$, $p=.85$ (pre-post, *d* [CI] = -.36 [-1.00, .28]; pre-delayed, *d* [CI] = -.23 [-.86, .42]; post-delayed, *d* [CI] = .07 [-.57, .70])

Table 5 shows comparisons with the groups from M&M.

[TABLE 5 HERE]

Baseline parity between all four groups was found on all measures (CMT-Read, $X^2(3)=.87$, $p=.33$; CMT-Listen, $X^2(3)=.99$, $p=.99$; SPR critical word, $X^2(3)=.08$, $p=.99$; SPR whole sentence, $X^2(3)=1.02$, $p=.60$).

Mismatched trials

Statistically significant change over time was observed only for CMT-Read, due to gains pre-post and pre-delayed, but with small ES:

- CMT-Read: $X^2(2)=6.22$, $p=.045$ (pre-post, Z=-1.829, $p=.07$, *d* [CI] = .68 [.01, 1.32]; pre-delayed, Z=-2.22, $p=.027$, *d* [CI] = .86 [.18, 1.51]; post-delayed, Z=-.631, $p=.53$, *d* [CI] = .11 [-.53, .75])
- CMT-Listen: $X^2(2)=3.25$, $p=.19$ (pre-post, *d* [CI] = .59 [-.08, 1.22]; pre-delayed, *d* [CI] = .43 [-.22, 1.06]; post-delayed, *d* [CI] = -.13 [-.77, .51])
- SPR critical word: $X^2(2)=1.26$, $p=.53$ (pre-post, *d* [CI] = -.25 [-.88, .39]; pre-delayed, *d* [CI] = -.18 [-.81, .46]; post-delayed, *d* [CI] = .07 [-.57, .71])

- SPR whole sentence: $X^2(2)=2.95$, $p=.23$ (pre-post, *d* [CI] = -.34 [-.97, .31]; pre-delayed, *d* [CI] = -.32 [-.96, .32]; post-delayed, *d* [CI] = .02 [-.62, .66])

Comparisons with the groups from M&M are given in Table 5 and summarized in Table 3.

Baseline parity was found in all measures (CMT-Read, $X^2(3)=2.26$, $p=.52$; CMT-Listen, $X^2(3)=.310$, $p=.96$; SPR critical word, $X^2(3)=.23$, $p=.97$; SPR whole sentence, $X^2(3)=1.11$, $p=.58$).

RTs in matched versus mismatched trials

Wilcoxon Signed Rank tests showed no between-trial differences for either critical word or whole sentence RTs in the SPR, and negligible ES with CIs for *d* passing through zero:

- SPR critical word
    - Pre: Z=-.805, $p=.42$, *d* [CI] = -.20 [-.84, .44]
    - Post: Z=-.765, p=.44, *d* [CI] = -.29 [-.92, .36]
    - Delayed: Z=-.52, $p=.60$, *d* [CI] = -.22 [-.85, .42]

- SPR whole sentence
    - Pre: Z=-.121, p=.904, *d* [CI] = .08 [-.56, .72]
    - Post: Z=-0.80, p=.936, *d* [CI] = .08 [-.56, .71]
    - Delayed: Z=-.322, p=.748, *d* [CI] = .12 [-.52, .75]

In contrast, M&M found small ES for between-trial differences at Post and Delayed in L2+L1(prac+EI) only for critical word and whole sentence processing.

Summary of findings

M&M found increased accuracy and speed of interpretation of the *Imparfait* at Delayed following a treatment of L1 EI plus L1 task-essential form-meaning mapping practice (in addition to L2 EI plus practice). We partially replicated that original study to examine the role played by L1 practice by removing the L1 EI but retaining the L1 practice (and the core L2 EI and practice). We used the original study's design, procedures, and materials.

Our L2+L1prac group's results patterned very similarly to those of the L2-only group from the original study, and tended not to pattern as well with L2+L1prac+EI from M&M (see Table 3 for summary of between-group differences).

In habitual contexts, we found no significant improvement over time on any measure for L2+L1prac. In ongoing contexts, we found no significant improvement in CMT-Listen or SPR, but we did find Pre-Delayed improvement in CMT-Read, matched and mismatched. In the original study, the L2-only group also showed limited improvement, mostly offline in CMT-Read (Pre-Delayed, ongoing matched) and also in CMT-Listen (Pre-Post, Habitual and Ongoing matched).

Our L2+L1prac group showed no online sensitivity to anomalies as only negligible matched-mismatched RT differences were found, consistent with the L2-only group in M&M.

