APTARA | MODL | modl12585 | Dispatch: July 22, 2019 | CE: XXX
Journal | MSP No. | No. of pages: 7 | PE: XXXX

The Desirable Difficulty Framework as a Theoretical Foundation for Optimizing and Researching Second Language Practice

YUICHI SUZUKI,[1] TATSUYA NAKATA,[2] and ROBERT DEKEYSER[3]

[1]*Kanagawa University, Faculty of Foreign Languages, 3-27-1, Rokkakubashi, Kanagawa-ku, Yokohama-shi, Kanagawa 221-8686, Japan Email: szky819@kanagawa-u.ac.jp*
[2]*Hosei University, Faculty of Letters, 2-17-1 Fujimi, Chiyoda-ku, Tokyo 102-8160, Japan Email: t-nakata@hosei.ac.jp*
[3]*University of Maryland College Park, School of Languages, Literatures, and Cultures, 3104 Jimenez Hall, College Park, MD 20742 Email: rdk@umd.edu*

This coda article offers unified theoretical accounts of the major findings of the empirical studies in this special issue of *Optimizing Second Language Practice in the Classroom: Perspectives from Cognitive Psychology*. We present a theoretical framework from cognitive psychology (desirable difficulty framework) and link it to the ideas of second language (L2) difficulty. We argue that practice condition, linguistic difficulty, and individual differences need to be taken into account for creating optimal, deliberate, and systematic L2 practice. The desirable difficulty framework may serve as a theoretical foundation to better understand the role of practice on L2 acquisition, as well as to gain insights into effective L2 teaching. Future directions for research are presented to further develop this emerging field of L2 practice.

Keywords: practice; desirable difficulty framework; second language difficulty; cognitive psychology; deliberate and systematic practice

INFORMED BY THEORIES AND FINDINGS from the second language acquisition (SLA) and cognitive psychology literature, the contributors in this special issue have elucidated the effects of various practice variables that induce different learning processes and outcomes. The practice activities examined in this issue can be subsumed in the broad definition of second language (L2) practice, that is, "specific activities in the second language, engaged in systematically, deliberately, with the goal of developing knowledge of and skills in the second language" (DeKeyser, 2007, p. 1). With this definition in mind, in this coda article, we further extend and develop what "systematic and deliberate" practice means. We do this by drawing on the desirable difficulty framework (Bjork, 1994, 2018; Schmidt & Bjork, 1992), which may serve as an overarching framework to create optimal, systematic, and deliberate practice conditions.

The goals of L2 learning are to develop knowledge and skills that are durable in the long term and transferrable to a new context. The desirable difficulty framework can provide an excellent account of how these goals can be achieved for a wide variety of cognitive and motor skills (Bjork, 1994; Schmidt & Bjork, 1992). The framework predicts that when learners experience optimal levels of difficulty during practice, their ~~postpractice~~ retention and transfer become maximized. Creating difficulty for learners sometimes slows down the initial rate of learning and leads to lower accuracy rates during practice, but it can eventually enhance long-term retention and transfer (Bjork, 2018). The desirably difficult practice may relate to the concept of deliberate practice,

The Modern Language Journal, 0, 0, (2019)
DOI: 10.1111/modl.12585
0026-7902/19/1–7 $1.50/0

which is considered to be an essential ingredient for achieving high levels of skills and knowledge in expert research (Ericsson, 2006); deliberate practice requires a learner's full attention and takes place outside one's comfort zone, bringing out near-maximal effort and allowing for increasing the current level of skills and knowledge. This is completely different from "both mindless, routine performance and playful engagement" (Ericsson, 2006, p. 692). The idea of desirable difficulty is useful when designing L2 practice because practice activities can be systematically designed to impose optimal, challenging levels of engagement on learners.

In L2 acquisition research, the concept of difficulty is often elusive and very difficult to pin down. Housen and Simoens (2016) propose that L2 difficulty is multi-componential, and is defined in relative terms as "how costly, demanding, or difficult a given language feature is for a given language learner in a given learning context, particularly in terms of the mental resources allocated and cognitive mechanisms deployed in processing and internalizing the feature" (p. 166). Their comprehensive view of difficulty captures myriads of interrelated factors involved in L2 practice such as practice conditions, linguistic features, and learner-related factors.

