

In Defense of Tasks and TBLT: Nonissues and Real Issues

MICHAEL H. LONG

University of Maryland

ABSTRACT

The first aim of this article, addressed in section 1, is to define what is meant, and not meant, by *task* and *task-based language teaching* (TBLT). The second is to summarize and evaluate 14 criticisms that have been made of both. Section 2 responds to five alleged problems with TBLT's psycholinguistic rationale, section 3 to six at the classroom level, and section 4 to three claimed problems with implementing TBLT in specific contexts. A few of the criticisms touch on important matters, but most, I will suggest, are *nonissues*. The third aim of the article is to identify some genuine problems in need of resolution—*real issues*—and briefly to illustrate research programs under way to address them.

I. WHAT IS MEANT, AND NOT MEANT, BY TASK AND TASK-BASED LANGUAGE TEACHING?

Task and *task-based language teaching* (TBLT) mean different things to different people. In some commercial textbooks and in some writing on pedagogy, *task* simply refers to a traditional, linguistically focused exercise or activity related to keep up with language teaching (LT) fashion. In other cases, especially among some critics of the use of tasks, it can mean anything the writer chooses. Bruton (2002a, p. 282), for instance, listed nine task types: “problem-solving, decision-making, spontaneous role-playing, etc.”; “information/opinion gap resolution”; “cued prompted interaction”; “question-answer exchanges”; “prepared role plays”; “focused receptive language (+/- itemized)”; “focused written language (+/- itemized) [reproduction]”; “understanding”; and “written expression.” Based on such a miscellany of activities, skills, modalities, pedagogic procedures, language, conversational moves, and cognitive processes, it would be impossible to define *task* or *task type* or know what would and would not qualify as examples of either.

A third use of *task* is to denote an activity that is communicative at some level, but whose purpose, overt or covert, is to practice specific linguistic items. Students may read a story about the chain of events that resulted in a traffic accident, and then answer comprehension questions on the text, as a way of improving their control of past time reference. Or they may role-play imaginary job interviews, not because they are preparing for real job interviews conducted in the second language (L2), but in order to improve their production and comprehension of questions. A more constrained version is the consciousness-raising task:

A pedagogic activity where the learners are provided with L2 data in some form and required to perform some operation on or with it, the purpose of which is to arrive at an explicit understanding of some linguistic property or properties of the target language. (R. Ellis, 1997, p. 160)

While varying in the degree of linguistic control exercised by the materials writer and in the degree of attention to language as object required of the learner, all such tasks in this third sense, aptly named “structure-trapping tasks” (Skehan, 1998, pp. 122–123), are components in the delivery of a traditional linguistic syllabus of some kind (grammatical, notional-functional, lexical, hybrid, etc.). In other words, they are part of task-*supported* LT. The linguistic syllabus may be covert, concealed in the communicative tasks, or alternatively, as advocated by R. Ellis (1994, 2003, 2009), one track in a dual structural and task hybrid:

Task-based teaching need not be seen as an alternative to more traditional, form-focused approaches but can be used alongside them ... (Long and Skehan view traditional structural teaching as theoretically indefensible while I see it as complementary to TBLT). (R. Ellis, 2009, pp. 221, 225)

In what follows, as developed in my own work over the past 30 years, *task* means something else again. Tasks in this fourth sense are the real-world communicative uses to which learners will put the L2 beyond the classroom—the things they will *do* in and through the L2—and the task syllabus stands alone, not as one strand in a hybrid of some kind. The real-world tasks may be required for academic purposes, for example, locating a journal in a university library, writing a lab report, or attending a graduate-level economics lecture. They may be for vocational training purposes, for example, in a noisy restaurant kitchen, preparing kitchen utensils and cooking ingredients at the direction of a master chef, attending a class for trainee computer technicians, or following written directions for the use of specialized automobile repair equipment. They may be for occupational purposes, either in the home country, for example, while employed in the tourist industry, welcoming and checking in hotel guests, renting surf boards, or leading a guided tour, or overseas, for example, while stationed at an embassy, interviewing visa applicants, issuing instructions to security personnel, or delivering an after-dinner speech. Whatever their main purpose and whether short or long term (including immigration to a new country), overseas stays will usually also involve a variety of “social survival” tasks, such as following street directions, using public transport, opening a bank account, renting an apartment, taking a driver’s test, visiting a doctor, or registering a child for school.