These results suggest that L1 task essential practice was not as beneficial as L1 practice plus L1 EI for learners' online processing and offline interpretation of the *Imparfait*. Further, the L1 practice component when in isolation seemed to contribute little compared to the L2-only intervention, as there were negligible ES throughout.

DISCUSSION AND CONCLUSIONS

Whereas the benefits of L2 EI, whether provided or induced, are well researched to date, the current study addressed the role of L1 EI. Extant classroom-based evidence suggested potential benefits of L1 EI for L2 learning (González, 2008; Spada et al., 2005). Although that research had different designs to the original study and this replication, our findings broadly align with it, extending it to show benefits of EI for the speed of *online* processing. This sheds some light on the extent to which, and circumstances under which, EI may interface at some level with online processing, but it also raises several questions for future research.

The findings support M&M's arguments that L1 EI helped to establish knowledge about L1 form-meaning mappings and this aided the automatization of L2 form-meaning mapping. The pre-practice EI may have raised awareness about L1 processing routines, including L1-entrenched attention and competing cues (Ellis, 2006; MacWhinney, 2005). For example, without EI, the habitual function of 'ed' may be difficult to induce from the input due to low saliency, multiple and complex form-function relations, and perhaps also prior instruction about the function of 'ed'. The EI may have enabled learners to establish more efficiently (with fewer trials) that an English context with 'ed' could map to a subsequent French *Imparfait* sentence.

We also note that this cross-linguistic complexity may explain, at least in part, why the *L2* EI + task essential practice component of our treatment did not seem to help learning, or, if it did, it helped only on the CMT-Read (the only measure to show gains). The L2 EI and practice may not have helped learning on other measures because concepts in the L1 may not have been sufficiently well established to facilitate mapping the appropriate L2 forms in more speeded contexts (CMT listening and SPR).

As noted by M&M, future research should also examine the extent to which the effectiveness of L1 EI and task-essential practice depends on the feature having cross-linguistically different form-function mappings. We also acknowledge that our outcome measures (English contexts followed by French stimuli) may have activated L1 representations. Future research requires measures that do not intentionally coerce the L1, such as oral production.

Although we isolated the L1 practice, we do not conclude that EI about the L1 *alone* was solely responsible for the benefits observed for in M&M. It is likely, we think, that the L1 EI in *combination* with practice in using that EI established and consolidated new L1 form-meaning mappings, which facilitated L2 form-meaning mapping during the L2 EI and practice. This also requires further research.

TABLES

Table 1

Habitual contexts for L2+L1prac group: Accuracy (CMTs) and reaction time (SPR) results

		CMT-Read	CMT-Listen	SPR critical word	SPR whole sentence
		Accuracy (max=5)	Accuracy (max=5)	RT (ms)	RT (ms)
MATCH					
Pre	M	3.97	3.84	617.41	26040.66
	(SD)	(.82)	(.79)	(207.89)	(17305.79)
Post	M	4.11	4.14	558.07	21214.41
	(SD)	(1.10)	(.95)	(253.98)	(9324.85)
Delayed	M	4.14	3.88	590.25	23512.08
	(SD)	(.98)	(1.23)	(192.49)	(8769.55)
MISMATCHED					
Pre	M	2.93	3.32	630.38	23091.64
	(SD)	(.88)	(.77)	(288.35)	(9035.67)
Post	M	3.23	3.37	621.09	20275.24
	(SD)	(1.41)	(1.12)	(178.13)	(7694.98)
Delayed	M	3.29	3.51	611.24	21540.82
	(SD)	(1.41)	(1.26)	(199.86)	(6632.42)

Table 2.

Habitual contexts: ES (Cohen's *d* including CIs for *d*) comparisons with Control, L2+L1prac+EI and L2-only from M&M, and **ES changes with effects adjusted for baseline differences**