By drawing on the desirable difficulty framework (Bjork, 2018; Schmidt & Bjork, 1992) and linking it to the cognitive difficulty framework (Housen & Simoens, 2016), we attempt to provide a unified framework and rationales for designing optimal practice conditions for L2 learning. Although a variety of practice-related variables are intertwined in classroom, understanding how each variable and its combination influence L2 learning processes provides a strong theoretical backbone for teachers to think and act upon. The empirical findings may prompt L2 teachers to reflect on their classroom teaching such as design of materials, intervals of repeated practice, and individual differences of learners (Rankin & Becker, 2006), as well as provide emotional support for their teaching (Sato & Loewen, 2018).

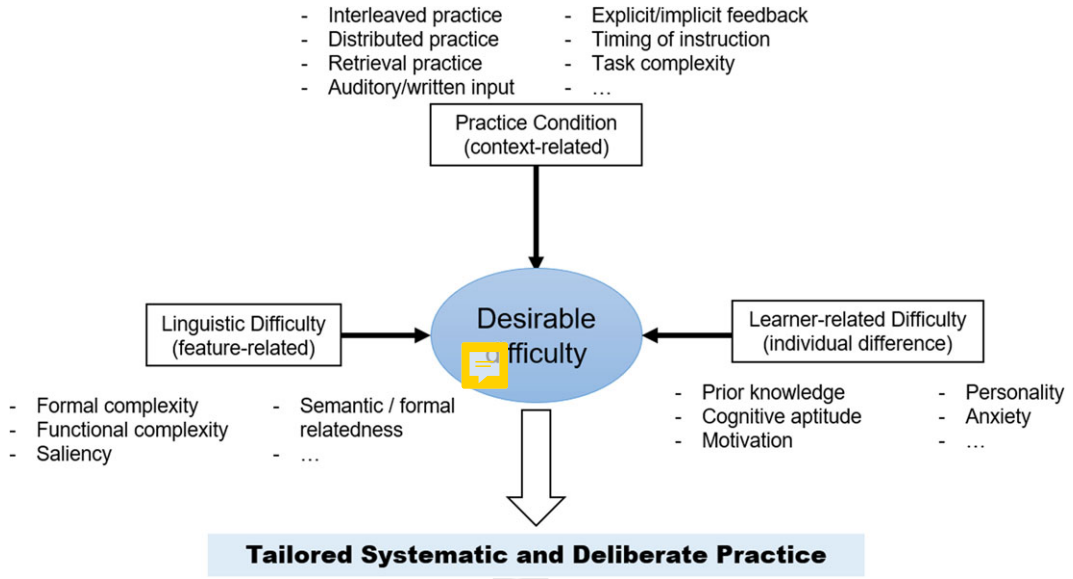
L2 COGNITIVE DIFFICULTY AND DESIRABLE DIFFICULTY FRAMEWORK

As described in the previous section, Housen and Simoen (2016) propose that L2 difficulty is a multi-componential concept that is broadly categorized as (a) linguistic difficulty, (b) learner-related difficulty, and (c) context-related difficulty (for a similar view, see DeKeyser, 2005, 2016). We adopted their framework and illustrate how an array of factors influence L2 difficulty (Figure 1).

Linguistic difficulty results from structural and conceptual properties of language features (i.e., formal complexity and functional complexity). In terms of formal complexity, for instance, *-ing* has no allomorph (simple), whereas *-ed* and *-s* have more than one allomorph (complex). Functional complexity concerns "the number and nature of the meanings and functions expressed" (Housen & Simoen, 2016, p. 168). The meanings of English articles, for instance, are more abstract than plural *s* for many L2 learners. Second, the learner-related difficulty concerns learners' abilities and capacity to learn L2 such as cognitive aptitudes (e.g., inductive ability, working memory, explicit and implicit learning abilities), conative and affective factors (e.g., motivation, personality, anxiety), and prior knowledge including first language (L1) knowledge. Last, the context-related difficulty is defined as the degree of the demands that L2 learning treatment (e.g., interleaved practice, explicit corrective feedback) and educational and social context (e.g., teacher and peer pressure to communicate in a classroom) induce. The context-related difficulty is also referred to as the practice condition, which is the central variable of interest throughout this special issue (see the next section).