Identified in the first stage of a task-based needs analysis (NA; Long, 2005, 2015, pp. 85–168; Serafini, Lake, & Long, 2015), these real-world communicative activities are *target tasks* for the learners concerned. Target tasks for two or more individuals or groups of learners may obviously differ greatly, which underscores the inadequacy of off-the-shelf commercial LT materials, most of which are written, on the basis of textbook writers’ intuitions, for all students and for

no students in particular. In the second stage of the NA, the needs analyst gathers samples of spoken or written language use by native speakers engaged in the most critical and/or frequent of the target tasks. Modified, elaborated (not linguistically simplified) versions of the samples subsequently become part of the task-based materials—the *pedagogic tasks*—produced for a course, and they constitute *the major source of new language* for the learners, new language that is relevant for their target discourse domains.

In the interests of economy of time and effort in both teaching and testing, groups of target tasks may be classified at a more abstract level into *target task types*. For example, filling out application forms for a driver's license, a library card, a job vacancy, a bank account, a mortgage, or a club membership (six target tasks) are all instances of filling out an application form (a target task type). Testable predictions are that (a) the ability to complete, say, two of the application forms will generalize to other application forms not specifically taught, and that (b) if students pass a test of their ability to complete one or both of those first two forms, the result can serve as a proxy for their untested ability with the other kinds. The generalizability of task-based abilities is an important research issue for both task-based teaching and testing—a *real issue*—to which we will return.

Instructional materials for TBLT take the form of *pedagogic tasks*—initially simple, progressively more complex approximations to the original target tasks. Multiple series of pedagogic tasks are sequenced for classroom use according to their intrinsic complexity—task complexity, not linguistic complexity (Long, 2015, pp. 223–247; Robinson, 2009, 2011). Each series culminates in the full target task or a simulation thereof, which serves as the exit task for a module. Collectively, the series of pedagogic tasks form a *task syllabus*. There is no linguistic syllabus, overt or covert, other than what Corder (1967) termed the internal “learner syllabus.”

The task syllabus is delivered in conformity with 10 (putatively universal) *methodological principles* (MPs) for LT:

- MP1: Use task, not text, as the unit of analysis
- MP2: Promote learning by doing
- MP3: Elaborate input
- MP4: Provide rich input
- MP5: Encourage inductive “chunk” learning
- MP6: Focus on form
- MP7: Provide negative feedback
- MP8: Respect learner syllabi and developmental processes
- MP9: Promote cooperative collaborative learning
- MP10: Individualize instruction

The MPs are motivated by what second language acquisition (SLA) research has shown about how children and adults learn L2s successfully (Long, 2009, 2015, pp. 300–328) and independently, by principles from the philosophy of education (Long, 2015, pp. 63–83). The MPs are realized at the local classroom level by pedagogic procedures (PPs). Selection of appropriate PPs from the many

available in each case is best left to the teacher, who is usually the expert on local circumstances, assuming that he or she is well trained and experienced. There are no universal or “best” PPs. Rather, choices should vary systematically to cater to individual learner differences (age, level of first or second language [L1 or L2] literacy, working memory, aptitudes for implicit or explicit learning, etc.), type of linguistic feature (salient or nonsalient, marked or unmarked, fragile or robust, etc.), and so on. To deal with persistent errors with a nonsalient target language feature, such as intrasentential clitics, a teacher of literate, analytically oriented adults might choose a PP for providing explicit negative feedback (MP7), for example, a simple pedagogic rule of thumb, but for a salient feature, such as adverb placement, one for providing implicit negative feedback, for example, recasts or clarification requests, or on indirect negative evidence.

Finally, student abilities are assessed using *task-based, criterion-referenced performance tests* (Long, 2015, pp. 329–341; Norris, 2009; Norris, Brown, Hudson, & Bonk, 2002; Van Gorp & Deygers, 2013). This is a straightforward process in the many cases, such as attending an economics lecture, conducting a visa interview, setting up laboratory equipment, or following street directions, where abilities to perform target tasks can be assessed directly and discretely, with unambiguous behavioral outcomes as achievement measures. Task-based assessment can be problematic, however, and involve choices among less desirable options, including whether or not to add a linguistic caboose when simply getting a message across may not suffice for acceptable task performance. A diplomat’s after-dinner speech in Japanese, for example, needs to be delivered in the correct register, with use of appropriate formulae and grammatical particles to mark politeness (Hillman, 2015). Task-based assessment is another real issue in TBLT, to which we will return.

