Studies in Second Language Acquisition **40** (2018), 459–475
doi:10.1017/S0272263117000171

Replication Study

ONLINE AND OFFLINE EFFECTS OF L1 PRACTICE IN L2 GRAMMAR LEARNING

A PARTIAL REPLICATION

Kevin McManus*

Pennsylvania State University

Emma Marsden

University of York

Abstract

This study partially replicates McManus and Marsden (2017), 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 role of the L1 practice. This design tested whether providing L1 task-essential practice only (alongside a core treatment of L2 EI plus L2 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). For accuracy and speed of online and offline L2 processing, the findings suggest that additional L1 practice without L1 EI was no more beneficial than L2 EI plus L2 practice alone, indicating that the original study's combination of additional L1 EI with L1 practice appeared to contribute to previously observed learning benefits.

Understanding how explicit information (EI) and practice shape second language development remains an active goal of instructed SLA research. Evidence indicates that,

The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at https://www.iris-database.org/iris/app/home/detail?id=york:932243.

This research was supported by a British Academy Postdoctoral Fellowship (PF130001) awarded to Kevin McManus, with Emma Marsden as Mentor. Many thanks to the learners who participated in this study, and to Angela O'Flaherty and Nigel Armstrong for help with recruitment. We would also like to thank the editors and two anonymous *SSLA* reviewers for their valuable comments, and Oriane Boulay, Margaret Coleman, and Sylvie Goodlow for their research assistance.

*Correspondence concerning this article should be addressed to Kevin McManus, Center for Language Acquisition, Department of Applied Linguistics, 209 Sparks Building, The Pennsylvania State University, University Park, Pennsylvania 16802. E-mail: kmcmanus@psu.edu

in some cases, EI benefitted learning very little (Sanz & Morgan-Short, 2004; Stafford, Bowden, & Sanz, 2012; VanPatten & Oikkenon, 1996). In other cases, EI accelerates 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; VanPatten, Collopy, Price, Borst, & Qualin, 2013). 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 crosslinguistically problematic target features (Spada, Lightbown, & White, 2005).

To this end, McManus and Marsden (2017) examined whether providing explicit 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 form-meaning mapping differences (Howard, 2005; Izquierdo & Collins, 2008; McManus, 2013, 2015).[1] In a 12-week Pretest, Posttest, Delayed Posttest design, 50 English-speaking learners of L2 French were assigned to 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 L2 task-essential practice. The L2+L1 group received the exact same core L2 treatment (as received by L2-only) as well as EI about the L1 and L1 task-essential practice. EI was delivered (a) prepractice (approximately five minutes) and (b) during the task-essential practice following incorrect answers. The prepractice EI used video, image, and sound to depict conceptual-semantic information (e.g., an ongoing event of a man eating an apple), followed by aural and written forms and information about how to interpret their meaning(s), that is, information about form-meaning connections. EI and practice were delivered using laptops (for a full description of the instruction, see McManus and Marsden, 2017, and IRIS for materials). Pretest, Posttest, and Delayed Posttest measures were online (self-paced reading, SPR) and offline (context-sentence matching tests, CMT, in listening and reading), measuring sensitivity to the *Imparfait*'s ongoing and habitual meanings, in which a French stimulus either matched or mismatched the ongoing/habitual meaning of a preceding event description in English (see [3] for example stimuli). Results showed no significant improvement over time for the Control group in speed and 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 (RTs) in mismatched trials for only the L2+L1 group.

While some previous research has provided EI about the L1 (but without L1 practice), with mixed learning outcomes (González, 2008; Horst, White, & Bell, 2010; Spada et al., 2005), McManus and Marsden (2017) is understood to be the first to have employed L1 task-essential practice. It is possible that L1 task-essential practice contributed to the online and offline benefits observed because 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. Similarly to McManus and Marsden (2017), our intervention provided L2 EI plus L2 practice. It additionally included L1 task-essential practice, like McManus and Marsden's (2017) L2+L1 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 EI plus L1 task-essential practice.