All three factors above contribute to L2 difficulty. Accuracy/error data *during* practice (e.g., proportion of correct responses on grammar exercises [Nakata & Suzuki, in this issue] or *during* vocabulary exercises [Strong & Boers, in this issue]) can be used as a measure of L2 difficulty. Perhaps, in future research, new technologies such as eye tracking (Révész & Gurzynski–Weiss, 2016) and pupillometry (Schmidtke, 2018) can also provide more sensitive measures of difficulty experienced by learners. In addition to the objectively measured data, perceived ease or difficulty of training on the part of learners may also be a valuable index of difficulty. The perceived difficulty is closely tied to the perceived effectiveness of training; learners tend to consider training to be effective when difficulty during training is low and learning-phase performance is successful (but see Nakata & Suzuki, in this issue for the "judgement of learning" data), although successful learning-phase performance does not often yield the most successful *outcomes* (e.g., Bjork, 1994). There is thus a need for learners to successfully self-regulate and monitor the quality of their performance (e.g., desirably difficulty, deliberate practice) during training for better learning and retention (Bjork, Dunlosky,

FIGURE 1
A Proposed Theoretical Framework for Systematic and Deliberate L2 Practice [Color figure can be viewed at wileyonlinelibrary.com]



& Kornell, 2013; Soderstrom & Bjork, 2015). It is, however, nearly infeasible to define the absolute difficulty level that yields optimal practice condition, as L2 difficulty is best captured in relative terms. The crux of the current proposal illustrated in Figure 1, then, is that all three major difficulty-related factors (i.e., linguistic, ~~learner related~~, and ~~context related~~) need to be taken into account for creating the optimal levels of L2 practice (tailored systematic and deliberate practice). The current issue can be seen as our endeavor to better understand how one practice condition (context-related difficulty) contributes to the effectiveness of practice on the acquisition of a certain aspect of L2 (linguistic difficulty) for a certain learner (learner-related difficulty).

L2 PRACTICE AND DIFFICULTY IN THIS ISSUE

The series of experiments in this special issue examined the effects of different practice conditions, which presumably induced different levels of difficulty for learners. Table 1 summarizes the practice conditions that were explored in this issue. Nakata and Suzuki compared the three practice schedules (blocked practice, interleaved practice, and increasing practice [blocking followed by interleaving]). Nakata and Suzuki originally hypothesized, based on the desirable difficulty framework, that the increasing schedule may be the most effective because, by ensuring a continuous match between task difficulty and learner's proficiency level throughout training, it may induce the appropriate level of difficulty. Their hypothesis, however, was not supported. While the results of a 1-week delayed posttest showed the advantage of interleaved practice over blocked practice, increasing practice was not significantly better or worse than the other two practice schedules. Another recent study, in contrast, showed that an increasing learning schedule led to significantly better outcomes than either blocked or interleaved practice for L2 learners' spoken grammar knowledge (Suzuki & Sunada, 2019), which lends direct support to the desirable difficulty framework. In other words, gradually increasing the practice difficulty to match learners' skill levels helps induce the appropriate level of difficulty throughout training and facilitates learning. The attenuated benefit of the increasing schedule found in Nakata and Suzuki may be in part due to a relatively high level of prior knowledge of the target structures. As indicated by the average pretest score (69.76%) on the grammaticality judgement task (GJT), the participants had a rather high level of prior knowledge of the target structures. As a result, they were perhaps able to benefit from the most demanding training schedule (interleaving) from the outset. This suggests

TABLE 1
Practice Condition and L2 Difficulty

	L2 Difficulty During Practice	
Practice Condition	Low	High
Interleaved practice (Nakata and Suzuki)	Blocked	Interleaved
Distributed practice (Kasprowicz ~~and Marsden~~; Li and DeKeyser)	Shorter spacing	Longer spacing
Written/aural input (Kim and Godfroid)	Written (written and auditory)	Aural only
Retrieval practice (Strong and Boers)	Retrieval (studying followed by testing)	Trial and error (testing followed by studying)
Corrective feedback (Yilmaz and Granena)	Explicit	Implicit

that not only practice schedules but also individual factors such as prior knowledge play a role in determining what the optimal learning conditions are for a given individual.