The whole approach—from NA, through syllabus design, materials writing, and choice of MPs and PPs, to assessment—is task-based throughout and constitutes genuine *task-based LT* (TBLT). A preliminary sketch was first presented in a plenary address to the 1983 Georgetown Round Table, subsequently published as Long (1985). The proposal was largely ignored until reports began to appear in popular journals, such as *TESOL Quarterly* (e.g., Long & Crookes, 1992). Since then, tasks and TBLT have been subjected to a number of critiques—some rational, some less so. The critics have included Widdowson (1993, 2003), R. Ellis (1994), Willis (1994), Wilkins (1994), Seedhouse (1999, 2005), Klapper (2003), Bruton (2002a, 2002b, 2005), and Swan (2005, 2011). If TBLT is discredited to an author’s satisfaction, the conclusion drawn is often that traditional linguistically focused approaches and present–practice–produce (PPP) are justified by default—a non sequitur. Swan’s 2005 polemic in *Applied Linguistics*, in particular, is widely cited and was republished in 2011. In his introduction to the reprint, Swan (2011, p. 91) described the piece as “a full-scale criticism of hard-core task-based instruction, and of the hypotheses which are held to justify it.” Many advocates of one form or another of TBLT are indicted by the critics. I cannot speak for them, of course, so will be “correcting the record” mostly in terms of TBLT as outlined above and presented in detail in Long (2015).

2. ALLEGED PROBLEMS WITH TBLT'S PSYCHOLINGUISTIC RATIONALE

2.1. Research Findings on Interlanguage Development and Teachability Are Insufficient to Discredit the Grammatical Syllabus and PPP

Traditional LT approaches using materials embodying a preset linguistic syllabus and delivered via various widely used incarnations of present–practice–produce (PPP)/focus on forms methodology are inconsistent with research findings on interlanguage development and teachability. Research has shown that L2 grammatical development does not follow an order externally imposed by a teacher or textbook. Instead, with occasional L1 influence on the duration of stages and/or the addition of substages, learners traverse developmental sequences, often producing utterances reflecting learner-created, non-target-like rules in the process. Part of the definition of “developmental sequence” in SLA is that it consists of a *fixed* series of stages, none of which can be omitted. The sequences are usually held to reflect underlying hierarchically related processing constraints (Johnston, 1985, 1997; Lenzing, 2015; Pienemann & Kessler, 2012). Evidence for the existence of developmental sequences is long-standing and robust (for review, see Ortega, 2009). They have been documented in a variety of typologically different L2s (e.g., Bettoni & Di Biase, 2015; Bonilla, 2015; Johnston, 1997; Pienemann, 2005), as has the failure of instruction to alter them (R. Ellis, 1989; Håkansson & Norrby, 2010; Mackey, 1999; Pienemann, 1984).

Contrary to what defenders of the structural syllabus and PPP—like Bruton (2002a, 2002b), Widdowson (2003, pp. 111–133), and Swan (2005)—and most commercial textbook writers assume, and as is only to be expected if developmental sequences are real, 40 years of instructed SLA (ISLA) research and classroom studies have demonstrated that teachers cannot teach whatever they want, whenever they want, if language learning is their goal. Stages in sequences cannot be skipped by presenting learners with full target-like grammatical structures and drilling them until they are “automatized.” In Pienemann’s terms, processability determines learnability, and learnability determines teachability (Pienemann, 1984). Students play a decisive role in the language-learning process; their readiness to learn is not determined by the day of the week or the page in a textbook. As Skehan (2002, p. 294) noted, PPP and the skill-builders’ view is that language acquisition is teacher-driven, whereas, with support from decades of SLA research, TBLT views language acquisition as learner-driven.

