Studies in Second Language Acquisition, **43** (2021), 574–605
doi:10.1017/S0272263120000698

Research Article

EXPLORING SYNTACTIC PRIMING AS A MEASURE OF IMPLICIT LANGUAGE APTITUDE

Shaofeng Li

Florida State University

Jiancheng Qian*

Zhengzhou University

Abstract

This study investigates the validity of syntactic priming as a measure of implicit language aptitude. Syntactic priming refers to the tendency to reproduce a linguistic structure due to a previous exposure to the structure. The validity of the construct was verified by collating evidence for divergent validity—whether it is dissociable from explicit aptitude; convergent validity—whether it is correlated with other measures of implicit aptitude; and predictive validity—whether it is predictive of learning attainment. One hundred sixty-six university EFL learners completed three tests of implicit aptitude: syntactic priming, sequence learning, and LLAMA_D; three tests of explicit aptitude: LLAMA_B, _E, and _F; and three tests of L2 proficiency: untimed grammaticality judgment, metalinguistic knowledge, and elicited imitation. The results showed that syntactic priming was dissociable from explicit aptitude, but it failed to converge with the other measures of implicit aptitude, and it also failed to predict L2 proficiency. The results also showed that priming was negatively correlated with sequence learning and that sequence learning was a negative predictor of learners' metalinguistic knowledge. On the other hand, the construct validity of explicit aptitude was strong. The results suggest the multidimensionality of implicit aptitude and the need for more research into the construct validity of syntactic priming as a cognitive ability for implicit learning.

INTRODUCTION

Syntactic priming, alternatively known as structural priming or structural persistence, refers to one's tendency to reuse a linguistic structure because of a previous encounter with the structure. For example, after hearing the sentence "The professor gave a book to the student," one is more likely to say, "The tour guide showed a map to the woman," even

* Correspondence concerning this article should be addressed to Jiancheng Qian, School of International Studies, Zhengzhou University, 100 Kexue Road, Zhengzhou, Henan 450001, P. R. China. E-mail: qjc2000@sina.com.

though "The tour guide showed the woman a map" is equally acceptable. Since Bock's (1986) seminal study, there has been an exponential growth in the empirical investigation of various aspects of priming. The large body of research has been synthesized in a number of meta-analytic (Mahowald et al., 2016) and narrative (Ferreira & Bock, 2006; Pickering & Ferreira, 2008) reviews. Originating in cognitive psychology as a method of examining the mechanism of speech production, priming was introduced into the field of second language acquisition (SLA) as a tool for examining whether learners incorporate the primed structures in their subsequent utterances in a second language and whether cross-linguistic influence is present in priming (Jackson, 2018). The research in both cognitive psychology and SLA has primarily focused on whether priming is effective in facilitating the production of certain linguistic targets, rather than whether priming can serve as a predictor of learning outcomes. This study examines priming as a measure of implicit language aptitude, an initiative that has not been undertaken in priming or aptitude research and that has the potential of contributing significantly to both fields of inquiry.

Implicit language aptitude, defined as a set of cognitive abilities to learn a second language unconsciously, is a recently emerged concept in SLA. The concept of implicit aptitude is couched in the theoretical and empirical basis for dissociating explicit and implicit learning and the abilities that underlie the two types of learning. According to dual-process theories of human learning, knowledge and skills are acquired in two distinct ways: through explicit learning characterized by effortful and deliberate processing of information and through implicit learning characterized by unconscious computation of the relationships between available materials in the environment (Evans & Frankish, 2009; Reber, 1993). One piece of evidence for the separation between the two types of learning is that cognitive abilities in the explicit domain that typically relate to attributes of "being smart" such as reasoning and associative memory (components of intelligence) are uncorrelated with implicit learning (Kaufman et al., 2010). In a similar vein, in second language research language aptitude ("intelligence" for language learning), which has been traditionally conceptualized as phonetic coding, analytic ability, and rote memory, has been found to be more strongly correlated with learning that occurs under explicit conditions compared with implicit conditions (Li, 2015, 2017). Therefore, there has arisen a need to examine the validity of the construct of implicit language aptitude and its role in second language learning. To date, there has been limited research on implicit aptitude, and existing studies have yielded mixed results for its validity. Syntactic priming is an ideal measure of implicit aptitude because it taps into the ability to derive syntactic abstractions from available linguistic input unconsciously. In what follows, we provide a discussion of the research and nature of implicit language aptitude, followed by an overview of the research on syntactic priming with a view to identifying its effects, nature, and relevance to second language learning and its potential as a measure of implicit aptitude.

IMPLICIT LANGUAGE APTITUDE

A discussion of implicit language aptitude needs to start with language aptitude in general. Language aptitude is a set of cognitive abilities that are predictive of learning rate and ultimate attainment in second language (L2) learning. According to meta-analytic and narrative syntheses of the research (Li, 2016; Skehan, 2015), aptitude is a strong predictor

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

of L2 attainment, with an average correlation of approximately .5 between composite aptitude scores and outcome measures (course grades and scores on standardized proficiency tests). Aptitude's predictive power is strong in comparison with other individual difference variables such as motivation and anxiety, whose average correlations with outcome measures are roughly .3 (Masgoret & Gardner, 2003) and –.3 (Teimouri et al., 2019; Zhang, 2019), respectively. However, the predictive power of aptitude is not uniform across learner groups and learning conditions. Specifically, it has been found to be more strongly correlated with initial than advanced L2 learning (Li, 2015); it is more likely to be drawn upon by adult learners than child language learners (DeKeyser, 2000; Granena & Long, 2012); high-aptitude learners benefit more from inductive instruction while low-aptitude learners more from deductive instruction (Erlam, 2005; Hwu & Sun, 2014). More generally, as with cognitive abilities for general academic learning, traditional language aptitude has been found to be a set of abilities more likely to be implicated in conscious learning (Li, 2015), thus motivating the need to identify and validate abilities for unconscious learning—implicit language aptitude.

A brief account of the history and measurement of implicit aptitude is necessary before synthesizing the research findings. Kaufman et al. (2010) were among the first to investigate implicit learning as a cognitive ability for second language learning. The primary objective of the study, however, was to examine the associations between implicit learning ability and academic achievements, not just language learning. In their study, implicit learning ability was measured through a serial reaction time (SRT) task in which learners responded to a dot that appeared at different locations on a computer screen. The locations where the dot appeared were based on two sequences: a target sequence and a control sequence. The target sequence appeared more frequently (in 85% of the stimuli) than the control sequence (15%), and learners' test performance was calculated as the difference between their mean reaction time for the target and control sequences. Because of the nature of the stimuli used in the SRT, the underlying ability that is tested is called sequence learning. Granena (2012) introduced SRT into the field of second language acquisition as a measure of implicit language aptitude. Granena's study also found that SRT loaded onto the same factor as LLAMA_D, a subtest of the LLAMA aptitude battery that measures the ability for phonological encoding. Since Granena's study, there has been a steady growth of interest in implicit aptitude in SLA research, and SRT and LLAMA_D have become the default measures of implicit aptitude.

The studies on implicit aptitude can be divided into two large categories: naturalistic and classroom, with the former referring to settings where the second language is the primary language of the community (e.g., learning Spanish in Spain), and the latter to settings where the second language is primarily learned in the classroom (e.g., learning Spanish in the United States). The studies can also be divided into correlational and experimental depending on whether the purpose is to explore the relationship between aptitude and L2 proficiency without manipulation of instructional treatments (correlational) or the relationship between aptitude and the effects of different instructional treatments (experimental). Five published studies were conducted in naturalistic settings, and all are correlational (Granena, 2013; Granena & Long, 2013; Suzuki & DeKeyser, 2015, 2017; Yi, 2018). Studies using SRT as a measure of implicit aptitude showed that (a) implicit aptitude correlated with late Chinese-Spanish bilinguals' implicit knowledge measured through a word-monitoring test but the effects were restricted to linguistic

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

structures involving agreement between sentence elements (e.g., subject-verb agreement) rather than structures involving form-meaning mapping (e.g., subjunctive mood) (Granena, 2013); (b) implicit aptitude was not a predictor of the Japanese proficiency of L1 Chinese speakers who stayed in Japan for around four years as a whole cohort (Suzuki & DeKeyser, 2017), but when they were divided into short- and long-residence learners based on whether they lived in Japan for less or more than 30 months, it was a near-significant predictor of long-residence learners' implicit knowledge measured using a word-monitoring test (Suzuki & DeKeyser, 2015); and (c) implicit aptitude was not a predictor of Chinese ESL learners' collocation judgments in a US context (Yi, 2018). Studies using LLAMA_D as a measure of implicit aptitude showed that LLAMA_D scores were correlated with late bilinguals' acquisition of lexis and collocations (Granena & Long, 2013) and early bilinguals' learning of explicit knowledge about structures involving agreement instead of form-meaning mapping (Granena, 2013).

What do we make of the findings of the naturalistic studies? First, it would seem that SRT is more likely to be predictive of the L2 proficiency of the learners who live in the country for a longer period and who therefore have more exposure to the L2. Although Granena (2013) found SRT predictive of late bilinguals' L2 attainment, the learners in the study had lived in the country for more than 10 years. Second, SRT seems more likely to be predictive of implicit knowledge than explicit knowledge, given that all significant results were found for implicit knowledge. Third, LLAMA_D seems important for vocabulary learning and for the learning of explicit grammar knowledge.

Next, we focus on classroom studies, where L2 learning occurs primarily in the classroom through instructed learning. Correlational classroom studies demonstrated that SRT was correlated with high school students' L2 French and German scores (Kaufman et al., 2010) and advanced L2 Spanish learners' reading and listening proficiency (Linck et al., 2013).[1] LLAMA_D was found to predict the fluency of L2 Spanish speech production (Granena, 2019) and EFL learners' pronunciation development (Saito et al., 2019). Furthermore, Granena (2019) found that SRT, labeled as "implicit learning" ability, interacted with a latent factor called "implicit memory" that consisted of LLAMA_D and Available Long-Term Memory (ALTM) such that among learners with high implicit learning ability, implicit memory was predictive of lexical variety in speech production. The findings of the correlational classroom studies seem to suggest that (a) SRT is predictive of advanced proficiency, with the caveat that the proficiency level of the learners in Kaufman et al.'s (2010) study is unclear, and (b) LLAMA_D is facilitative of the phonological aspects of L2 proficiency.