Two studies in this issue compared the effects of short and long spacing. Generally speaking, longer-spaced learning imposes more difficulty compared to shorter-spaced learning (Suzuki, 2017; Toppino & Gerbier, 2014). Li and DeKeyser compared 1-day and 7-day practice intervals for multiple dimensions of Mandarin pronunciation skill acquisition. They found that longer spacing was more advantageous for the acquisition of simpler knowledge/skills (declarative knowledge), whereas shorter spacing led to the better retention of the more complex knowledge/skill (procedural knowledge). While Li and DeKeyser referred to the skill retention theory (Kim, Ritter, & Koubek, 2013), their findings can also be accounted for by the desirable difficulty framework. During practice, the memory decay between the training sessions was greater in the 7-day interval group than in the 1-day interval group. More demanding practice (7 days) is optimal for simple skills, and less demanding practice condition (1 day) is ideal for more complex skills. On the other hand, Kasprowicz, Marsden, and Sephton in this issue did not find any significant difference between the short (3.5-day) and long (7-day) practice intervals distributed over 3 weeks. During training, both short and long spacing intervals resulted in relatively high proportions of correct responses (>75%). When applying the desirable difficulty framework, the very high rates of learning-phase performance suggest that none of the two conditions introduced sufficient difficulty to enhance retention and therefore neither was optimal. Another potential factor responsible for their findings is individual differences in language-analytic ability (note that Li and DeKeyser controlled for individual difference factors in their statistical models).

A major source of difficulty in processing input is its modality. Kim and Godfroid convincingly showed the advantage of written modality over aural input for developing a stronger representation of L2 knowledge. This may also be explained by the desirable difficulty framework by stipulating that written input has eased the cognitive processing burden during incidental exposure. No data during practice was available in their study because their focus was on resulting (explicit and implicit) knowledge after the exposure. Based on their findings, however, Kim and Godfroid suggest that learners can benefit from receiving both aural and written input. Note that their study involved the learning of a semi-artificial language, and participants had no prior knowledge of the target structures. If the proficiency level of the learners were higher, aural-only practice might induce an appropriate level of difficulty and facilitate learning.

Earlier research has demonstrated positive effects of retrieval on learning. Strong and Boers in this issue compared the effects of retrieval (studying followed by testing) and trial-and-error conditions (testing followed by studying) on the learning of L2 phrasal verbs and found advantage of retrieval. From the desirable-difficulty vantage point, the trial-and-error condition might have been too demanding for learners because unsuccessful inferences generated during the learning-phase interfered with the correct forms. For instance, learners who erroneously produced *call out* instead of *call off* during the learning phase tended to produce *call out* on the posttest, despite

receiving corrective feedback. If the target items are less confusing (e.g., the distinction between *cancel* and *scream* may be less confusing than the distinction between *call off* and *call out*), however, more demanding practice conditions might lead to desirable difficulty and facilitate retention.

While comparing interactive practice with explicit or implicit corrective feedback, Yilmaz and Granena examined how a myriad of cognitive aptitudes predict the pretest–posttest improvement. One of the critical issues for corrective feedback research is the extent to which learners notice their error when they receive implicit feedback. The awareness during practice (e.g., feedback awareness, rule awareness) was examined by administering a ~~postexperiment~~ questionnaire. During practice, more learners in the explicit group became aware of the presence and quantity of feedback than in the implicit group. This suggests that by clearly indicating that learners made an error, explicit feedback provided scaffolding and presumably imposed a less challenging condition for noticing to take place, from the perspective of the desirable difficulty framework. As a corollary, their study found no effect of cognitive aptitudes on the learners' noticing in the explicit feedback condition. In contrast, in the implicit feedback condition (where more cognitive demands were imposed on learners for noticing the corrective feedback), attention control and phonemic coding ability predicted the degree of learners' noticing (rule awareness and feedback awareness). This pattern of findings is consistent with the general aptitude–treatment–interaction finding: the more the treatment puts the information-processing burden on the learner, the more important the aptitude is (DeKeyser, 2013). The learning condition that is made sufficiently easy through instruction (e.g., pointing out that the error was made) leads to no aptitude effects.