	CMT-Read			CMT-Listen			SPR critical word			SPR whole sentence		
	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control
MATCHED												
Pre	-.09	-.07	-.08	.10	.17	.05	.01	-.04	-.04	.11	.01	.19
	[-.74, .57]	[-.72, .59]	[-.75, .59]	[-.56, .75]	[-.49, .82]	[-.61, .72]	[-.64, .67]	[-.69, .62]	[-.71, .62]	[-.54, .77]	[-.65, .66]	[-,49, .85]
Post	-.22	.25	.34	-.02	-.11	.47	1.18	.07	.02	.77	.18	.37
	[-.87, .44]	[-.41, .90]	[-.34, 1.00]	[-.67, .64]	[-.77, .54]	[-.22, 1.13]	[.45, 1.86]	[-.58, .73]	[-.65, .68]	[.08, 1.43]	[-.48, .83]	[-.31, 1.03]
Delayed	-.77	-.24	.65	-.96	-.29	.19	1.50	.59	.21	.92	.62	.70
	[-1.43, -.08]	[-.89, .42]	[-.04, 1.32]	[-1.63, -.25]	[-.94, .38]	[-.48, .86]	[.73, 2.20]	[-.09, 1.25]	[-.46, .87]	[.21, 1.58]	[-.06, 1.28]	[.00, 1.36]
Pre-post *d* change	**-.13**	**.32**	**.42**	**-.12**	**-.28**	**.42**	**1.17**	**.11**	**.06**	**.66**	**.17**	**.18**
Pre-delayed *d* change	**-.68**	**-.17**	**.73**	**-1.06**	**-.46**	**.14**	**1.49**	**.63**	**.25**	**.81**	**.61**	**.51**
MISMATCHED												
Pre	-.17	-.03	-.05	-.21	-.10	-.03	-.02	.02	.02	.10	-.42	-.05
	[-.82, .49]	[-.69, .62]	[-.72, .61]	[-.86, .45]	[-.75, .55]	[-.70, .63]	[-.67, .64]	[-.63, .67]	[-.64, .69]	[-.55, .76]	[-1.08, .25]	[-.72, .61]
Post	-.25	-.05	.10	-.50	.13	.27	.87	.14	.17	.49	.16	.11
	[-.90, .41]	[-.70, .61]	[-.56, .77]	[-1.16, .17]	[-.53, .78]	[-.40, .94]	[.17, 1.53]	[-.52, .79]	[-.50, .83]	[-.19, 1.14]	[-.50, .81]	[-.55, .78]
Delayed	-.95	.09	.16	-1.04	-.07	.84	1.00	.14	.14	.28	.13	.38
	[-1.62, -.24]	[-.57, .74]	[-.51, .83]	[-1.71, -.32]	[-.72, .59]	[.13, 1.51]	[.28, 1.67]	[-.52, .80]	[-.53, .80]	[-.38, .93]	[-.53, .78]	[-.30, 1.04]
Pre-post *d* change	**-.08**	**-.02**	**.15**	**-.29**	**.23**	**.30**	**.89**	**.12**	**.15**	**.39**	**.58**	**.16**
Pre-delayed *d* change	**-.78**	**.12**	**.21**	**-.83**	**.03**	**.87**	**1.02**	**.12**	**.12**	**.18**	**.55**	**.43**

Table 3

Summary interpretations of between-group ES, based on Plonsky & Oswald (2014)

			L2+EIprac vs. L2+L1prac+EI	L2+EIprac vs. L2-only	L2+EIprac vs. Control
HABITUAL					
Matched	CMT-Read	Post	N	N	N
		Delayed	S (<)	N	S* (>)
	CMT-Listen	Post	N	N	N
		Delayed	S/M (<)	N	N
	SPR-Critical	Post	M (slower)	N	N
		Delayed	L (slower)	N/S* (slower)	N
	SPR-Whole sentence	Post	S (slower)	N	N
		Delayed	S (slower)	S* (slower)	S (slower)
Mismatched	CMT-Read	Post	N	N	N
		Delayed	S/M (<)	N	N
	CMT-Listen	Post	N	N	N
		Delayed	M (<)	N	S (>)
	SPR-Critical	Post	S (slower)	N	N
		Delayed	M (slower)	N	N
	SPR-Whole sentence	Post	N	N	N
		Delayed	N	N	N
ONGOING					
Matched	CMT-Read	Post	N	N	S (>)
		Delayed	N	N	M/L (>)
	CMT-Listen	Post	N	N	N
		Delayed	S (<)	N	S* (>)
	SPR critical word	Post	L (slower)	N	N
		Delayed	M (slower)	N	N
	SPR whole sentence	Post	M (slower)	N	S* (slower)
		Delayed	S (slower)	N	N
Mismatched	CMT-Read	Post	N	N	N
		Delayed	N	N	N
	CMT-Listen	Post	N	N	S* (>)
		Delayed	S (<)	N	M ((>)
	SPR critical word	Post	S (slower)	N	N
		Delayed	S (slower)	N	N
	SPR whole sentence	Post	S (slower)	N	N
		Delayed	S* (slower)	N	S (slower)

Note. N = Negligible: $d < .60$, and CI passes through zero; S = Small: $d \geq .60 < 1.00$; S* = Small $d \geq .60 < 1.00$, but ES passes through zero; M = Medium: $d \geq 1.00 < 1.40$; L = Large: $d \geq 1.40$; (>) and (<) = ES direction of change in CMT, e.g. (>) = more accurate performance for L2+L1prac than comparison group; (faster) and (slower) = ES direction of change in SPR, e.g. (faster) = faster RT for L2+L1prac than comparison group.