The experimental classroom studies investigated whether SRT was predictive of treatment effects in different learning conditions. Using a pretest-treatment-posttest design, the studies by Granena and Yilmaz (2019) and Yilmaz and Granena (2019) reported that SRT was predictive of the effects of implicit feedback measured using a self-paced reading task and an untimed written grammaticality judgment test in the learning of Spanish gender agreement but not in the learning of object differential marking; SRT did not predict the effects of explicit feedback while phonological short-term memory did. Both Hamrick (2015) and Tagarelli et al. (2016) investigated the relationship between SRT and the learning of novel structures built in the learners' native language—called the semiartificial grammar paradigm. However, the two studies showed very different results. Hamrick's study investigated the associations between SRT and LLAMA_B (a test of rote

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

memory), on one hand, and the incidental learning of three Persian structures translated into English matrix sentences by native English speakers, on the other. During the treatment, learners read each sentence and rated whether it was easy or difficult—the purpose was to focus learners' attention on meaning and encourage incidental learning; they were not informed that there would be a test. The researcher found a significant effect for LLAMA_B at the time of the immediate posttest and for SRT at the time of the delayed posttest. Tagarelli et al. (2016) conducted a similar study with a different semiartificial language: English words with German word order. They investigated three linguistic structures that differed in complexity and two treatment types: incidental and instructed. In both learning conditions, learners were told to read some sentences, judge whether they were semantically plausible, and read them aloud. In the incidental condition, learners were simply told that they would be tested after learning the materials, whereas in the instructed condition, they were taught rules, asked to write two sentences following the rules, and encouraged to search for rules when learning the materials. The only significant result was a strong negative correlation between SRT and the learning of the most complex structure in the incidental conditions. SRT also showed a negative correlation with other outcome measures, although the correlations were nonsignificant. The negative correlations may be due to the explicit nature of the treatments—even in the incidental conditions learners were informed of an upcoming test, which, according to Hulstijn (2005), constitutes explicit instruction. Therefore, it would seem that implicit aptitude may have a negative impact on explicit learning. However, Tagarelli et al. (2016) administered two SRT tasks and used reaction time data from one task but accuracy data from the other. Thus, the results should be interpreted with caution.

To sum up, the findings of experimental classroom research, together with other strands of research, seem to suggest the following about implicit aptitude. First, sequence learning tested through SRT is predictive of the ultimate attainment of naturalistic learners who lived in the country where the second language is spoken for a longer period (Granena, 2013; Granena & Long, 2013; Suzuki & DeKeyser, 2015), but not learners who lived in the country for a shorter period (less than three years) (Suzuki & DeKeyser, 2015; Yi, 2018). It would seem that more exposure to the second language is more likely to result in implicit knowledge, which in turn is more likely to be correlated with implicit aptitude. This finding is also supported by correlational classroom research showing a significant correlation between sequence learning and advanced L2 proficiency (Linck et al., 2014), which is more likely to be implicit than beginning level L2 proficiency. Second, sequence learning is sensitive to the nature of the instruction learners receive in that implicit instruction draws on implicit aptitude while explicit instruction involves explicit aptitude (Granena & Yilmaz, 2019; Hamrick, 2015; Yilmaz & Granena, 2019). It should be pointed out that the learners investigated in the reviewed experimental studies were beginners, suggesting that implicit aptitude has a role to play even at beginning stages of L2 learning, at least under tightly controlled experimental conditions. Third, sequence learning is correlated with implicit knowledge measured by word-monitoring tasks and oral production tasks (Granena, 2013, 2019; Suzuki & DeKeyser, 2015), but the experimental studies also showed that sequence learning is correlated with learners' grammaticality judgment scores (Yilmaz & Granena, 2019), which have been considered to represent explicit knowledge in SLA research (Ellis, 2005). Fourth, implicit memory measured by LLAMA_D is predictive of the learning of collocations and the phonological

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

aspects of the L2 system (Granena & Long, 2013; Yi, 2018). It is also correlated with learners' metalinguistic knowledge (Granena, 2013). These findings suggest that the ability LLAMA_D represents is likely explicit, or at least not entirely implicit. Fourth, sequence learning is predictive of structures involving linear relationships between sentence elements (agreement structures), but not structures involving form-meaning mapping (nonagreement structures). This finding is obtained by both naturalistic and experimental classroom studies by Granena and Yilmaz (Granena, 2013; Yilmaz & Granena, 2019). However, the distinction between agreement and nonagreement structures seems to be partly based on whether the linguistic structure involves meaning, but nearly all structures involve meaning to some extent and the so-called nonagreement structures such as the subjunctive mood also involve agreement between different sentence elements.

The review of the literature shows that to date, SRT and LLAMA_D have been utilized to measure implicit aptitude. However, the research has demonstrated that SRT (a) involves stimuli that are semantically vacuous, (b) concerns only the linear relationship between sentence elements, and (c) is less likely to be predictive of structures involving form-meaning mapping. LLAMA_D has demonstrated characteristics of explicit aptitude in that it is predictive of metalinguistic knowledge and of the lexical aspects of L2 knowledge. There is therefore a need to identify new measures of implicit aptitude that overcome the limitations of existing measures, and priming is an ideal candidate. In the following sections, we introduce the literature on syntactic priming regarding its effects and nature and its potential as a measure of implicit aptitude.

SYNTATIC PRIMING

Priming refers to facilitated performance due to a previous event. Priming has been examined primarily from the perspective of whether it is effective in influencing subjects' linguistic production. The purpose of this study is to investigate priming as a cognitive ability for language learning. In this case, learners' performance in a priming task will serve as an independent variable hypothesized to predict learning outcomes. To use priming tasks as measures of a cognitive ability, it is important to have a thorough understanding of the literature on priming, including whether priming is effective, what factors constrain its effectiveness, and how priming tasks have been designed.[2] The best way to start the literature review is to summarize the findings of the meta-analysis conducted by Mahowald et al. (2016). The meta-analysis aggregated the results of all empirical research on syntactic priming and found a significant effect for priming in the absence of lexical boost (the same verb is used in the prime and the target), with a weighted mean odds ratio of 1.67 (95% CI: 1.63, 1.72). The finding suggests that when there is no lexical overlap between the prime and the target, participants are .67 times more likely to produce the target structure in primed conditions compared with unprimed conditions. The meta-analysts clarified that an odds ratio of 1.67 is equivalent to .28 if converted to Cohen's *d*, which is a small effect based on Cohen's criteria (1988). However, when there is a lexical overlap, the mean odds ratio became 3.26, which means the chances of producing the primed structure are 2.26 times larger under primed conditions than unprimed conditions. This effect is .6 in Cohen's *d*, which constitutes a medium effect.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

In addition to overall effects, the meta-analysts performed moderator analyses to ascertain whether the effects of priming varied as a function of systematic methodological differences between the sampled studies. We would like to focus on several moderators that are key to understanding the major paradigms of priming research and that are relevant to the priming tasks used in our own study. First, in terms of the impact of the language of priming tasks, no significant differences were found between L1 and L2 priming, but lexical boost had a significantly larger effect in L2 priming than in L1 priming. Second, regarding the effect of lags (fillers between the prime and the target), different findings were obtained when the studies were divided into nontreatment and treatment studies. In nontreatment studies where primes and targets were not manipulated such as those following Bock's model (1986), priming effects decreased significantly with the increase of the number of lags, and the insertion of lags also made the influence of lexical boost disappear. Typically in these studies, the participant responds to an experimental trial by repeating a sentence (the prime) containing the linguistic structure to be primed and then describing a picture (the target). The most examined linguistic structures in these studies are the dative construction and transitive verbs in English. The dative construction has two variants or two alternating structures: prepositional object, as in "The professor gave a book to the student," and double object, as in "The professor gave the student a book." Transitive verbs can be used in either the active voice: "The lightning bolt hit the house," or the passive voice: "The house was hit by the lightning bolt." In these studies, the same participants respond to trials for both alternatives of the same structure, and some of the studies involve more than one linguistic structure. The priming tasks used in this study followed the methods in this paradigm.

Regarding the effect of lags, an opposite pattern was obtained from the analysis of the results of the treatment studies conducted by Kaschak and associates (e.g., Kaschak et al., 2011). Unlike the nontreatment studies, these treatment studies demonstrated that the number of lags was a positive predictor of priming effects such that increasing the number of lags led to larger priming effects. These studies are typically interventional, consisting of a pretest, a training session, and a posttest. In the training phase, participants were "biased" toward one variant, such as the prepositional object of the dative construction, by responding to target trials that sought to manipulate their language production (such as by filling out a template that only allows the prepositional object, "The teacher gave a book____."). Also, in these studies participants only receive training on one variant of a linguistic structure. The methodological variation may account for the opposite effects of lags between treatment and nontreatment studies. In studies following Bock's model, the same learners were primed for multiple structures and for more than one variant of the same structure, and the effects of priming may be limited to trials where there was no lag because trials containing the other variant may cancel the longer-term effects of the variant in question. In studies following Kaschak's model, there was no distraction resulting from cross-priming because there was only one structure in the treatment. Lags in the treatment studies increased the space between primes and targets, alleviating participants' processing burden and leading to greater priming effects—a phenomenon called spacing effects.

One important moderator that has figured prominently in the research but that was not investigated in Mahowald et al.'s (2016) meta-analysis is the linguistic target of priming,

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

that is, some structures are more easily primed than others. For example, it has been found that "the datives generally yielded clearer patterns of priming [and larger effects] than the transitives" (Bock & Griffin, 2000, p. 183). Furthermore, within a particular target structure, one variant may show larger effects than the other. For example, in English, the prepositional object has consistently generated greater priming effects than the direct object, and the passive voice showed greater effects than the active voice. In either case, it is the less frequent variant, such as the prepositional object and the passive, that led to more priming gains—a phenomenon called "inverse frequency" (e.g., Bock & Loebell, 1990; Kaschak et al., 2011).

SYNTATIC PRIMING AS IMPLICIT LEARNING

The primary objective of the present study is to examine the validity of priming as an ability for implicit learning, and therefore a fundamental question for such an initiative is whether priming implicates implicit learning. To answer this question, it is necessary to evaluate the truthfulness of two theoretical accounts of the mechanism of priming: the "transient activation account" and the "implicit learning account" (Bock & Griffin, 2000, p. 180). The transient activation account states that priming is a matter of temporary activation of long-term memory, and that its effects are short-lived. Evidence for this view should show that the effects of priming will disappear or decline steeply if primes and targets are separated. However, this did not happen. The studies by Bock and Griffin (2000) and Bock et al. (2007) show that long lags (number of fillers between the target and prime) did not cause substantial decline of effects. Bock and Griffin (2000) investigated the number of lags as an independent variable and found that the effects of priming could last as long as 10 lags after primes were presented. The effects were strongest at lag 2, followed by lags 1, 0, and 10, suggesting that a shorter lag is not equivalent to a stronger effect. One caveat is that Mahowald et al.'s (2016) meta-analysis showed a negative effect of a longer lag on priming effects. However, the counterevidence for the caveat is that the negative effect was restricted to studies where the priming of alternative structures may have cancelled the long-term effects of priming. The studies where participants received training on a single structure showed a positive effect for a larger number of lags, and these studies also showed that the treatment effects could last as long as one week. Also, the moderator—number of lags—was created by the meta-analysts based on the methodological features of the primary research, and most of the studies did not examine lags as an independent variable. Therefore, the finding is synthesis-based, not study-based, and synthesis-based findings are suggestive rather than conclusive (Li et al., 2012).