Some critics, such as Wilkins (1994), recognized TBLT’s consistency with findings on developmental sequences and other processes in interlanguage development, as well as the problems such findings pose for structural syllabuses and accompanying methodology. However, despite this, they claimed that the research in support of TBLT to date was insufficient and defended continued use of a synthetic linguistic syllabus and linguistically simplified materials. If their focus was on task-*supported* LT, and therefore, on pedagogic tasks unmotivated

by a NA and unrelated to learners' real-world target tasks (see, e.g., Widdowson, 2003, pp. 111–133), they claimed that tasks are defined too vaguely, and that their classroom use is inauthentic. In fact, however, work on approximations to real-world tasks can be very realistic in genuine *task-based* LT classrooms or other instructional environments. Widdowson asserted that “authenticity” is inevitably a chimera because “the classroom contexts within which language has usually to be learnt are totally different from those within which the language is used” (Widdowson, 2003, p. 112). That may be true of PPP classrooms, but it is not true of TBLT. Wilkins maintained that structurally graded, inauthentic materials are not inevitably bad, and noted that “no evidence is cited [by Long & Crookes, 1992] that learners cannot learn language from what the authors would regard as non-authentic material” (Wilkins, 1994, p. 51). Instead of (a) a straw man charge about “authentic” materials, (b) questioning whether the considerable volume of published research evidence for TBLT (see below) is sufficient, and (c) requiring the logical impossibility of proving a negative (here, that something cannot occur), a more constructive approach for defenders of the status quo would be to provide empirical support for their position.¹

Not all critics are intent on defending the grammatical syllabus and PPP. Some are familiar with the SLA literature and sympathetic to the idea of building a syllabus around tasks rather than linguistic structures, but with caveats. Klapper (2003), for example, recognized the inability of instruction to change the route of L2 development (confirmed in Klapper & Rees, 2003), that “the linearity of PPP methodology and the incremental syllabus is difficult to defend in light of insights from SLA which indicate that language learning is a far more complex affair” (p. 37), and that “what we know about language learning strongly suggests the primacy of meaning negotiation supported by a focus on form, as proposed by TBLT” (p. 40). He certainly did not reject TBLT by claiming that the research behind it is “insufficient.” Rather, he argued for a task-based methodology that *incorporates* the skill-building model underlying PPP, whereby declarative knowledge is first proceduralized, and then automatized:

In cognitive theory, acquiring a second language is the same as learning any complex skill: a range of sub-skills must be practiced in “controlled” processing until they can be integrated into automatic or fluent performance—in much the same way as we learn to play the piano, drive a car or juggle. (Klapper, 2003, p. 39)

Many cognitive theorists (e.g., Rastelli, 2014) would deny that learning a language is like learning to play an instrument, drive, or juggle. They would also reject Klapper's (2003) second controversial assertion, reflecting a strong interface position, that “when procedural knowledge is ‘automatized’ or fine-tuned, it is implicit” (p. 39). Nevertheless, Klapper's position is more considered and more interesting than the wholesale rejection of TBLT espoused by some of the early critics.

Klapper (2003) eventually proposed what he described as a weak version of TBLT, a “hybrid model,” “which accepts the primacy of the communicative focus

but reinstates declarative knowledge and practice at the appropriate point in the task cycle” (p. 40). He claims that this is “the most effective way to make forms salient to students and thereby to speed up the acquisitional process” (p. 40). He was careful, however, to distinguish his proposal from conventional PPP:

It does not assume, like traditional language syllabuses that linguistic forms can be learnt in a pre-ordained order over a short period of time. In such a model, there would be no structural syllabus independent of the task syllabus, rather the forms to be practiced would arise from the task context; but planning would be required to ensure grammatical structures were regularly revisited and recycled, especially those that were poorly represented in classroom input and task instruction. (Klapper, 2003, p. 40)

At first sight, this might appear similar to arguments by TBLT advocates for abandoning the structural syllabus and PPP, and their (partial) replacement by *focus on form*, which is also reactive and often triggered by task work, not the other way around. The difference, however, and it is a major difference, is that there is no assumption in genuine TBLT that performance of individual grammatical structures can be developed to native-like levels rooted in a separate explicit knowledge system via the massive practice required for automatization, or that if accomplished, the new underlying knowledge will morph into the separate implicit system. In other words, not just the methodology but the underlying psycholinguistic rationale for TBLT is quite different. Klapper’s (2003) proposal, like that of R. Ellis (1994, 2003, 2009), is for a hybrid model, one that may yet turn out to be correct but that some would see as fatally flawed because it seeks to meld two oppositional psycholinguistic positions. In my view, if there is a place for skill acquisition theory in TBLT, it is for the use of task repetition to improve task performance, not performance of individual grammatical structures before their time.