MOTIVATION FOR THE PRESENT STUDY

A small number of studies have provided EI about both L2 and L1, finding some offline benefits 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 cloze test results showed that L2+L1 EI was more beneficial than L2-only EI. However, interpretation of the results is inconclusive because the test (a written composition, identical at Pretest and Posttest) did not test knowledge types other than those taught during the intervention (largely controlled, offline, explicit) without Delayed Posttests. Spada et al.'s (2005) L1-L2 contrastive EI explicitly directed learners' attention to L1-L2 differences and similarities. The 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, and no L1 practice was provided. In sum, previous research providing L1 EI has indicated potential advantages. Thus, the extent to which the learning gains observed by McManus and Marsden (2017) were due to providing L1 EI, L1 practice, or their combination remains unclear.

An additional motivation for the current study relates to the nature of the outcome measures to date. (Although see Andringa & Curcic, 2015; Dracos, 2013; Henry, 2015; Marsden, Williams, & Lui, 2013, using eye-tracking, SPR, and cross-modal priming respectively; but these studies did not incorporate L1 information or practice.) That is, research into L1 EI has not measured instructional effects on online processing, which is particularly relevant given that many interventions have aimed to influence input processing.

In terms of potential learning mechanisms at play, McManus and Marsden (2017) argued that the L1 EI may have helped establish explicit knowledge about L1 form-meaning mappings, specifically that in English perfectivity and habituality can be expressed by the same morpheme (-ed), but ongoingness is expressed by a different morpheme (-ing). However, French form-meaning mappings are different: Habituality and ongoingness can be expressed by the same morpheme (-ait), but perfectivity is

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

expressed by a different morpheme. As the task-essential practice required learners to make explicit, meaning-based judgements about stimuli, the practice was claimed to have helped proceduralize explicit knowledge about L1 and L2 form-meaning mappings (DeKeyser, 2015). Considering evidence of L1 activation during L2 processing and production (see Runnqvist, Gollan, Costa, & Ferreira, 2013; Sanoudaki & Thierry, 2014), this knowledge may have helped the L2+L1 group parse and/or anticipate cues needed to disambiguate the *Imparfait*'s habitual from ongoing meanings more accurately and faster than the L2-only group. Given the preceding, it is unclear whether EI and practice play different roles. For example, L1 practice alone may have been sufficient to establish such knowledge, 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 examine 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 L1 task-essential practice alone could result in similar learning gains as L1 EI plus L1 practice, that is, results would pattern with McManus and Marsden's (2017) 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, 2015; Marsden & Chen, 2011; Sanz & Morgan-Short, 2004), that is, this account assumes that L1 EI can benefit L2 learning; or (b) L1 EI does not, in fact, affect or interact with L2 knowledge/processing mechanisms, but practice in interpreting the L1 does.

Our hypothesis that L1 EI could interact with accruing L2 knowledge is compatible with views that foreground a role for L1 in L2 processing (Ellis, 2006; O'Grady, 2013), including evidence of L1 influence in L2 sentence processing (Roberts & Liszka, 2013; Tokowicz & Warren, 2010) and convergence of L1 and L2 systems (Brown, 2015; Brown & Gullberg, 2013). We additionally assume roles for L2 EI plus L2 practice in L2 learning (DeKeyser, 2015). Given that the preceding may be particularly relevant for features with crosslinguistic differences in form-meaning mappings and processing routines, like the target feature investigated here, it is possible that the L1 EI in McManus and Marsden (2017) could have directed learners' attention to L1 form-meaning mappings and L1-entrenched processing routines to improve the effectiveness of the L1 practice. L1 practice, in turn, could promote remapping of L2 form-meanings and more accurate and efficient L2 processing during the L2 EI and practice. In short:

1. Results that pattern like those from McManus and Marsden's (2017) 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 for offline performance (Marsden & Chen, 2011; Sanz and Morgan-Short, 2004) or beneficial for online processing (Andringa and Curcic, 2015).
2. Results that pattern more closely to those from McManus and Marsden's (2017) 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 or the

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

combination of L1 EI plus L1 practice. This would corroborate findings by González (2008), Horst et al. (2010), and Spada et al. (2005) that L1 EI 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 bachelor of arts honors degree in French, recruited from two large universities in England. Every participant was an English native speaker, had completed A-level French (a British, subject-based, secondary school leaving qualification, equivalent to level B2 in the Common European Framework of Reference for Languages), 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. These are very similar participant characteristics to the participants in McManus and Marsden (2017).