Last, an interesting pattern of findings was observed in two studies in this special issue. They pertain to the significant effects of aptitudes and its roles for allowing for transfer of learning. The transfer of learning discussed here concerns the extent to which gains from the treatment will be used in new contexts/tests (see also the Introduction article in this issue). In ~~Kasprowicz, Marsden, and Sephton's~~ study, after engaging in referential form-meaning mapping activities, participants completed a sentence–picture matching task and an auditory GJT as posttest measures. The aptitude (language-analytic ability) was significantly correlated only with the GJT score. Unlike the picture matching task, which was similar to the activities during the training, GJT required participants to demonstrate their knowledge in a novel format. Higher language-analytic ability might have helped learners to apply what they had learned from the training. A similar pattern of findings was also observed in Yilmaz and Granena's study. While the aptitudes were not a significant predictor of pretest–posttest performance change in an oral production task (which was similar to training as both required oral production), GJT performance gains were predicted by the aptitudes. These findings suggest that relatively near skill/knowledge transfer (e.g., transferring skill/knowledge learned from an oral production task to GJT) may be facilitated by higher aptitudes. The recurring theme of this issue, then, is that what constitutes the optimal, desirably difficult practice conditions depends on learner-related factors (e.g., prior knowledge and aptitudes) and target skills and items (e.g., declarative vs. procedural knowledge; phrasal verbs vs. single words). The interactions between these factors are at the heart of studying optimal L2 practice.

FUTURE DIRECTIONS OF L2 PRACTICE RESEARCH

There are several directions for future research on L2 practice in the overarching framework that we proposed in this coda article. Following the previous section's framework for understanding various practice-related factors, future research should investigate a variety of practice conditions and more structures with different degrees of linguistic difficulty, as well as expanding the scope of learner-related difficulty factors and delve into the synergy among them. Among the practice conditions, for instance, one of the most focused areas of L2 interaction research is corrective feedback. In their desirable difficulty framework, Schmidt and Bjork (1992) argued that providing delayed (summarized) feedback, rather than immediate (continuous) feedback during practice, may slow down initial learning but result in better retention in the long run. The timing of corrective feedback is still underexplored in L2 research, but some interesting lines of work have been emerging (Li, Zhu, & Ellis, 2016; Quinn & Nakata, 2017). This is an interesting, developing area of corrective feedback research in the scope of L2 practice and desirable difficulty framework. If we further extend the applicability of desirable difficulty framework, it may be worth exploring how this framework relates to and informs task complexity research, especially in the context of task-based language teaching (TBLT; e.g.,

Jackson & Suethanapornkul, 2013). In task complexity research, researchers typically aim to identify task complexity levels that induce good task performance (e.g., higher linguistic complexity, accuracy, fluency). Little attention, however, has been paid to how challenging tasks, perhaps even with degraded performance, ~~results~~ in the learning and retention of skills and knowledge through TBLT instruction. The current desirable difficulty framework has the potential to reach a large area of L2 research.

In addition to examining the main effect of practice condition, we can elucidate different learning processes during L2 practice in quest of complex interactions between multiple variables. The empirical studies in this special issue explored such interactions. Yet, we emphasize that more efforts should be made to investigate how the effects of different practice conditions are moderated by target knowledge and skills within the same study. This area merits attention for future research because the effects of practice conditions may differ depending on types of knowledge (e.g., declarative, procedural, automatized knowledge) or linguistic domains examined (e.g., grammar, pronunciation, lexis, pragmatics). Even within the same domain the putative psychological processes may vary significantly, which warrants more systematic investigations into the interaction between practice condition and knowledge.

As a case in point, Li and DeKeyser's study revealed that declarative and procedural knowledge of Mandarin tones are susceptible to learning schedules to different degrees. The findings suggest that the practice condition may need to be determined depending on target skills. It may also be useful to examine whether different linguistic structures benefit differently from practice. Nakata and Suzuki (2019), for instance, pursued this line of work and examined whether spacing differentially affects the learning of semantically related and unrelated lexical sets, and found that semantically unrelated sets benefit more from spacing. Another recent study (Suzuki & Sunada, 2019) also examined whether blocking, interleaving, and increasing schedules differently affect the learning of comprehension and production skills. They found the advantage of increasing schedule over the blocking and interleaving schedules for measures of production accuracy and comprehension speed (reaction time), but not for measures of production speed or comprehension accuracy.