The "implicit learning account" holds that priming involves implicit learning because it does not require any effortful, conscious information retrieval or processing.[3] Bock and Griffin (2000) stated:

> The relevant kind of learning appears to be implicit or procedural, inasmuch as it does not depend on specific intentions to replicate a sentence's structure in new words, does not require an effort to remember the priming sentence (Bock, 1986), and does not require explicit attention to the form of a priming sentence (Bock et al., 1992). (p. 180)

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

Similarly, Pickering and Ferreira (2008) believed that the learning that happens in syntactic priming is "tacit, incidental, and automatic" (p. 447). One argument that Bock and her colleagues have made based on a series of studies they conducted is that priming involves syntactic abstraction that is unrelated to semantic connections between primes and targets. Thus, in this view it is syntax rather than semantics that is behind the phenomenon of priming, and as the researchers argued, syntax is likely abstract and implicit, while semantics, which relates to episodic memory, is likely explicit. The claim that priming is a matter of syntax rather than semantics rests on the following empirical evidence. Bock (1986) reported that participants were more likely to produce passive sentences following passive primes and that manipulating the animacy (which concerns meaning) of the agent of the prime did not have any impact on whether they used the passive when describing target pictures. Similar findings were obtained by other studies such as Huang et al. (2016), which reported that manipulating the animacy of the recipient had little influence on the production of the dative alternatives.

Another argument for an implicit learning account is that it is syntax rather than lexical repetition that is at work, and lexicon is more likely to be explicit. This claim has been examined in two streams of research. One relates to lexical boost, which refers to the finding that the repetition of the same verb between the prime and the target leads to larger priming effects. Although the effects of lexical boost are likely due to explicit learning, studies where there was no lexical overlap between primes and targets also demonstrated significant priming effects. Also, the effects of lexical boost disappear once lags are inserted between primes and target (Mahowald et al., 2016). The other stream of research concerns the speculation that priming occurs because of participants' repeated exposure to the same particle that induces them to produce the linguistic target. Bock and Loebell (1990) examined this possibility by investigating the priming effects of two sentence structures on the production of prepositional objects. Both structures contain "to" but have different constituent structures. One is the prepositional object, as in "Susan brought a book to Susan," and the other is the infinitive, as in "Susan brought a book to study." The results showed that only primes with prepositional objects were effective in leading to more frequent production of prepositional objects, not primes with infinitives, suggesting that the driving force for priming is the abstract syntactic rule rather than a mere repetition of the particle "to."

One way to establish the validity of a construct is to identify its relationship with competing variables or variables in a different paradigm and with variables in the same paradigm which supposedly tap into the same or similar construct—the former is known as divergent validity and the latter convergent validity. Evidence for the divergent validity of priming as an implicit learning mechanism is provided by Ferreira et al.'s (2008) study that investigated learners with amnesia and that showed a separation between the learning mechanism of priming and that of explicit learning. The study showed that these learners performed as well as normal controls (learners with no learning disability) in syntactic priming, but their performance in the recognition of the meaning of priming sentences was significantly worse than normal controls. The researchers interpreted the findings as suggesting that the learners had intact procedural memory that contributed to their excellent performance in syntactic priming but they suffered from deficits in declarative memory that resulted in their poor recall of past events. Ferreira et al.'s (2008) interpretations were based on the assumption that

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

procedural memory underlies implicit learning whereas declarative memory is responsible for explicit learning.

While priming has been shown to differ from explicit learning—evidence for divergent validity, there has been scant evidence for convergent validity. Kaschak et al. (2011) is perhaps the only psycholinguistic study that investigated the relationship between priming and other implicit learning tasks. In this study, learners were divided into two groups receiving priming training in the two variants of the dative construction: prepositional object and double object. They also completed a Simon task as a measure of implicit learning, where they reproduced colored squares two seconds after they appeared on a computer screen. Unbeknownst to the participants, the locations of the squares were based on a fixed sequence, and the task consisted of a study phase and a test phase including new and old strings. The participants' implicit learning scores were calculated by subtracting their accuracy scores on the new strings from the scores for the old strings. For both the prepositional object and double object, the contribution of implicit learning (performance on the Simon task) to priming effects was near-significant, $p = .07, .06$, respectively, with implicit learning being a negative predictor of the priming of the direct object (the authors did not report the direction of the correlation for the prepositional object). The authors interpreted the overall effect of implicit learning as being nonsignificant, while pointing out that (a) measures of implicit learning are typically poorly correlated or uncorrelated, and (b) the Simon task may have tapped into a different kind of ability from the ability underlying syntactic priming.

To sum up, there is theoretical justification and empirical evidence for the implicit nature of the learning that results from syntactic priming. Priming does not impose any demand for effortful, conscious information retrieval or processing on the participant. Priming involves abstract syntactic processes that are unrelated to semantics and lexicon (Bock & Loebell, 1990; Huang et al., 2016). Priming effects are sustainable and long-lasting (Mahowald et al., 2016). Similar to the findings on other implicit learning tasks, learners with cognitive impairments in explicit memory performed as well as learners with no cognitive disability when completing tasks of syntactic priming (Ferreira et al., 2008; Kaschak et al., 2011). However, more research is needed on the convergent validity of priming—how it is related to other implicit learning tasks—and there has been no research on its predictive validity, that is, whether it can serve as a predictor of learning outcomes. This study will address these issues.

SYNTATIC PIRMING AS IMPLICIT LANGUAGE APTIUTDE

Syntactic priming is an ideal candidate of implicit aptitude because it involves implicit learning, and because it involves form-meaning mapping, a defining feature of language learning. Although priming studies (e.g., Bock & Loebell, 1990; Huang et al., 2016) have shown that manipulating meaning did not have an impact on the effectiveness of priming, there is no denying that for priming to happen, the target must be consistent with the prime in terms of the basic event or discourse structure. For example, for a prime of double object ("The professor gave the student a book") to be successful, the event of the target (picture) must depict an event that has the basic elements of a sentence containing the double object. In other words, priming is verbal and happens in context, and thus, unlike sequence

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

learning, which is nonverbal, priming is not semantically vacuous. Therefore, priming matches the mechanism of the processing of a natural language.

The mechanism of priming fits the currently popular meaning-based approaches to language instruction that emphasize the importance of learning a second language through exposure to input and through using language as a tool to communicate meaning, rather than excessive grammar instruction and language-focused mechanical drills—characteristics of traditional language instruction. Priming has the following attributes that coincide with meaning-based instruction:

- Priming is an unobtrusive technique in that it does not impose a processing demand and yet it exerts some kind of influence by orienting the learner toward the target feature.
- Priming affords positive evidence that the learner can incorporate in their own interlanguage.
- Priming provides a syntactic template that alleviates learners' onus in selecting and retrieving the appropriate sentence-building procedure, thereby freeing up learners' cognitive resources for form-meaning mapping.
- Priming provides a forum where learners have an opportunity to practice retrieved or newly acquired linguistic knowledge in immediate production.

In light of the congruity between the mechanisms of priming and meaning-oriented instruction, priming can serve the dual function of *a learning task* (McDonough, 2006) and *a measure of the cognitive ability* for implicit learning that occurs in meaning-oriented approaches. To date, priming has been investigated for the former purpose, not the latter (but see Woltz (2003) for a discussion of semantic priming as an ability for implicit learning). However, the findings about priming as a learning task accord with current explanations about the role of explicit and implicit learning in major theoretical models of second language acquisition (DeKeyser, 2020; Long, 2015). In L2 research, L2 priming has been found to be more likely to occur for structures learners have more previous knowledge about (McDonough, 2006; McDonough & Fulga, 2015), suggesting that it facilitates learning that happens at more advanced stages—a feature of implicit learning. The research also demonstrates that in L2 priming tasks, the effects of lexical boost (namely, priming is more effective when the prime and the target contain the same key words) were evident among beginners instead of advanced learners (Kim & McDonough, 2008), and that the effects of priming with lexical boost were short-lived while the effects of priming without lexical boost were sustainable (Jackson, 2018; McDonough, 2011). These findings suggest that priming with no lexical boost, which belongs to the implicit paradigm, happens in advanced L2 learning and generates long-term effects, and that priming with lexical boost, which suggests explicit learning, is characteristic of initial learning. The findings corroborate the claims of current SLA theories about explicit and implicit learning, which are more likely to occur at initial and advanced stages, respectively.

THIS STUDY

The primary objective of this study is to determine whether priming is a cognitive ability for implicit learning. To establish the construct validity of priming as a cognitive ability for implicit learning, it is important to collect evidence for three types of validity: convergent validity, divergent validity, and predictive validity. Convergent validity refers

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

to whether priming is correlated with other measures of implicit aptitude, such as LLAMA_D and SRT, which have been used as tests of implicit aptitude in previous research (e.g., Granena, 2013). Divergent validity concerns whether a certain variable is not or less correlated with variables in another domain—one that is hypothesized to be distinct from the domain the investigated variable falls into. In this case, priming as a measure of implicit aptitude should be uncorrelated or less correlated with cognitive abilities for explicit learning, which have been measured using traditional aptitude tests such as MLAT, LLAMA (B, E, F), and PLAB (Li, 2016). Predictive validity pertains to whether priming is predictive of learning outcomes measured using tests of L2 attainment. To examine the three types of validity of the construct of syntactic priming as a device for implicit learning, the following research questions were formulated:

RQ1. What is the relationship between priming and other measures of implicit aptitude?
RQ2. What is the relationship between priming and measures of explicit aptitude?
RQ3. What is the relationship between priming and L2 attainment?

METHODS

PARTICIPANTS

One hundred sixty-six learners of English as a foreign language from a large Chinese university participated in the study. These were English majors in their second and fourth year of study with an average age of 20.71 years. The curriculum consisted of language-focused courses such as reading, listening, speaking, and writing as well as content-based courses such as English literature and linguistics. The language instruction the learners currently and previously (in secondary schools) received were heavily form-oriented, and the use of meaning-oriented tasks was uncommon. Among the learners, only four reported having visited an English-speaking country, and the average length of stay was two months. The two cohorts of learners showed no significant differences in their grammar knowledge represented by their scores on an untimed written grammaticality judgment test, an elicited imitation test, and a metalinguistic knowledge test (see Appendix for the two cohorts' test scores). Therefore, they were combined as one group in all statistical analyses with a view to increasing the sample size, which will potentially enhance the robustness of the findings.