Unlike the Klapper and the Ellis studies, Swan (2005) questioned not just the implications for LT drawn from them, but also the underlying SLA theory and research findings themselves. To defend the intuition-based sequences in grammatical syllabuses favored by commercial textbook authors (of whom he is one of the best-selling and most respected), he starts by questioning the existence and immutability of developmental sequences, quoting DeKeyser for the purpose:

As far as developmental factors are concerned, evidence for inbuilt acquisition sequences currently lacks generality ... if such sequences exist, there seems no good evidence that they cannot be interfered with (DeKeyser, 1998, 57–8). (Swan, 2005, p. 379)

He also dismissed Pienemann’s teachability hypothesis, due to “the lack of wide-ranging empirical evidence” (p. 381). Contrary to what the quote implies, however, DeKeyser (1998) explicitly recognized on the very same page from which the quote was drawn both “the existence of acquisition orders, which is well

established in the literature” (p. 57) and that instruction “does not have a substantial impact on route (order) of acquisition” (p. 57). He does not deny the existence of studies showing the failure of instruction to alter the orders. Rather, DeKeyser suggests that, absent adequate information about the form the teaching took,

it is safe to assume ... that none received instruction along the lines of what skill acquisition theory would seem to imply: explicit teaching of grammar, followed by FonF [focus on form] activities to develop declarative knowledge, and then gradually less focused communicative exercises to foster proceduralization and automatization. (DeKeyser, 1998, p. 58)

Neither DeKeyser’s conjecture (and it was just a conjecture) concerning a possible *cause* of the failure of instruction to alter the orders nor Swan’s dismissal of the research findings alters the fact that studies that have compared classroom and naturalistic learners have found acquisition sequences to be broadly the same. There may currently be “no good evidence” in Swan’s view that they cannot be “interfered with” (to refute this would, again, require proving a negative), but more to the point, there is no evidence that they can be.

Having discounted decades of empirical findings, Swan reassured readers that “traditional structure grading is informed by pedagogic experience and expertise” (Swan, 2005, p. 394). Their source is unspecified. In fact, he repeatedly settled arguments to his satisfaction by appeals to what is “sensible” (to whom is left unstated), to “experience,” which differs from one individual to the next, and to “common sense,” which would obviate the need for debate if it really were common. Such an attitude to research is distressingly common among TBLT’s loudest critics, who frequently demand more studies, while offering none of their own. More than 25 years ago, Nunan was able to write:

One of the strengths of task-based language teaching is that the conceptual basis is supported by a strong empirical tradition. This distinguishes it from most methods and approaches to pedagogy, which are relatively data-free.” (Nunan, 1991, p. 283)

Twenty-five years later, the research basis for TBLT is many, many times greater. As an experienced classroom teacher and teacher educator myself (and one, incidentally, who learned a great deal working closely with Michael Swan early in my teaching career), I am certainly aware of, and open to, the insights that come from practice. However, classroom experiences vary across individuals and contexts and over time, and are usually subjective and impressionistic. They can motivate research, but are no substitute.

2.2. Incidental Focus on Form Is the Basis for TBLT, and It Does Not Work

Operationalizing and demonstrating incidental (and implicit) language learning under laboratory conditions is notoriously difficult, and harder still in the wild. However, even most skeptics recognize that most lexical items and collocations

are learned that way—by adults, not just by children—and some scholars (e.g., N. Ellis, 2005; Ellis & Wulff, 2015; Rastelli, 2014) maintain that implicit, statistical, usage-based learning underlies most grammatical learning as well. In an assertion that assumes a radically different view of how acquisition occurs, that is, with conscious awareness at the moment of learning, Swan wrote that TBLT is often justified by the claim that “‘linguistic regularities’ are acquired through ‘noticing,’ during communicative activity, and should therefore be addressed primarily by incidental ‘focus on form’ during task performance,” adding that “there is no compelling evidence for the validity of the model” (Swan, 2005, p. 376).