DESIGN

The study design, timescale, procedure, and outcome measures were exactly as reported by McManus and Marsden (2017). 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, totaling 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 practice in interpreting equivalent English forms for ongoingness and habituality. These treatment components are briefly described in the following text (see McManus and Marsden [2017] for descriptions and examples and IRIS for full materials, www.iris-database.com).

EI about the L2 was provided prepractice (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 the treatment groups in McManus and Marsden (2017) (see Table S3, Supplementary Materials). Identical to McManus and Marsden (2017), the prepractice EI depicted conceptual-semantic information using 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 prepractice or during practice.

Immediately following the L2 prepractice EI, participants undertook forced-choice, task-essential practice that focused attention on meaning contrasts expressed by different verb forms. Images plus bracketed infinitives pinpointed the verb form to which participants should attend. The stimulus's meaning was selected from two fixed options in training sessions 1–3

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

(as in, e.g., VanPatten, 2002) and three in session 4. Practice lasted approximately 40–50 minutes per session. L1 (k = 112) and L2 (k = 384) practice items were interweaved to promote attention to L1 and L2 form-meaning mappings. Sample sentences are shown in (a) for past ongoingness and (b) for past habituality. The French items drew on 48 lexical verb types, each occurring 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

All measures were identical to those used in McManus and Marsden (2017).

Context-Matching Tests (Listening and Reading)

All participants completed two context-matching tests, administered with EPrime (Schneider, Eschman, & Zuccolotto, 2012): first listening (CMT-Listen), then reading (CMT-Read), each with 24 target and 8 filler trials. Each trial consisted of (a) a written English context: two sentences describing either a habitual or an ongoing activity and (b) a French stimulus: a two-clause French sentence that either matched (k = 12) or mismatched (k = 12) the meaning of the English context. Sample context-stimulus parings are shown in (3).

(3) Context (ongoing): Last Thursday Peter drove to the supermarket to buy some food for tomorrow's party. On his way there he spotted his mother who was walking to work.

Matched stimulus: Quand Peter conduisait au supermarché, il a vu sa mère
"When Peter was driving to the supermarket, he saw his mother"
Mismatched stimulus: Quand Peter conduisait au supermarché, il voyait sa mère
"When(ever) Peter drove/used to drive to the supermarket, he saw his mother"

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"), and 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. Answers could not be changed. The task was untimed and took approximately 20–25 minutes.

Self-Paced Reading Test

The SPR test was administered using EPrime after the CMTs and used 16 items from the CMT-Listen, with eight context-stimulus matches and eight mismatches. Half the items

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

were followed by yes/no comprehension questions, which did not focus on the target feature or the context-stimulus match/mismatch, but increased the likelihood that participants focused on meaning (see Keating & Jegerski, 2015). For each trial, the English context appeared for 10 seconds before an X appeared in the center of the screen. A spacebar press brought up the first and then each subsequent word of the French stimuli in the center of the screen. After the last word, the next screen displayed "END." Reaction times were collected from each word.

DATA SCORING AND ANALYSIS

CMT responses were coded as follows: five points per correct response (i.e., one or two for matched, and four or five for mismatched trials); three points per midway responses (three for matched and mismatched trials); and one point per incorrect response (four or five for matched, and one or two for mismatched trials). Cronbach's alphas were CMT-Listen version A ($\alpha = .81$), version B ($\alpha = .86$), CMT-Read version A ($\alpha = .74$), and 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. RTs 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 analyzed the raw RT data, removing critical word RTs less than 150 ms and greater than 2000 ms.

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 Pretest, Posttest, and Delayed Posttest scores. If a statistically 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 McManus and Marsden (2017) (L2+L1, L2-only, and Control). Having established baseline parity, we then compared the current study's results with the patterns of results found by McManus and Marsden (2017) 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. These analyses mirrored that of McManus and Marsden (2017).