INSTRUMENTS

A total of six tests were utilized to measure learners' implicit and explicit aptitude and their L2 English proficiency. The measures of cognitive aptitude, which served as predictor or independent variables, included syntactic priming, serial reaction time, and the LLAMA aptitude battery, which consists of four subtests (B, D, E, and F). Among the tests, syntactic priming, serial reaction time, and LLAMA_D were hypothesized to measure implicit aptitude, and the three remaining subtests of the LLAMA were intended to measure explicit aptitude. The measures of outcome variables included (a) an untimed written grammaticality judgment test, (b) a metalinguistic knowledge test, and (c) an elicited imitation test (Ellis, 2005). The first two (untimed grammaticality judgment and

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

metalinguistic knowledge) were intended to measure explicit knowledge, which is analyzed, metalinguistic, conscious, and accessible under conditions without time constraint. The elicited imitation test was posited to measure implicit knowledge—knowledge that is tacit, intuitive, unconscious, and accessed in spontaneous production (but see Suzuki & DeKeyser [2015] who argued that elicited imitation measures automated explicit knowledge instead of implicit knowledge). The tests are described in the following sections.

TESTS OF COGNITIVE ABILITIES (PREDICTING VARIABLES)

Syntactic Priming

As pointed out in earlier sections, priming research has centered on whether learning occurs during a priming task. However, in this study priming was examined as an individual difference variable and the focus is on whether it is predictive of learning outcomes. The priming task was developed in the learners' L1 instead of L2 to prevent the impact of their L2 proficiency on their priming performance.[4] The task stimuli targeted two linguistic structures: the dative construction (prepositional object vs. double object) and transitive verbs (active voice vs. passive voice), which have been investigated in many priming studies including those by Bock and associates (e.g., Bock et al., 2007). The two structures, exemplified in Table 1, also exist in Chinese (Cai et al., 2015). Similar to English, Chinese has two variants for the dative construction: prepositional object and double object, with the former being formed by "Subject + Verb + Direct Object + Preposition + Indirect Object," and the latter by "Subject + Verb + Indirect Object + Direct Object." Transitive verbs also have two variants: active voice and passive voice. In active voice, the subject is the performer of the action, whereas in passive voice the subject is the receiver of the action. In Chinese, the passive construction is formed by "Subject (receiver of action) + Bei (passive marker, similar to 'by') + Agent (performer of action) + Verb." It is noteworthy that in the field of priming, there has been little research on Chinese, and

TABLE 1. Examples of the dative construction and transitive verbs

Structure		Examples
Dative	Double Object	jiaoshou jiegei yisheng yi ben shu Professor lend doctor a [classifier] book A professor lent a doctor a book.
	Prepositional Object	jiaoshou jie le yi ben shu gei yisheng Professor lend [perfective] a [classifier] book to doctor A professor lent a book to a doctor.
Transitive	Active	daodan jizhong le feiji Missile hit [perfective] airplane. A missile hit an airplane.
	Passive	feiji bei daodan jizhong le Airplane by missile hit [perfective] An airplane was hit by a missile.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

the limited research (Cai et al., 2011; Huang et al., 2016) has examined only the dative construction, not transitive verbs (active/passive voice).

Following previous studies (Bock, 1986; Huang et al., 2016), the priming task consisted of two phases: a study phase and a test phase. During the study phase, the learners listened to some sentences and viewed some pictures, and they were informed that they needed to memorize the materials and identify them later in the test phase. The materials consisted of 50 items, half of which were sentences and half pictures. The items were presented in a random order, with the constraint that no more than two of the same type occurred consecutively. The study phase aimed primarily to make the learners believe that the whole activity was a memory task where items studied initially would be subsequently tested. The test phase, which is the core component of the priming task, contained 128 items, including 64 target items and 64 distractors. The target items did not appear in the study phase, and all the items presented in the study phase served as distractors in the test phase. Among the 64 target items, 32 concerned the dative construction, and 32 related to transitive verbs. Among the 32 items for each structure, 16 pertained to one variant (direct object or active voice), and 16 to the other variant (prepositional object or passive voice). Each target item consisted of two components: a prime followed by a target. The prime was an audio sentence that contained the target structure, and the learner was told to repeat the sentence and decide whether it appeared in the study phase. The target was a picture that the learner was asked to describe verbally using provided clues (names of people and objects). After the verbal description, the learner was required to judge whether the picture was presented in the study phase. The expectation was that the learner would reproduce the target structure in their verbal description after being exposed to the structure in the first event—the prime. The distracting items followed the same format as the target items, but they were sentences with intransitive verbs (e.g., "wushi zai tiaowu" meaning "The witch was dancing"), which are unrelated to the two linguistic structures. Figure 1 illustrates the steps and components of an item in the priming task.

The researchers created the stimuli by consulting previous priming studies (Cai et al., 2011; Huang et al., 2016) as well as Chinese grammar books (Li & Thompson, 1981), corpora, newspapers, and dictionaries. Each noun and verb used for both the audio and picture stimuli appeared exactly twice in the task. For the picture stimuli for both linguistic structures, half had the agent on the left and half on the right, to prevent the potential influence of the positions of the agent and patient on learners' utterances. For transitive verbs, whether the agent and patient are human or nonhuman was controlled, which led to four configurations: human agent + human patient (e.g., "The pirate is chasing the dancer"); human agent + nonhuman patient ("The father broke the glasses"); nonhuman agent + human patient ("The clock awakened the singer"); and nonhuman agent + nonhuman patient ("The rocket hit the plane"). The job title of the character, the name of the object, and the verb of the action were provided in the picture. Also, dative pictures had a preamble stating the subject and the verb of the sentence to prevent the use of a third, alternative structure, the *ba-* structure. Scoring of the priming task was based on the number of picture description responses that were consistent with the primes, that is, a response was considered as primed if the target description had the same structure as the prime sentence (i.e., a passive response following a passive prime). For each of the four

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

Prime

Step 1 Listen: "mingxing bei dashu bandao le" ("The celebrity was tripped by a tree.")

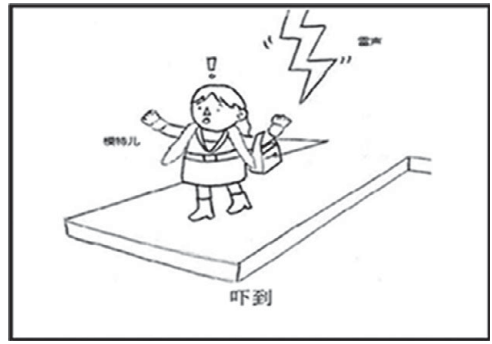
Step 2 Repeat the sentence.

Step 3 Memory Check: "Have you heard the sentence?"

Target

Step 1 Describe the picture.

(Expected description: "mote bei leisheng xiadao le" ["The model was scared by a thunder."])



Step 2 Memory Check: "Have you seen this picture?"

FIGURE 1. The template of a priming trial.

structures—prepositional object, direct object, active, and passive—the maximum score was 16, and the minimum was zero.

Serial Reaction Time Task

A serial reaction time task, which was developed by the researchers, was used to measure sequence learning, a default measure of implicit aptitude in SLA (Granena, 2013; Kaufman et al., 2010; Linck et al., 2013). During the task, the learners responded to a black dot that appeared on four horizontally displayed locations. The locations of the dot were based on two 12-digit number sequences: a target sequence 1–2–1–4–3–2–4–1–3–4–2–3–) and a control sequence (1–2–4–3–1–4–2–1–3–2–3–4–). The target sequence appeared in 85% of the stimuli and the control sequence in 15%. The stimuli were divided into eight blocks, each consisting of 10 12-digit sequences. Each sequence generated an average of 11 triplets (such as 121, 214, 143, etc.), yielding 110 triplets in total for each block. The control sequence was built into the stimuli by varying the third digit of a target triplet (e.g., from 121 to 124), and as much as possible, the control triplets were evenly distributed among the eight blocks of trials, with each block including an average of 16.5

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

control triplets (15% of 110). Altogether, the task comprised of 880 triplets, 748 of which were target triplets and 132 control triplets. The learners were told to press one of the four keys corresponding to the locations on the computer screen as quickly as possible, and they were informed that it was a game where speed and accuracy were both important. The test was scored by subtracting the learners' mean reaction time for the target stimuli from their mean reaction time for the control stimuli: SRT score = control−target (Granena, 2013; Suzuki & DeKeyser, 2015). In this scoring method, a larger value represents higher sequence learning ability because it suggests that the learner was faster in responding to the target stimuli (because he/she had learned the rules) and slower in responding to the control stimuli which were less familiar due to their lower frequency.

LLAMA

The LLAMA aptitude battery, which is modeled on the MLAT (Carroll & Sapon, 2002), is a free online test kit that has been extensively used to measure language aptitude in recent research. Created based on a language spoken in Central America, the test is intended to be language-neutral to learners who are unfamiliar with the language. It has four subtests: B, D, E, and F, each with a study phase and a test phase. LLAMA_B is a test of rote memory, during which the learner has 2 minutes to study 20 picture-word associations and is then tested on the same stimuli. LLAMA_D, which has been claimed to be a measure of implicit aptitude, requires the learner to listen to 10 sound sequences in the study phase and then distinguish old and new sequences in the test phase. LLAMA_E measures the ability to learn sound-symbol associations. During the study phase, the learner listens to the recordings of 24 syllables and studies the scripts of the syllables. In the testing phase, the learner listens to some syllable combinations and chooses the correct spellings for each syllable sequence. LLAMA_F measures the ability for inductive learning or language analytic ability. The learner studies 20 sentences and their meanings represented by pictures for the purpose of extracting the grammar rules underlying the sentences and is then tested on the rules. The score of each subtest is automatically generated by the software program.