Table 4

Ongoing contexts for L2+L1prac group: Accuracy (CMTs) and reaction time (SPR) results

		CMT-Read	CMT-Listen	SPR critical word	SPR whole sentence
		Accuracy (max=5)	Accuracy (max=5)	RT (ms)	RT (ms)
MATCH					
Pre	M	3.86	3.75	578.76	28953.38
	(SD)	(.46)	(.89)	(222.77)	(11036.86)
Post	M	4.67	4.25	500.09	25155.75
	(SD)	(.52)	(.94)	(158.99)	(9769.59)
Delayed	M	4.77	4.19	529.59	26013.36
	(SD)	(.32)	(1.03)	(226.91)	(14738.98)
MISMATCHED					
Pre	M	2.81	2.70	621.14	27989.50
	(SD)	(.79)	(.76)	(192.61)	(12458.95)
Post	M	3.47	3.32	563.26	24481.22
	(SD)	(1.13)	(1.29)	(265.93)	(7898.55)
Delayed	M	3.59	3.15	581.39	24639.04
	(SD)	(1.01)	(1.27)	(243.64)	(7769.94)

Table 5
Ongoing contexts: ES (Cohen's *d* including CIs for *d*) comparisons with Control, L2+L1prac+EI and L2-only from M&M, and **ES changes with effects adjusted for baseline differences**

	CMT-Read			CMT-Listen			SPR critical word			SPR whole sentence		
	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1prac+EI	L2+L1prac vs. L2-only	L2+L1prac vs. Control
MATCHED												
Pre	.36	-.09	-.03	-.11	-.10	-.07	-.02	-.03	-.03	.52	.23	.40
	[-.31, 1.01]	[-.74, .57]	[-.70, .63]	[-.76, .55]	[-.75, .56]	[-.73, .60]	[-.67, .63]	[-.69, .62]	[-.70, .63]	[-.15, 1.18]	[-.44, .88]	[-.28, 1.06]
Post	-.13	.00	.93	-.04	-.10	.34	1.76	.44	.15	1.25	.16	.61
	[-.78, .53]	[-.65, .65]	[.21, 1.61]	[-.70, .61]	[-.75, .56]	[-.34, 1.00]	[.95, 2.48]	[-.23, 1.09]	[-.52, .82]	[.51, 1.94]	[-.50, .81]	[-.08, 1.27]
Delayed	-.24	.25	1.39	-.77	.05	.61	1.25	.02	.21	.87	.52	.49
	[-.89, .43]	[-.41, .90]	[.63, 2.10]	[-1.43, -.08]	[-.60, .70]	[-.09, 1.27]	[.51, 1.93]	[-.63, .68]	[-.46, .87]	[.17, 1.54]	[-.16, 1.17]	[-.20, 1.15]
Pre-post *d* change	**-.49**	**.09**	**.96**	**.07**	**.00**	**.41**	**1.78**	**.47**	**.18**	**.73**	**-.07**	**.21**
Pre-delayed *d* change	**-.60**	**.34**	**1.42**	**-.66**	**.15**	**.68**	**1.27**	**.05**	**.24**	**.35**	**.29**	**.09**
MISMATCHED												
Pre	-.22	-.22	-.40	-.15	-.13	-.13	.01	-.02	.04	.49	-.07	.40
	[-.87, .44]	[-.87, .44]	[-1.06, .28]	[-.80, .51]	[-.79, .53]	[-.79, .54]	[-.64, .67]	[-.67, .64]	[-.63, .70]	[-.18, 1.14]	[-.72, .59]	[-.28, 1.06]
Post	-.27	.26	-.01	.07	.03	.67	.80	-.06	-.06	.95	.34	.42
	[-.92, .39]	[-.40, .91]	[-.67, .66]	[-.59, .72]	[-.63, .68]	[-.03, 1.34]	[.10, 1.46]	[-.71, .60]	[-.72, .61]	[.24, 1.61]	[-.32, .99]	[-.26, 1.08]
Delayed	-.58	.31	.19	-.73	.17	.91	.91	.03	-.20	.61	.19	.78
	[-1.23, .10]	[-.36, .96]	[-.48, .85]	[-1.39, -.04]	[-.49, .82]	[.19, 1.59]	[.20, 1.58]	[-.63, .68]	[-.86. .47]	[-.08, 1.36]	[-.47, .84]	[.07, 1.45]
Pre-post *d* change	**-.05**	**.48**	**.39**	**.22**	**.16**	**.80**	**.79**	**-.04**	**-.10**	**.46**	**.41**	**.02**
Pre-delayed *d* change	**-.36**	**.53**	**.59**	**-.58**	**.30**	**1.04**	**.90**	**.05**	**-.24**	**.12**	**.26**	**.38**

REFERENCES

Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing: Evidence from differential object marking acquisition. *Studies in Second Language Acquisition, 37*(2), 237-268.