Following McManus and Marsden (2017) and recent discussion on decreasing the probability of Type II errors for low-stakes outcomes and as p-values can be strongly influenced by sample size (Plonsky & Oswald, 2014), the alpha was set at 0.10 for the Kruskall-Wallis H and Friedman tests. This alpha-level choice also reflects our small sample size and the study's exploratory nature. 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 *d*. CIs for *d* that do not pass through zero can be considered reliable indicators of change. Within-group ES were calculated in relation to the mean and standard deviation of the Pretest as a baseline (and Posttest for ES at Delayed). As within-group ES tend to be larger than between-group, we also present, in the Supplementary Materials, within-group ES corrected for the

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

dependence (correlation) between the means (Morris & DeShon, 2002, Table S2). Between-group ES adjusted for Pretest differences are additionally provided in the Supplementary Materials. Between-group ES are provided for each of McManus and Marsden's (2017) groups using the mean and standard deviation of the relevant group from McManus and Marsden (2017) as the "comparison/control" group. We draw on Plonsky and Oswald's (2014) field-specific benchmarks: within subject: $d \geq .60 < 1.00$ (small), $\geq 1.00 < 1.40$ (medium), and ≥ 1.40 (large); between subject: $d \geq .40 < .70$ (small), $\geq .70 < 1.00$ (medium), and ≥ 1.00 (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.

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 McManus and Marsden (2017) are presented in Table 2.

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, $X^2(3) = .29$, $p = .96$) (for between-groups ES and CIs for *d* at baseline see Supplementary Materials, Table S1. All CIs for *d* passed through zero).

MISMATCHED TRIALS

Friedman tests showed no statistically significant improvement over time on any measure with only negligible ES. CIs for *d* 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])

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

TABLE 1. Habitual contexts for L2+L1prac group: Accuracy (CMTs) and reaction time (SPR) results

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

- 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 McManus and Marsden (2017) are presented in Table 2.

Baseline parity was found between the four groups on all measures (CMT-Read, $X^2(3) = .58$, $p = .90$; CMT-Listen, $X^2(3) = .82$, $p = .85$; SPR critical word, $X^2(3) = .63$, $p = .73$; SPR whole sentence, $X^2(3) = 1.29$, $p = .52$).

RTs IN MATCHED VERSUS MISMATCHED TRIALS

Wilcoxon Signed Rank tests showed no statistically significant between-trial (matched *vs.* 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, McManus and Marsden's (2017) L2+L1 group had RTs that were significantly slower in mismatched compared to matched trials at both Post and Delayed (medium ES). Between-trial differences for whole sentence processing at Post and Delayed were negligible, 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 3. All participants scored 100% on the SPR comprehension questions.

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

TABLE 2. Habitual contexts: ES (Cohen's *d* including CIs for *d*) comparisons with Control, L2+L1, and L2-only from McManus and Marsden (2017), and *ES changes with effects adjusted for baseline differences*