TESTS OF L2 PROFICIENCY (CRITERION VARIABLES)

Untimed Grammaticality Judgment Test

Adapted from Ellis (2005), this test was intended to measure learners' explicit grammar knowledge about L2 English. The test required learners to judge whether a given item was grammatical or ungrammatical and correct the error if it was ungrammatical. The test targeted 17 grammatical features such as the subjunctive mood, the simple past, embedded questions, and so on, and for each structure, a grammatical sentence and an ungrammatical sentence were created, yielding a total of 34 items. Students were not given a time limit for the test because the absence of time pressure would encourage the use of explicit knowledge. One point was given if a grammatical item was judged to be grammatical, and if an ungrammatical item was judged to be ungrammatical and the error was corrected.[5]

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

Metalinguistic Knowledge Test

This test measures learners' knowledge about the rules of the 17 grammatical structures that are also tested in the grammatical judgment and elicitation imitation tests. There were 17 items, each consisting of an ungrammatical sentence and four possible descriptions of the error in the sentence, and learners were required to choose the best description. Each item was allocated one point, and the total possible score was 17. As with the untimed grammaticality judgment test, no time limit was imposed for this test. Both the metalinguistic knowledge test and the untimed grammaticality judgment test were posited to measure learners' explicit knowledge.

Elicited Imitation Test

This test was intended to measure learners' implicit knowledge (Ellis, 2005; Ellis et al., 2019). The test included 34 items, half grammatical and half ungrammatical, targeting the same 17 linguistic structures. For each item, the learner was required to listen to the recording of an English sentence (e.g., "I remember meet President Trump at a party"), decide whether it was true, not true, or whether she/he was uncertain, and repeat the sentence in correct English (the correct response should be "I remember meeting President Trump at a party"). The learner was allowed 8 seconds to respond, and the test would proceed to the next item if there was no response within the time limit. A time limit was imposed to limit access to explicit knowledge and encourage implicit knowledge. The recordings of learners' oral production were transcribed verbatim and scored following the rule stipulating that each response receives one point if the target structure is produced in the obligatory context.

PROCEDURE

The study took place in several large computer labs at the site of data collection. The data collection was completed following this sequence: a background questionnaire, the priming task, the serial reaction time task, the LLAMA aptitude tests, the elicited imitation test, the untimed written grammaticality judgment test, and the metalinguistic knowledge. The whole session lasted approximately 2 hours. To prevent the possible influence of the tests of explicit aptitude/knowledge on tests of implicit aptitude/knowledge, the latter always preceded the former. Among the instruments, the priming task, the serial reaction time task, and the elicited imitation test were developed using DMDX (Forster & Forster, 2003)—software used extensively in psychological and SLA research.

ANALYSIS

An initial analysis was conducted on the data on priming to determine whether priming was effective and which of the scores relating to the four primed linguistic targets would be used in subsequent analysis to explore priming's relationships with other variables. After the initial analysis, the data were subjected to structural equation modeling analysis to map the complicated relationships between the observed and latent variables. The paths between the different variables in the tested models were determined based on theoretical

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

hypotheses and findings of previous research. In the case of failure to confirm an initial model, the paths were modified to identify a model that best fit the data (Hiver & Al-Hoorie, 2020). It is worth clarifying that because the research on implicit aptitude is in its infancy, its conceptualization and measurement are highly hypothetical. Therefore, the goal of this study is more exploratory than confirmatory and is both theory-testing and theory-building. As Kline (2016) pointed out, the ideal scenario where a hypothesized model is confirmed is uncommon, and often the goal of a study has to be changed from confirming a model to "discover[ing] a model" (p. 11). One could argue that correlation analyses should be used to probe the relationships between the investigated variables once the initial measurement or structural model is not identified. However, simple correlation analysis can only demonstrate the zero-order relationships between observed variables, not the complicated relationships between observed and latent variables and between latent variables. Also, performing multiple correlations are subject to error, a limitation that structural equation modelling analysis is designed to overcome (Schumacker & Lomax, 2010).

RESULTS

THE RELATIONSHIP BETWEEN SYNTACTIC PRIMING AND OTHER COGNITIVE ABILITIES

Learners' Priming Performance

Before exploring the associations between priming and other variables, several decisions were made about the priming data. First, unlike existing research where learners' responses are often coded as categorical and analyzed as frequency counts, this study treats the data as continuous by calculating a score for each individual learner because the primary objective of the study is to detect individual variation in their potential of being primed. Second, following previous research (Bock & Griffin, 2000; Huang et al., 2016), learners' priming performance was operationalized as their production of the less frequent structures, which were the passive voice and the double object in this case. This practice makes sense if we assume that learners who are more likely to be primed for a less frequently used structure are more able to learn a second language because of their stronger ability to learn less familiar linguistic features and to abandon entrenched habits. Third, to be entered in further analysis to explore its correlations with other cognitive variables and with learning outcomes, priming must be effective, that is, learners' production of a certain structure after receiving primes on that structure must be significantly better than after being primed on the alternative structure. There are two methods to test the effects of priming: primary and secondary. The primary method is to ascertain whether the mean score of primed production of the target structure (e.g., passive) is significantly higher in the primed condition (passive) than the unprimed or alternative condition (active). The secondary method is to see whether learners' production of the target structure under the primed and unprimed condition are highly correlated. If they are highly correlated, then it means learners' production of the structure is consistent across different priming conditions (probably because of personal preference) and is not due to the kind of primes they received. Hence, we are looking for

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

TABLE 2. Learners' priming scores

	Primed production			
	Mean	SD	Mean	SD
Primes	Active		Passive	
Active	11.41	2.27	3.51	1.62
Passive	9.74	2.42	4.37	1.58
Correlation (*r*)	.69* (between primed actives)		.51* (between primed passives)	
	Prepositional object		Direct object	
Prepositional object	11.13	4.18	2.87	3.46
Direct object	11.43	4.50	2.98	3.73
Correlation (*r*)	.90* (between primed PO's)		.88* (between primed DO's)	

*$p < .05$.

a lower correlation between learners' production of the target structure in the primed and unprimed condition.

The learners' priming scores for each structure and priming condition and their scores under primed and alternative conditions are displayed in Table 2. As can be seen, for the active-passive priming, priming was successful, that is, learners were more likely to produce active sentences after active primes than after passive primes, and they produced more passive sentences after passive primes than after active primes. Pair Samples T-tests showed that the differences were significant: $t(134) = 10.41$, $p < .01$ for active primes, and $t(134) = 6.27$, $p < .01$ for passive primes. The correlation between learners' production of active sentences under primed and unprimed conditions was .69, and for passive sentences, it was .51, suggesting that although the learners had a tendency to use a particular structure across the two conditions, their use of a certain structure was influenced by the primes. By contrast, priming did not seem to work for the dative construction: learners were overwhelmingly in favor of using the prepositional object rather than the direct object regardless of the type of primes they received; T-tests also showed no significant differences. The correlations between the learners' production of PO sentences after PO and DO primes ($r = .90$) and between their production of DO sentences after DO and PO primes ($r = .88$) were very strong, suggesting that their use or nonuse of the structures was consistent across the priming conditions and that they did not respond to the primes to which they were exposed. Based on the analysis in the preceding text, we decided to use the passive scores as a proxy of learners' priming performance in all subsequent analyses given that the passive is a less frequent structure and that priming was effective for this structure.

Priming and Other Cognitive Abilities

The analyses in this section concern the relationship between priming and other measures of implicit aptitude and measures of explicit aptitude (see Table 3 for the correlation matrix and Table 4 for the descriptive statistics). The analyses sought primarily evidence for the convergent and divergent validity of priming as a measure of implicit aptitude. If

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

TABLE 3. Correlation matrix

	Priming	SRT	LLAMA_B	LLAMA_D	LLAMA_E	LLAMA_F	UGJT	MKT	EIT
Priming	–								
SRT	−.22*								
LLAMA_B	.09	.04							
LLAMA_D	.10	.06	.30**						
LLAMA_E	.00	.07	.26**	.13					
LLAMA_F	.05	.04	.46*	.19*	.34**				
UGJT	.00	.00	.09	.08	.15	.26**			
MKT	−.03	−.16	−.02	−.02	.13	.22*	.41**		
EIT	−.00	.08	.11	.11	.16*	.20*	.44**	.23**	–

* $p < .05$; ** $p < .01$.
SRT: serial reaction time (sequence learning); UGJT: untimed grammaticality judgment test; MKT: metalinguistic knowledge test; EIT: elicited imitation test.

TABLE 4. Descriptive statistics

Tests	n	Mean	SD	Range	Reliability
Priming	134	4.37	1.58	0–8	.51
LLAMA_B	162	44.23	21.94	0–100	N/A
LLAMA_D	163	30.89	15.42	0–65	N/A
LLAMA_E	163	55.15	24.99	0–100	N/A
LLAMA_F	162	55.74	25.01	0–100	N/A
Serial reaction time	151	4.33	12.84	−31–41	.97/.95[#]
Untimed grammaticality judgment	140	24.58	3.70	15–32	.79
Metalinguistic knowledge test	166	12.17	1.63	6–16	.72
Elicited imitation test	159	15.36	4.49	2–26	.74

N/A: not available.
[#].97 for stimuli based on the target sequence and .95 for stimuli based on the control sequence.

our hypothesized model (Figure 2) is correct, then priming, sequence learning, and LLAMA_D would form a higher-order latent factor, which we can call "implicit aptitude," and the remaining subtests of the LLAMA would underlie a factor called "explicit aptitude." The two latent variables should be weakly correlated or uncorrelated. However, the hypothesized model was not identified, and modified models with the three hypothesized implicit measures in one latent factor all showed poor fit indices. The final confirmed model displayed in Figure 3 (see Table 5 for the estimates) showed a good fit for the data: $X^2 = 2.83$, $p = .95$, NFI = .97, RFI = .91, CFI = 1.00, and RMSEA = .00. In this model, sequence learning is significantly and negatively correlated with priming, and neither implicit measure is significantly correlated with explicit aptitude, which is represented by the four subtests of the LLAMA. LLAMA_D, which was hypothesized to measure implicit aptitude, turned out to converge with measures of explicit aptitude rather than implicit aptitude. Thus, in terms of divergent validity, priming is indeed unrelated to explicit aptitude. However, in terms of convergent validity, it is negatively correlated with sequence learning and the two failed to converge under the same factor, suggesting that they do not belong to the same construct.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

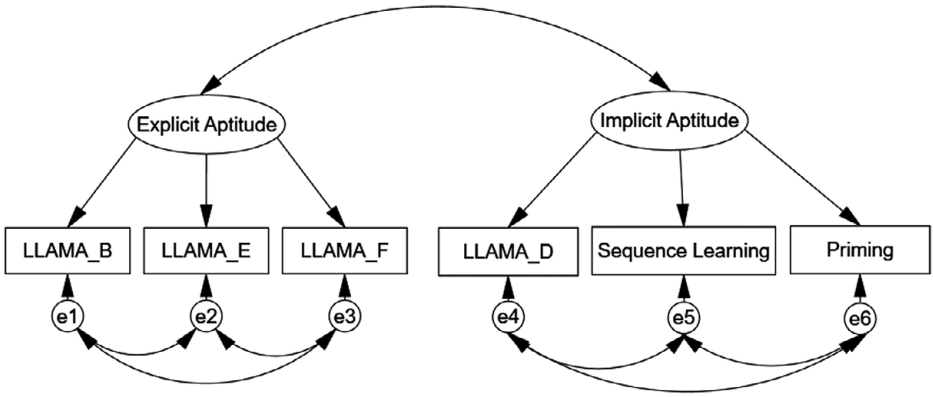


FIGURE 2. Hypothesized measurement model for explicit and implicit aptitude.