In addition to research on interactions between multiple variables, more detailed analysis of feature-related and learner-related variables themselves are necessary. For instance, working memory, which is one type of learner-related variables (aptitude), is a multifaceted, complex system. Research on working memory thus should specify which component or function (e.g., attention control, updating, inhibition) is measured. The more fine-tuned measurement is crucial for pinning down exactly why certain variables interact with treatment. On the other hand, combining multiple components (for instance, see Kasprowicz ~~& Marsden,~~ in this issue, where language-analytic ability was used to subsume the inductive and deductive learning ability and metalinguistic knowledge about L1) may serve practical purposes and be more useful for practitioners. In order to pursue rigorous investigations into the role of L2 practice, fundamental understanding of the nature of L2 cognitive aptitude (e.g., the assumptions of stability and lack of trainability of aptitudes) will also be crucial (see Hayashi, in this issue).

CONCLUSIONS

This special issue collectively revealed a number of factors that influence the effects of L2 practice and contains useful pedagogical implications. However, we would like to close this issue with a word of caution. As DeKeyser (2007) remarked in his book on L2 practice, "there is the tendency to overgeneralize from research on the acquisition of certain elements of a certain language by certain kinds of learners without taking into account the big differences in psycholinguistic difficulty that characterize the acquisition of vocabulary versus grammar versus formulas, or of rules versus items versus prototypes" (p. 289). For L2 teachers, who are teaching in classrooms where all those complex factors are intertwined, their macro and micro teaching decisions (e.g., syllabus planning, choosing practice conditions, or feedback options) should be informed by empirically supported evidence in principle; however, in practice, teachers need to ~~exercise caution~~ when applying research findings to their context (Rankin & Becker, 2006; Spada, 2019). In turn, in the face of a number of factors influencing L2 learning and outcomes (and the dynamic interactions between these factors), researchers need to carefully control for any extraneous variables and rigorously design empirical research that can shed light on the target areas of investigations. More efforts and systematic research are necessary to further advance our understanding of L2 practice.

REFERENCES

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: The MIT Press.

Bjork, R. A. (2018). Being suspicious of the sense of ease and undeterred by the sense of difficulty: Looking back at Schmidt and Bjork (1992). *Perspectives on Psychological Science, 13*, 146–148

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444.

DeKeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language Learning, 55*, 1–25

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.

DeKeyser, R. M. (2013). Aptitude. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 27–31). New York: Routledge.

DeKeyser, R. M. (2016). Of moving targets and chameleons: Why the concept of difficulty is so hard to pin down. *Studies in Second Language Acquisition, 38*, 353–363.

Q4

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). Cambridge: Cambridge University Press.

Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38*, 163–175.

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63*, 330–367.

Kim, J. W., Ritter, F. E., & Koubek, R. J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science, 14*, 22–37.

Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *Modern Language Journal, 100*, 276–295.

Nakata, T., & Suzuki, Y. (2019). Effects of massing and spacing on the learning of semantically related and unrelated words. *Studies in Second Language Acquisition, 41*, 287–311.

Quinn, P. G., & Nakata, T. (2017). The timing of oral corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language teaching and learning* (pp. 35–47). New York: Routledge.

Rankin, J., & Becker, F. (2006). Does reading the research make a difference? A case study of teacher growth in FL German. *Modern Language Journal, 90*, 353–372.

Révész, A., & Gurzynski–Weiss, L. (2016). Teachers' perspectives on second language task difficulty: Insights from think-alouds and eye tracking. *Annual Review of Applied Linguistics, 36*, 182–204.

Sato, M., & Loewen, S. (2018). Do teachers care about research? The research–pedagogy dialogue. *ELT Journal, 73*, 1–10.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.

Schmidtke, J. (2018). Pupillometry in linguistic research: An introduction and review for second language researchers. *Studies in Second Language Acquisition, 40*, 529–549.

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance an integrative review. *Perspectives on Psychological Science, 10*, 176–199.

Spada, N. (2019). Balancing methodological rigor and pedagogical relevance. In R. M. DeKeyser & P. G. Botana (Eds.), *Doing SLA research with implications for the classroom: Reconciling methodological demands and pedagogical applicability* (pp. 201–215). Philadelphia/Amsterdam: John Benjamins Publishing Company.

Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning, 67*, 512–545.

Suzuki, Y., & Sunada, M. (2019). Dynamic interplay between practice type and practice schedule in a second language: The potential and limits of skill transfer and practice schedule. *Studies in Second Language Acquisition.*

Q5

Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. *The Psychology of Learning & Motivation, 60*, 113–189.