	CMT-Read			CMT-Listen			SPR critical word			SPR whole sentence		
	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control
MATCHED												
Pretest	−.09 [−.74, .57]	−.07 [−.72, .59]	−.08 [−.75, .59]	.10 [−.56, .75]	.17 [−.49, .82]	.05 [−.61, .72]	.01 [−.64, .67]	−.04 [−.69, .62]	−.04 [−.71, .62]	.11 [−.54, .77]	.01 [−.65, .66]	.19 [−,49, .85]
Posttest	−.22 [−.87, .44]	.25 [−.41, .90]	.34 [−.34, 1.00]	−.02 [−.67, .64]	−.11 [−.77, .54]	.47 [−.22, 1.13]	1.18 [.45, 1.86]	.07 [−.58, .73]	.02 [−.65, .68]	.77 [.08, 1.43]	.18 [−.48, .83]	.37 [−.31, 1.03]
Delayed Posttest	−.77 [−1.43, −.08]	−.24 [−.89, .42]	.65 [−.04, 1.32]	−.96 [−1.63, −.25]	−.29 [−.94, .38]	.19 [−.48, .86]	1.50 [.73, 2.20]	.59 [−.09, 1.25]	.21 [−.46, .87]	.92 [.21, 1.58]	.62 [−.06, 1.28]	.70 [.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												
Pretest	−.17 [−.82, .49]	−.03 [−.69, .62]	−.05 [−.72, .61]	−.21 [−.86, .45]	−.10 [−.75, .55]	−.03 [−.70, .63]	−.02 [−.67, .64]	.02 [−.63, .67]	.02 [−.64, .69]	.10 [−.55, .76]	−.42 [−1.08, .25]	−.05 [−.72, .61]
Posttest	−.25 [−.90, .41]	−.05 [−.70, .61]	.10 [−.56, .77]	−.50 [−1.16, .17]	.13 [−.53, .78]	.27 [−.40, .94]	.87 [.17, 1.53]	.14 [−.52, .79]	.17 [−.50, .83]	.49 [−.19, 1.14]	.16 [−.50, .81]	.11 [−.55, .78]
Delayed Posttest	−.95 [−1.62, −.24]	.09 [−.57, .74]	.16 [−.51, .83]	−1.04 [−1.71, −.32]	−.07 [−.72, .59]	.84 [.13, 1.51]	1.00 [.28, 1.67]	.14 [−.52, .80]	.14 [−.53, .80]	.28 [−.38, .93]	.13 [−.53, .78]	.38 [−.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**

Note. Positive ES in SPR indicate that L2+L1prac was *slower* than the other group. Negative ES in CMTs indicate that L2+L1 was *less accurate* than the other group.

TABLE 3. Ongoing contexts for L2+L1prac group: Accuracy (CMTs) and reaction time (SPR) results

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

MATCHED TRIALS

There was statistically significant change over time in 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 4 shows comparisons with the groups from McManus and Marsden (2017).

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, $X^2(3) = 1.02$, $p = .60$) (for between-groups ES and CIs for *d* at baseline see Supplementary Materials, Table S1. All CIs for *d* passed through zero).

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 = -1.83, $p = .07$, *d* [CI] = .68 [.01, 1.32]; Pre-Delayed, Z = -2.22, $p = .027$, *d* [CI] = .86 [.18, 1.51]; Post-Delayed, Z = -.63, $p = .53$, *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])

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

TABLE 4. Ongoing contexts: ES (Cohen's *d* including CIs for *d*) comparisons with Control, L2+L1prac+EI, and L2-only from McManus and Marsden (2017), and *ES changes with effects adjusted for baseline differences*

	CMT-Read			CMT-Listen			SPR critical word			SPR whole sentence		
	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control	L2+L1prac vs. L2+L1	L2+L1prac vs. L2-only	L2+L1prac vs. Control
MATCHED												
Pretest	.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]
Posttest	−.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 Posttest	−.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												
Pretest	−.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]
Posttest	−.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 Posttest	−.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**

Note. Positive ES in SPR indicate that L2+L1prac was *slower* than the other group. Negative ES in CMTs indicate that L2+L1 was *less accurate* than the other group.

- 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 McManus and Marsden (2017) are given in Table 4.

Baseline parity between all groups was found in all measures (CMT-Read, $X^2(3) = 2.26$, $p = .52$; CMT-Listen, $X^2(3) = .31$, $p = .96$; 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. ES were negligible ES and CIs for d passed through zero:

- SPR critical word
  - Pre: Z = -.81, $p = .42$, d [CI] = -.20 [-.84, .44]
  - Post: Z = -.77, $p = .44$, d [CI] = -.29 [-.92, .36]
  - Delayed: Z = -.52, $p = .60$, d [CI] = -.22 [-.85, .42]
- SPR whole sentence
  - Pre: Z = -.12, $p = .90$, d [CI] = .08 [-.56, .72]
  - Post: Z = -.80, $p = .94$, d [CI] = .08 [-.56, .71]
  - Delayed: Z = -.32, $p = .75$, d [CI] = .12 [-.52, .75]

In contrast, McManus and Marsden (2017) found small ES for between-trial differences at Post and Delayed in L2+L1 for critical word and whole sentence processing.