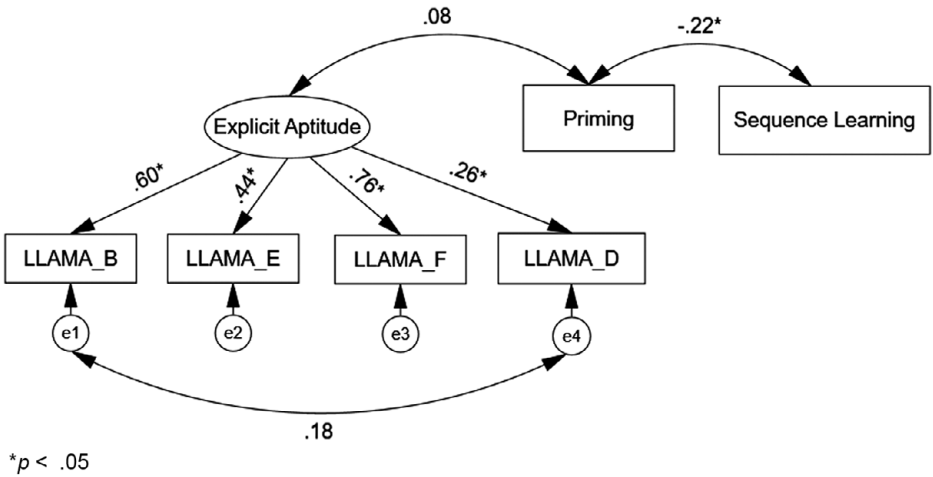


FIGURE 3. Identified measurement model for explicit and implicit aptitude.

THE RELATIONSHIP BETWEEN SYNTACTIC PRIMING AND L2 PROFICIENCY

Based on the preceding measurement model regarding the structure of cognitive ability and on the findings of previous research (Granena, 2013; Suzuki & DeKeyser, 2017), a structural model mapping the relationship between aptitude measures and outcome measures was hypothesized (Figure 4). In this model, explicit aptitude is predictive of explicit knowledge and the two measures of implicit aptitude are predictive of implicit knowledge. However, the model was not confirmed, and an alternative model was identified (Figure 5 and Table 6), which showed a good fit, $X^2 = 14.87$, $p = .92$, NFI = .91, CFI = 1.00, and RMSEA = .00. In this model, explicit aptitude formed by LLAMA_B, _D, _E, and _F is significantly predictive of L2 proficiency, a latent factor

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

TABLE 5. The measurement model for cognitive abilities

Variable A	Variable B	Estimates	Significance (p)
Explicit aptitude	→LLAMA_F	$.76^{\beta}$	.00*
	→LLAMA_E	$.44^{\beta}$	.00*
	→LLAMA_B	$.59^{\beta}$	.00*
	→LLAMA_D	$.26^{\beta}$	.01*
LLAMA_B (errors)	↔LLAMA_D (errors)	$.18^{r}$	.09
Explicit aptitude	↔Priming	$.08^{r}$	.42
Sequence learning	↔Priming	$-.22^{r}$	.02*

$^{\beta}$: standardized regression coefficient; ^{r}: correlation; *$p < .05$.

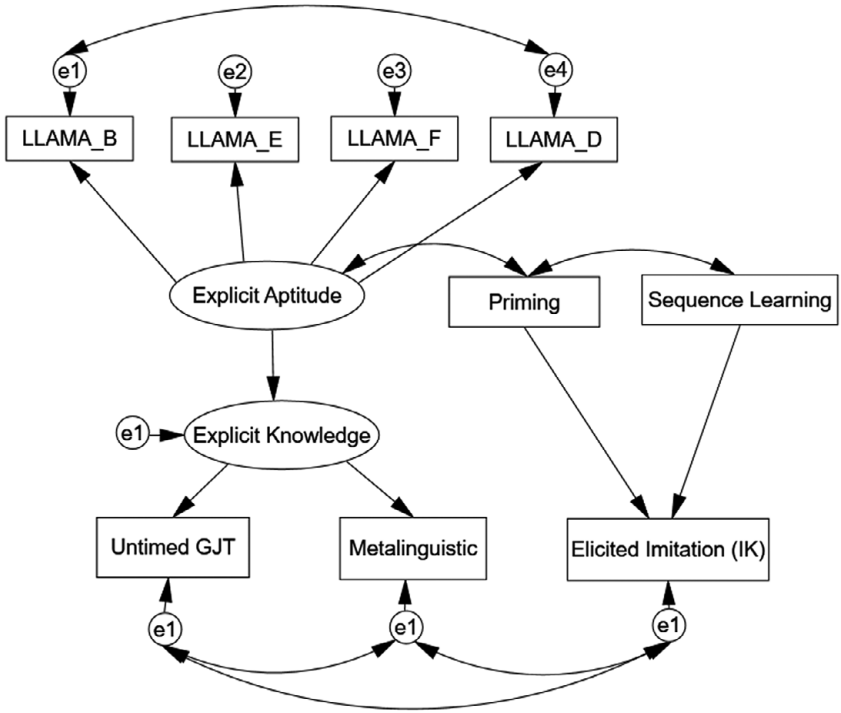
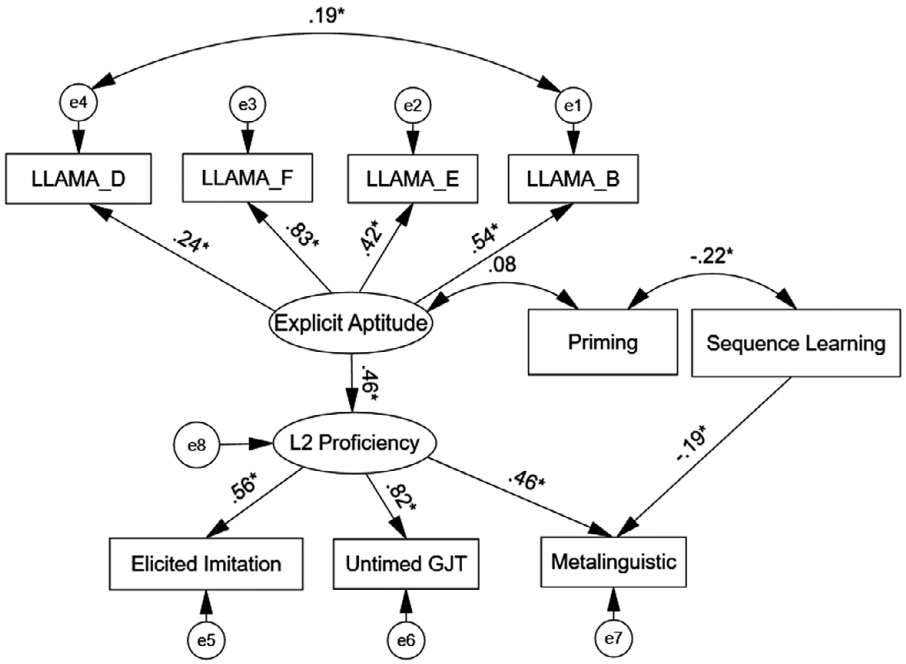


FIGURE 4. Hypothesized structural model for cognitive ability and L2 proficiency.

represented by the three proficiency tests. Elicited imitation, which was hypothesized to measure implicit knowledge, was found to load onto the same factor as the two measures of explicit knowledge. For the untimed written grammaticality judgment test, scores based on ungrammatical items were used rather than composite scores or scores based on grammatical items because ungrammatical items showed a better fit for the data. Regarding the two measures of implicit aptitude, sequence learning was negatively predictive of metalinguistic knowledge, but priming was not a significant predictor of any outcome measure.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press



grammaticality judgment test; metalinguistic: metalinguistic knowledge test; *p < . 05.

FIGURE 5. Identified structural model for cognitive ability and L2 proficiency.

TABLE 6. The structural model for cognitive abilities and L2 proficiency

Variable A	Variable B	Estimates (β)	Significance (p)
L2 proficiency	→Elicited IMITATION	.56	.00*
	→Untimed grammaticality judgment	.82	.00*
	→Metalinguistic knowledge	.46	.00*
Explicit aptitude	→L2 proficiency	.46	.00*
Sequence learning	→Metalinguistic knowledge	−.19	.02*

*p < .05.

DISCUSSION

This study is the first to investigate whether syntactic priming is a valid measure of implicit language aptitude. The validity of the construct rests on evidence for divergent validity, that is, whether it is different from explicit aptitude; convergent validity, namely whether it is correlated with other measures of implicit aptitude such as sequence learning; and predictive validity, which refers to whether it is predictive of learning outcomes. The study found evidence for divergent validity, but there was no evidence for the other two types of validity. In the following sections, we interpret the findings regarding the associations between priming and other variables by referring to theories, previous findings, and the methodological features of this study.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

DIVERGENT VALIDITY

Priming is distinct from explicit aptitude, which constitutes evidence for divergent validity. More specifically, there is a lack of associations between priming, on the one hand, and the factor of explicit aptitude (r = .08) and the observed variables (i.e., the LLAMA tests) subsumed by the factor (r's ranging from -.01 to .10, n.s.), on the other. The finding adds further evidence to the conclusions reached by priming researchers (Bock & Griffin, 2000; Pickering & Ferreira, 2008) about the implicit nature of the learning resulting from priming. As far as this study is concerned, the following features and evidence contribute to the "implicit learning" account of priming. First, there was no device drawing learners' attention to the target structures. When asked about the objective of the task, almost all participants indicated that it was a memory task, and none mentioned the target structures. Second, there was no semantic connection between the prime and the target. Therefore, any priming effects must have been due to the syntactic abstractions rather than episodic memory. Third, because there was no lexical overlap between the prime and target, there was no lexical boost which may have contributed to explicit learning. Fourth, the objective of the priming was further masked by the fact that half the items were distractors, and the target items were dispersed among the four structures that served as distractors to each other.

This study also shows that sequence learning, another measure of implicit aptitude, was unrelated to explicit aptitude measured by the LLAMA aptitude tests. This finding reinforces existing evidence for the divergent validity of sequence learning. For example, Suzuki and DeKeyser (2017) found no correlation between sequence learning and LLAMA_F (language analytic ability), Granena (2019) reported no correlations between sequence learning and any LLAMA scores, and Hamrick (2015) found no correlation between sequence learning and LLAMA_B (rote memory).