Author. (XXXX).

Cumming, G. (2016). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. London: Routledge.

DeKeyser, R. (2007). Situating the concept of practice. In R. DeKeyser (Ed.), *Practicing in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 1-18). Cambridge: Cambridge University Press.

DeKeyser, R., & Prieto Botana, G. (2015). The effectiveness of Processing Instruction in L2 grammar acquisition: A narrative review. *Applied Linguistics, 36*, 290-305.

Ellis, N. C. (2006). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking and perceptual saliency. *Applied Linguistics, 27*(2), 1-31.

Field, A. (2013). D*iscovering statistics using IBM SPSS Statistics*. London: Sage.

Fernández, C. (2008). Reexamining the role of explicit information in Processing Instruction. *Studies in Second Language Acquisition, 30*, 277-305.

González, P. (2008). Towards efective instruction on aspect in L2 Spanish. *International Review of Applied Linguistics in Language Teaching, 46*(2), 91-112.

Henry, N. (2015). *Morphosyntactic processing, cue interaction, and the effects of instruction: An investigation of processing instruction and the acquisition of case markings in L2 German.* (Unpublished doctoral dissertation), The Pennsylvania State University.

Henry, N., Culman, H., & VanPatten, B. (2009). More on the effects of explicit information in instructed SLA. A partial replication and a response to Fernández (2008). *Studies in Second Language Acquisition, 31*, 559-575.

Horst, M., White, J., & Bell, P. (2010). First and second language knowledge in the language classroom. *International Journal of Bilingualism, 14*, 331-349.

Izquierdo, J., & Collins, L. (2008). The facilitative role of L1 influence in tense-aspect marking: A comparison of Hispanophone and Anglophone learners of French. *Modern Language Journal, 92*, 350-368.

MacWhinney, B. (2005). The emergence of linguistic form in time. *Connection Science, 17*, 191-211.

McManus, K., & Marsden, E. (2016, FirstView). L1 explicit information can improve L2 online and offline performance. *Studies in Second Language Acquisition*, 1-34.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 65*(4), 878-912.

Runnqvist, E., Gollan, T., Costa, A., & Ferreira, V. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, 129, 256-263.

Sanoudaki, E. & Thierry, G. (2014). Juggling two grammars. In E. M. Thomas & I. Mennen (Eds.), *Advances in the study of bilingualism* (pp. 214-230). Bristol: Multilingual Matters.

Sanz, C., & Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 24*, 35-78.

Spada, N., Lightbown, P. M., & White, J. (2005). The importance of form/meaning mappings in explicit form-focused instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in Instrcuted Second Language Acquisition* (pp. 199-234). Berlin: Mouton de Gruyter.

Stafford, C. A., Bowden, H. W., & Sanz, C. (2012). Optimizing language instruction: Matters of explicitness, practice, and cue learning. *Language Learning, 62*(3), 741-768.

VanPatten, B. (2002). Processing Instruction: An update. *Language Learning, 52*, 755-803.

VanPatten, B., & Borst, S. (2012). The role of explicit information and grammatical sensitivity in Processing Instruction: Nominative-Accusative case marking and word order in German L2 *Foreign Language Annals, 45*(1), 495-510.

VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

Supplementary materials

Table S1

Number of instances of EI received during training when selecting incorrect answers by group

	L2+L1prac+EI (n=17)	L2-only (n=17)	L2+L1prac (n=19)
Session 1: Ongoingness, past vs. present			
After incorrectly responding ‘MAINTENANT’	9	5	13
After incorrectly responding ‘DANS LE PASSÉ’	8	10	10
Session 2: Habituality, past vs. present			
After incorrectly responding ‘MAINTENANT’	23	21	29
After incorrectly responding ‘DANS LE PASSÉ’	7	9	7
Session 3: Ongoing vs. Habitual, past only			
After incorrectly responding ‘ONGOING / INTERRUPTED’	60	63	57
After incorrectly responding ‘REGULARLY REPEATED'	25	28	26
Session 4: Ongoing vs. Habitual vs. Complete, past only			
After incorrectly responding ‘ONGOING / INTERRUPTED’	111	111	108
After incorrectly responding ‘REGULARLY REPEATED'	71	77	68
After incorrectly responding ‘COMPLETE':	87	93	83