SUMMARY OF FINDINGS

McManus and Marsden (2017) found increased accuracy and speed of interpretation of the *Imparfait* at Delayed following a treatment of L1 EI plus L1 task-essential 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 L2 practice). We used the original study's design, procedures, and materials.

Our L2+L1prac group's results patterned very similarly to McManus and Marsden's (2017) L2-only group and tended not to pattern as well with the L2+L1 group (see Tables 2 and 4).

In habitual contexts, change over time was negligible (all measures). In ongoing contexts, change over time in CMT-Listen and SPR was also negligible (all CIs for d passed through zero), but we did find CMT-Read improvement for Pre-Post and Pre-Delayed, matched (large ES) and mismatched (small ES). Comparisons with grammar instruction meta-analyses showed that the gains in ongoing matched contexts (Pre-Post $d = 1.65$, Pre-Delayed 2.30) fell within the broad ranges of mean ES reported in Shintani (2015) ($d = 1.71 - 2.65$) and exceeded those reported in Spada and Tomita (2010) (d range = .73 - .88).[2] However, gains in CMT-Read *mismatched* contexts were much

https://doi.org/10.1017/S0272263117000171 Published online by Cambridge University Press

smaller than in Shintani and fell slightly below Spada and Tomita's ranges at Pre-Post ($d = .68$), but within at Pre-Delayed ($d = .86$). These findings pattern with McManus and Marsden's (2017) L2-only group (limited to a few offline measures, specifically in CMT-Read, Pre-Delayed, Ongoing matched, and 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 McManus and Marsden's (2017) L2-only group.

These results suggest that L1 practice was not as beneficial as L1 EI plus L1 practice for learners' online processing and offline interpretation of the *Imparfait*. Further, the L1 practice without L1 EI seemed to contribute little extra compared to the L2-only intervention (negligible ES throughout).

DISCUSSION AND CONCLUSIONS

Whereas the benefits of L2 EI and L2 practice are well researched to date, the current study addressed the role of L1 practice in L2 learning. Classroom-based evidence has suggested benefits of L1 EI (González, 2008; Spada et al., 2005), but has not examined L1 practice in L2 learning. Although that research had different designs to McManus and Marsden (2017) and this replication, our findings broadly align with it, extending it to show L1 EI plus practice benefited L2 offline and online performance more than L1 practice alone. In short, L1 practice without L1 EI provided few learning benefits, suggesting a role for L1 EI for learning the L2 features investigated here that have crosslinguistic differences for L1 English learners. 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 McManus and Marsden's (2017) arguments that the combination of L1 EI plus L1 practice helped establish knowledge about L1 form-meaning mappings, which aided L2 form-meaning re-mapping, and, possibly, their proceduralization. The prepractice EI in McManus and Marsden (2017) may have raised awareness about L1 processing routines, including L1-entrenched attention and competing cues (Ellis, 2006). 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-meaning relations, and perhaps also prior instruction about the function of "-ed" (in L2 or L1 classes). The EI may have enabled learners to establish more efficiently (with fewer trials) how English "-ed" could map to a subsequent French *Imparfait* sentence.

We also note that this crosslinguistic complexity may explain, at least partly, why the *L2* EI plus task-essential practice did not seem to help learning reliably (only in the CMTs), unlike in previous research using input-based task-essential practice (e.g., DeKeyser & Prieto Botana, 2015; Marsden, 2006; Marsden & Chen, 2011; Shintani, 2015; VanPatten, 2015). The benefits of L2 EI plus practice may not have been detectable by more time-sensitive/speeded tests because of the influence of L1-entrenched processing routines.