While priming and sequence learning were found to be separate from explicit aptitude, LLAMA_D, which was hypothesized to be a measure of implicit aptitude, turned out to be part of explicit aptitude. As with other LLAMA tests, LLAMA_D includes a study phase and a testing phase. A critical aspect of the test (and other LLAMA tests) is that during the study phase, learners were informed of the prospect of being tested on the materials they were presented with, which, according to Hulstijn (2005), is a defining feature of explicit instruction because it encourages intentional/conscious learning. Suzuki (this issue) found that even with changed instructions that encouraged implicit/incidental learning, accuracy scores on LLAMA_D still represent explicit learning. Evidence for the explicit nature of LLAMA_D is also available from other studies, for example, it was found to be correlated with explicit (metalinguistic) knowledge (Granena, 2013). Based on the finding of this and other studies, we conclude that LLAMA_D is at least an inconsistent measure of implicit aptitude and that it is likely a measure of explicit aptitude.

CONVERGENT VALIDITY

There is no evidence for convergent validity. The three measures of implicit aptitude—priming, sequence learning, and LLAMA_D—failed to load on the same factor. LLAMA_D clustered with explicit aptitude, as discussed in the preceding text. Therefore, caution must be exercised when using LLAMA_D as a measure of implicit aptitude and

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

TABLE 7. A side-by-side comparison between priming and sequence learning

Dimensions of comparison	Priming	Sequence learning
Definition	Tendency to be influenced by a recent experience	Tendency to learn from repeated practice
Mechanism	Habit shattering	Habit developing
Amount of input	One encounter	Repeated exposure
Evaluation of performance	Accuracy	Reaction time
Role in learning	Learning elemental material	Proceduralizing
Stage of learning	Early	Advanced
Domain of learning	Verbal	Nonverbal

when interpreting related findings. Priming and sequence learning not only failed to converge on one factor but also showed a significant negative correlation, suggesting that learners who were more likely to be primed were less likely to learn from the serial reaction time task. An explanation for the negative correlation can be sought by considering the learning mechanisms of the two types of learning tasks (see Table 7 for a comparison). Priming happens in verbal tasks, and it concerns the ability to be influenced by a recent event, especially one that involves a less frequent phenomenon such as the passive voice (compared with the active voice).[6] Therefore, priming is about the ability to change habits and the tendency to respond to less familiar stimuli. Sequence learning happens in nonverbal tasks through repeated exposure to arbitrarily combined, meaningless strings; it relies on frequency and unconscious computations of the probabilistic relationships and contingencies between stimuli. Thus, it would seem that priming concerns habit-shattering while sequence learning is about habit-forming. Priming may contribute to the learning of "elemental materials" (Woltz, 2003, p. 98) at early stages of learning, while sequence learning is key to proceduralizing knowledge through repeated practice (Squire & Zola, 1996). Also, priming is based on accuracy, that is, whether learners produced the target structure, while sequence learning is based on latency (reaction time), namely how fast learners respond to target and control stimuli. The extent to which accuracy and latency are comparable is unknown, and there is evidence that psychometric tests based on accuracy and latency lead to inconsistent construct validity when they are used to measure the same construct (Rey-Mermet et al., 2019). While we hope the previously mentioned comparison can shed some light on the negative correlation between the two cognitive abilities, evidently more research is needed to arrive at a clearer understanding of their relationship.

In fact, the lack of correlations, or rather positive correlations, between measures of implicit learning/aptitude is the norm rather than anomaly. For example, Kaschak et al. (2011) found no correlation between implicit learning measured through the Simon task and syntactic priming, and they found a near-significant (p = .06) negative correlation between the Simon task and the priming of the double-object structure. Morgan-Short and Buffington (this issue) show that the three tasks that have been used to measure procedural memory (implicit aptitude) failed to converge, and one of them even clustered with declarative memory (explicit aptitude). The lack of correlations between measures of

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

implicit aptitude/learning poses a great challenge for the establishment of construct validity. We would like to discuss this phenomenon from the following perspectives. First, like implicit learning, implicit aptitude lacks individual variation, which makes it difficult to achieve significant correlations between different measures of the construct. Second, implicit learning is subtle and happens without awareness, which makes it hard to capture or measure. Third, the measures of implicit aptitude often lack reliability because of the uncontrolled nature of the testing tasks (Ward et al., 2013). Fourth, the lack of correlations may suggest that implicit learning may happen in multiple ways, and the different types of implicit learning may represent different mechanisms.

With respect to convergent validity, unlike implicit aptitude, explicit aptitude seems to be a robust construct where all components loaded significantly under the same latent factor. What is common about the four LLAMA tests[7] is that they all encourage explicit learning: in all four subtests, there is a study phase where learners engage in the conscious processing of verbal and nonverbal materials and a test phase where they are required to consciously retrieve the learned knowledge. The convergent validity of the LLAMA battery seems robust in that in previous research all the subtests, except for LLAMA_D, consistently loaded on the same factor (Granena, 2015, 2019), indicating that they measure different dimensions of the same construct. LLAMA_D has been an inconsistent component of explicit aptitude because (a) it loaded on separate factors from the other three subtests in some studies (Artieda & Muñoz, 2016; Granena, 2012), and (b) when it loaded on the same factor as other subtests, its factor loadings were the weakest (Granena, 2015; this study). The findings seem to suggest that LLAMA_D is the "least explicit" among the four subtests and should be excluded from the LLAMA battery if the purpose is to measure explicit aptitude.

PREDICTIVE VALIDITY

We hypothesized that syntactic priming would predict implicit knowledge measured by the elicited imitation test. However, the hypothesis was not confirmed. It is possibly because the knowledge reflected in this test is explicit knowledge: the test scores loaded under the same factor as the two measures of explicit knowledge; it was also significantly correlated with LLAMA_B and LLAMA_E, two measures of explicit aptitude. These results suggest that the elicited responses from the learners represent explicit rather than implicit knowledge. It is possible that these learners, who had received heavy form-focused instruction throughout their learning experience, did not possess implicit knowledge and accessed their explicit knowledge instead during the elicited imitation test. Zhang (2015) conducted a study to explore Chinese university EFL learners' explicit and implicit L2 English knowledge using similar tests as this study. Although the author claimed that the learners' implicit and explicit knowledge were separable, a closer inspection of the findings revealed that they were strongly correlated, $r = .86$. Kim and Nam (2017) reported that Korean EFL learners' explicit knowledge measured through a metalinguistic knowledge test was separable from their implicit knowledge operationalized as elicited imitation, with a correlation of .15. However, the factor based on two timed grammaticality tests (written and oral), referred to as "weak" implicit knowledge, was strongly correlated with elicited imitation, $r = .74$. The point here is that the distinction between learners' explicit and implicit knowledge is equivocal in foreign

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

language settings, and likely the bulk of L2 knowledge learners possess is explicit rather than implicit.

The explicit nature of the learners' L2 knowledge may also explain why sequence learning was not predictive of the latent factor of L2 proficiency. Before delving further into the relationships between sequence learning and the proficiency measures, we would like to distinguish two types of L2 knowledge: analyzed knowledge and metalanguage, both of which are explicit knowledge (Ellis, 2005). Analyzed knowledge is knowledge learners are aware of, and metalanguage refers to language used to describe how language works—linguistic rules. We argue that the latent factor of L2 proficiency, which is the common variance of the three proficiency measures: untimed grammaticality judgment, metalinguistic knowledge, and elicited imitation, represents analyzed knowledge. The metalinguistic knowledge test measures both metalanguage and analyzed knowledge, and it is analyzed knowledge that contributes to the latent factor. Based on this distinction, sequence learning is not a significant predictor of the common core—analyzed knowledge. The lack of correlations between sequence learning and analyzed knowledge (measured using grammaticality judgment tests) has also been obtained by previous correlational studies investigating the relationship between sequence learning and L2 attainment at fixed time points (Granena, 2013; Suzuki & DeKeyser, 2017). However, it is inconsistent with the results of an experimental study conducted by Yilmaz and Granena (this issue), who found sequence learning a significant moderator of the effects of implicit feedback measured using a grammaticality judgment test. It is possible that in their study, when making grammaticality judgments, learners relied on implicit rather than explicit knowledge, given that the knowledge resulted from implicit feedback about a linguistic structure that involves the arbitrary associations between nouns and gender markers. It is also possible that sequence learning leads to explicit knowledge, that is, rule knowledge that is learned unconsciously may turn into conscious knowledge. Therefore, sequence learning may facilitate the learning of both implicit and explicit knowledge, although it is more likely to contribute to the former. Certainly, this hypothesis needs to be empirically tested.

Regarding the negative predictive path between sequence learning and metalinguistic knowledge, the finding constitutes predictive validity from another perspective, that is, as a measure of implicit aptitude, sequence learning is fundamentally different from the mechanism for the learning of metalanguage; learners who are good at learning grammar rules unconsciously are not good at understanding or learning L2 metalanguage or they tend not to do so. There is little literature to refer to regarding the associations between implicit aptitude and metalanguage. Granena (2013) is perhaps the only relevant study, which showed that metalinguistic knowledge was uncorrelated with sequence learning but was significantly and positively correlated with LLAMA_D. However, the metalinguistic knowledge test in Granena's study was not a pure test of metalanguage because it consisted of three components: grammaticality judgment, error correction, and rule description; only rule description measures metalanguage. Also, LLAMA_D is possibly a component of explicit aptitude, as this study showed.

While the predictive validity of implicit aptitude is unclear, explicit aptitude showed a clear, significant effect on the outcome measures, which underlie the common factor of analyzed (explicit) knowledge. Explicit aptitude was measured by means of the LLAMA test battery, which, together with other tests of traditional aptitude such as the MLAT and

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

PLAB, has been found to be a set of abilities that are more strongly, if not only, correlated with the learning resulting from explicit than implicit instructional treatments (Li, 2015; Yilmaz & Granena, this issue). It is also more likely to be implicated at initial stages of learning than advanced learning (Li, 2016), with the former being more likely to be explicit than the latter.