As noted by McManus and Marsden (2017), future research should also examine whether the effectiveness of L1 EI plus L1 task-essential practice depends on the feature having crosslinguistically different form-meaning 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 did not isolate the L1 EI and so do not conclude that L1 EI *alone* was solely responsible for the benefits observed in McManus and Marsden (2017). The L2+L1prac's limited development suggests that the combination of L1 EI plus L1 practice led to the L2+L1 group's superior L2 performance in McManus and Marsden (2017), but future research should isolate L1 EI to test this.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0272263117000171

NOTES

[1]French and English form-meaning mapping differences are described in McManus and Marsden (2017) (see also McManus, 2013, 2015). Briefly, past habituality, ongoingness, and perfectivity are expressed in English and French by verbal morphology, but with critical form-meaning mapping differences: First, ongoingness and habituality can be expressed by one morpheme in French, but not in English, as in (1); second, perfectivity and habituality can be expressed by one morpheme in English, but not in French, as in (2).

(1) Elle mangeait[a-ONG/b-HAB] une pomme quand j'ai appelé[a]/quand nous étions petits[b]
   (a) She was[a-ONG] eating an apple
   (b) She ate[b-HAB] an apple
(2) He played[a-HAB/b-PERF] football every Saturday[a]/once last year[b]
   (a) Il jouait[a] au foot
   (b) Il a joué[b] au foot

[2]To the best of our knowledge these meta-analyses used between-subject ES calculations, uncorrected for within-subject correlation. Thus, comparisons with those given in our main text are relevant (rather than the corrected versions in the Supplementary Materials).

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*, 237–268.

Brown, A. (2015). Universal development and L1–L2 convergence in bilingual construal of manner in speech and gesture in Mandarin, Japanese, and English. *The Modern Language Journal*, *99*, 66–82.

Brown, A., & Gullberg, M. (2013). L1–L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition*, *16*, 477–494.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction*. London, UK: Routledge.

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

Dracos, M. (2013). *The effects of form-focused training and working memory on the L2 processing and learning of morphological cues* (Unpublished doctoral dissertation). University Park, PA: Pennsylvania State University.

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

Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition*, *30*, 277–305.

González, P. (2008). Towards effective instruction on aspect in L2 Spanish. *International Review of Applied Linguistics in Language Teaching*, *46*, 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). University Park, PA: 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.

Howard, M. (2005). Les contextes prototypiques et marqués de l'emploi de l'imparfait par l'apprenant du français langue étrangère. In E. Labeau & P. Larrivée (Eds.), *Nouveaux développements de l'imparfait* (pp. 175–197). Amsterdam, The Netherlands, and New York, NY: Rodopi.

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.

Keating, G., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition*, *37*, 1–32.

Marsden, E. (2006). Exploring input processing in the classroom: An experimental comparison of processing instruction and enriched input. *Language Learning*, *56*, 507–566.

Marsden, E., & Chen, H-Y. (2011). The roles of structures input activities in processing instruction and the kinds of knowledge they promote. *Language Learning*, *61*, 1058–1098.

Marsden, E., Williams, J., & Liu, X. (2013). Learning novel morphology: The role of meaning and orientation of attention at initial exposure. *Studies in Second Language Acquisition*, *25*, 619–654.

McManus, K. (2013). Prototypical influence in second language acquisition: What now for the aspect hypothesis? *IRAL: International Review of Applied Linguistics in Teaching*, *51*, 299–322.

McManus, K. (2015). L1–L2 differences in the acquisition of form-meaning pairings: A comparison of English and German learners of French. *Canadian Modern Language Review*, *71*, 51–77.

McManus, K., & Marsden, E. (2017). L1 explicit information can improve L2 online and offline performance. *Studies in Second Language Acquisition*, *39*, 1–34. doi: 10.1017/S027226311600022X

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105–125.

O'Grady, W. (2013). The illusion of language acquisition. *Linguistic Approaching to Bilingualism*, *3*, 253–285.

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

Roberts, L., & Liszka, S. A. (2013). Processing tense/aspect agreement violations online in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, *29*, 413–439.

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (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, UK: 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.

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics*, *36*, 306–325.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, *60*, 263–308.

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 Instructed 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*, 741–768.

Tokowicz, N., & Warren, T. (2010). Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology*, *22*, 1092–1106.

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

VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching*, *53*, 91–109.

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*, 495–510.

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

VanPatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit information, grammatical sensitivity, and the first-noun principle: A cross-linguistic study in processing instruction. *The Modern Language Journal*, *97*, 506–527.