CONCLUSION

The study was undertaken to validate syntactic priming as a measure of implicit aptitude. The results showed some evidence for divergent validity, that is, for being separate from explicit aptitude. However, it failed to show convergent validity with the other measure posited to gauge implicit aptitude. The lack of convergent validity of measures of implicit aptitude found in this study, together with similar findings of other studies in this issue (Buffington & Morgan-Short; Godfroid & Kim), leads us to endorse Gebauer and Mackintosh's (2007) "modular view of implicit learning" as opposed to "the notion of general implicit learning ability" (p. 48). Thus, we recommend conceptualizing explicit and implicit aptitude in different ways: measures of explicit aptitude share common variance while measures of implicit aptitude are likely distinct. Priming also lacks predictive validity—it is not predictive of L2 attainment, which we attributed to learners' lack of implicit knowledge. Interestingly, sequence learning was a negative predictor of metalinguistic knowledge, which constitutes evidence for construct validity in that explicit knowledge in the form of metalanguage is learned through a completely different avenue. As this study is the first to examine syntactic priming as an ability for implicit learning, these findings are preliminary and by no means conclusive. We hope to stimulate more interest in the notion of priming as implicit aptitude and call for more research on this topic.

Next, we would like to make the following recommendations regarding the methods of the task used to measure learners' priming performance. First, for priming to be investigated as a predictor, priming must be effective, and the researcher must recognize that not all linguistic structures are "primeable." In this study, the priming task targeted two structures: dative and transitive verbs, and priming effects were only evident for the latter. Second, primed responses must be coded as a continuous variable instead of a categorical variable. Third, priming normally involves two alternatives of the same structure such as passive voice versus active voice, and a decision must be made on which structure to be used as a proxy of learners' priming performance. We recommend the structure that is less frequent in the language in which the stimuli are created. It would be unreasonable to use a composite score because there is often a negative correlation between them, which means that they represent different abilities.

Finally, we recommend that future research address the following issues when designing a study to map the relationships between implicit/explicit aptitude and implicit/explicit knowledge. First, a study investigating the predictive validity of implicit aptitude should involve learners that likely possess implicit knowledge, such as advanced learners or learners in a naturalistic setting where opportunities for the development of implicit knowledge abound. In this study, the learners did not seem to have much implicit knowledge, which made it difficult to investigate the predictive validity of implicit aptitude. Second, given the multidimensionality of implicit aptitude and the hypothesis

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

that different implicit learning abilities are responsible for different aspects of language learning, it is necessary to include multiple measures of implicit knowledge with a view to exploring whether priming is predictive of certain measures of implicit knowledge but not others. This study used only one test of implicit knowledge, and it is unclear whether including other measures such as a task of free oral production or a word-monitoring task (Suzuki & DeKeyser, 2015) would have changed the results. Third, other measures of explicit aptitude such as working memory or alternative instruments such as the MLAT and the PLAB should be utilized to cross-validate the findings.

NOTES

[1]One anonymous reviewer pointed out that the L2 Spanish studies reviewed in this study may not be representative of classroom research because of the prevalence of Spanish speakers in the United States.

[2]The studies investigated L1 priming unless otherwise specified.

[3]One reviewer argued that although cognitive psychologists regard priming as a measure of implicit learning, it should be viewed as a measure of implicit memory. This is because priming primarily implicates the activation of preexisting syntactic knowledge, especially when it is tested in learners' L1.

[4]Using learners' L1 rather than their L2 as the language of test stimuli is typical in the measurement of cognitive abilities in L2 learning. The purpose is to prevent the influence of learners' L2 proficiency on their performance in the measured ability. For example, the MLAT (Carroll & Sapon, 2002), the most influential test of language aptitude, is developed in English—the native language of foreign language learners in the United States, for whom the test is intended. In the research on working memory, there is evidence that L2 working memory tests show stronger associations with L2 learning than L1 working memory tests (Linck et al., 2014), which prompted the call to use L1 rather than L2 stimuli in the measurement of working memory (Juffs & Harrington, 2012). However, L2 stimuli would be ideal if the impact of learners' L2 proficiency on their performance in the cognitive test could be controlled or minimized.

[5]This scoring method is different from Ellis (2005), where learners were not required to make corrections to incorrect sentences and answers were scored as correct or incorrect.

[6]An anonymous reviewer speculated that a higher priming score may represent weaker implicit learning ability because it suggests that the learner has a lesser command of his/her native language. This may explain the negative correlation between priming and sequence learning. This is an interesting perspective that needs to be empirically tested.

[7]See Bokander and Bylund (2020) for the internal validity of the LLAMA tests.

REFERENCES

Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences*, *50*, 42–48.

Bock, J. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.

Bock, J., Dell, G., Chang, F., & Onishi, K. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*, 437–458.

Bock, J., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*, 177–192.

Bock, J., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171.

Bokander, L., & Bylund, E. (2020). Probing the internal validity of the LLAMA language aptitude tests. *Language Learning*, *70*, 11–47.

Cai, Z., Pickering, M., Yan, H., & Branigan, H. (2011). Lexical and syntactic representations in closely related languages: Evidence from Cantonese-mandarin bilinguals. *Journal of Memory and Language*, *65*, 431–445.

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press

Cai, Z. G., Pickering, M. J., Wang, R., & Branigan, H. P. (2015). It is there whether you hear it or not: Syntactic representations of missing arguments. *Cognition*, *136*, 255–267.
Carroll, J., & Sapon, S. (2002). *Manual for the MLAT.* Second Language Testing, Inc.
DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*, 499–533.
DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 83–104). Routledge.
Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141–172.
Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*, 147–171.
Ellis, R., Li, S., & Zhu, Y. (2019). The effects of pre-task explicit instruction on the performance of a focused task. *System*, *80*, 38–47.
Evans, J., & Frankish, K. (2009). *In two minds: Dual processes and beyond*. Oxford University Press.
Ferreira, V., Bock, J., Wilson, M., & Cohen, N. (2008). Memory for syntax despite amnesia. *Psychological Science*, *19*, 940–946.
Ferreira, V. S., & Bock, J. K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*, 1011–1029.
Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. Behavior Research Methods, *Instruments, & Computers*, *35*, 116–124.
Gebauer, G., & Mackintosh, N. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 34–54.
Granena, G. (2012). *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment (Unpublished PhD dissertation)*. The University of Maryland.
Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, *63*, 665–704.
Granena, G. (2015). Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied Psycholinguistics*, *37*, 577–600.
Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency. *Studies in Second Language Acquisition*, *41*, 313–336.
Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, *29*, 311–343.
Granena, G., & Yilmaz, Y. (2019). Corrective feedback and the role of implicit sequence learning ability in L2 online performance. *Language Learning*, *69*, 127–156.
Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9–15.
Hiver, P., & Al-Hoorie, A. (2020). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of You, Dornyei and Csizer (2016). *Language Learning*, *70*, 48–102.
Huang, J., Pickering, M., Yang, J., Wang, S., & Branigan, H. (2016). The independence of syntactic processing in Mandarin: Evidence from structural priming. *Journal of Memory and Language*, *91*, 81–98.
Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, *27*, 129–140.
Hwu, F., Wei, P., & Sun, S. (2014). Aptitude-treatment interaction effects on explicit rule learning: A latent growth curve analysis. *Language Teaching Research*, *18*, 294–319.
Jackson, C. N. (2018). Second language structural priming: A critical review and directions for future research. *Second Language Research*, *34*, 539–552.
Juffs, A., & Harrington, M. (2012). Aspects of working memory in L2 learning. *Language Teaching*, *44*, 137–166.
Kaschak, M. P., Kutta, T. J., & Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. *Psychonomic Bulletin & Review*, *18*, 1133–1139.
Kaufman, S., DeYoung, C., Gray, J., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*, 321–340.
Kim, J., & Nam, H. (2017). Measures of implicit knowledge revisited: Processing modes, time pressure, and modality. *Studies in Second Language Acquisition*, *39*, 431–457.

Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, *29*, 149–154.

Kline, R. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.

Li, C., & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar*. Los Angeles, CA: University of California Press.

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, *36*, 385–408.

Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition*, *38*, 801–842.

Li, S. (2017). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors and researchers* (pp. 41–70). John Benjamins.

Li, S., Shintani, N., & Ellis, R. (2012). Doing meta-analysis in SLA: Practices, choices, and standards. *Contemporary Foreign Language Studies*, *384*, 1-27.

Linck, J., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*, 530–566.

Linck, J., Osthus, P., Koeth, J., & Bunting, M. (2014). Working memory and second language comprehension and produciton: A meta-analysis. *Psychonomic Bulletin & Review*, *21*, 861–883.

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

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27.

Masgoret, A., & Gardner, R. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, *53*, 123–163.

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition*, *28*, 179–207.

McDonough, K. (2011). Eliciting wh-questions through collaborative syntactic priming activities during peer interaction. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research: insights from psycholinguistics* (pp. 131–151). John Benjamins.

McDonough, K., & Fulga, A. (2015). The detection and primed production of novel constructions. *Language Learning*, *65*, 326–357.

Pickering, M., & Ferreira, V. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford University Press.

Re-Mermet, A., Souza, A., Gade, M., & Bastian, C. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*, *148*, 1335–1372.

Saito, K., Suzukida, Y., & Sun, H. (2019). Aptitude, experience, and second language pronunciation proficiency development in classroom settings. *Studies in Second Language Acquisition*, *41*, 201–225.

Schumacker, R., & Lomax, R. (2010). *A beginner's guide to structural equation modelling*. Routledge.

Skehan, P. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics*, *36*, 367–384.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, *93*, 13515–13522.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*, 860–895.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*, 747–790.

Tagarelli, K., Ruiz, S., Moreno Vega, J., & Rebuschat, P. (2016). Variability in second language learning: The roles of individual differences, learning conditions, and linguistic complexity. *Studies in Second Language Acquisition*, *38*, 293–316.

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*, 363–387.

Ward, E., Berry, C., & Shanks, D. (2013). Age effects on explicit and implicit memory. *Frontiers in Psychology*, *4*, 1–11. https://doi.org/10.3389/fpsyg.2013.00639.

Woltz, D. (2003). Implicit cognitive processes as aptitudes for learning. *Educational Psychologist*, *38*, 95–104.

Yi, W. (2018). Statistical sensitivity, cognitive aptitudes, and processing of collocations. *Studies in Second Language Acquisition*, *40*, 831–856. First view.

Yilmaz, Y., & Granena, G. (2019). Cognitive individual differences as predictives of improvement and awareness under implicit and explicit feedback conditions. *Modern Language Journal*, *103*, 686–702

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition*, *37*, 457–486.

Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *Modern Language Journal*, *103*, 763–781.

APPENDIX

Year 2 and Year 4 Students' Test Scores

	Untimed grammaticality judgment		Metalinguistic knowledge		Elicited imitation	
Level	Mean	SD	Mean	SD	Mean	SD
Year 2	24.92	3.73	12.15	1.65	15.18	3.84
Year 4	24.14	3.66	12.19	1.63	15.61	5.29

https://doi.org/10.1017/S0272263120000698 Published online by Cambridge University Press