There are two problems with these assertions. First, the rationale for TBLT is far broader, involving, among other things, (a) recognition of the immense diversity of learner populations and learner needs and, hence, the inadequacy of one-size-fits-all courses and textbooks; (b) a set of widely attested SLA research findings, explanations for them, and implications derived from them; and (c) consistency with a set of nine principles from the philosophy of education (see Long, 2015, pp. 3–15). Focus on form (MP6) is but one of the 10 MPs for *implementing* TBLT by those already convinced by the rationale, a way of drawing learners’ attention to grammar (in the broadest sense) in harmony with the internal “learner syllabus.” The cognitive-interactionist theory of ISLA underpinning TBLT embraces not just incidental focus on form, but (among other processes and abilities) what theory and research findings in SLA show to be a necessary, principled, symbiotic relationship between *incidental and intentional learning* (Long, 2015, pp. 30–62).

Second, a considerable body of research has demonstrated that incidental focus on form works. By 2000, five years before Swan’s (2005) polemic first appeared, Norris and Ortega (2000) published a widely cited statistical meta-analysis of 49 studies of the relative effectiveness of explicit and implicit instruction and (a related but different issue) of focus on forms and focus on form. Swan (2005, p. 379) cited the review correctly as having found “little advantage for on-line over off-line instruction,” seemingly without realizing how this statement contradicts his original assertion as to the inadequacy of incidental learning. Eleven of the original studies, plus 34 new ones reported in the following decade prompted a second statistical meta-analysis (Goo, Granena, Novella, & Yilmaz, 2015), which confirmed the earlier results. In addition, a large body of research—well over 60 studies—on recasts, another device for achieving incidental focus on form, has shown that the implicit negative feedback they provide leads to substantial learning gains in grammar and vocabulary (for an excellent narrative review, see Goo & Mackey, 2013; and for two statistical meta-analyses, Li, 2010; Mackey & Goo, 2007).

In other words, Swan’s (2005) claim is wrong. Focus on form works, and the gains it produces in the learning of grammar and vocabulary are not statistically significantly smaller than those achieved by traditional *focus on forms*, which Swan favors. The gains are achieved, moreover, without the latter’s side effects: among other problems, irrelevance to student needs, at best only accidental developmental appropriateness, exposure to non-native-like L2 models, little opportunity for

communicative practice, and in all too many classrooms, mind-numbing monotony. The results from a combined total of well over 140 empirical studies and several statistical meta-analyses of the issue to date in these two areas alone should suffice to meet critics' demand for "compelling evidence." Meanwhile, nothing approaching the quality and scope of this body of empirical work exists in support of the favored PPP model.

The research findings notwithstanding, several critics assert that *focus on forms*, not focus on form, is optimal, with TBLT perhaps useful at advanced levels, when students already know the language and just require more practice. But even DeKeyser, a leading proponent of skill-based language learning, recognized that although the four conditions are not requirements:

Skill Acquisition Theory is most easily applicable to what happens in (a) high-aptitude adult learners engaged in (b) the learning of simple structures at (c) fairly early stages of learning in (d) instructional contexts. (DeKeyser, 2015, p. 101)

If true, the vast majority of the world's language learners—those of average or lower aptitude, those of intermediate or higher proficiency, those not studying one of the minority of the world's languages for which instruction is available and fortunate enough to have access to that instruction, or those learning naturalistically—must be acquiring their L2 some other way.

2.3. Acquisition in TBLT Only Takes Place Online During Communicative Language Use

Swan cited Doughty (2001, pp. 226–227) as the justification for a related criticism that, in TBLT, "acquisition only takes place on-line during communication" (Swan, 2005, p. 379). But Doughty claimed no such thing; nor is such a belief part of the psycholinguistic rationale for TBLT. Doughty (2001) provided a review of the psycholinguistic underpinnings of focus on form. In the passage Swan referred to, Doughty was discussing theory and research findings on working memory and short-term memory pertaining to the estimated period of perhaps 40 seconds following production during which speakers' (and by inference, L2 learners') selective attention span might be expected to be capable of utilizing online implicit negative feedback in the form of recasts for form-function mapping. This is the "window of opportunity" in which "cognitive comparison" (Saxton, 1997) of spoken output and input containing negative feedback might successfully occur in real time, offering L2 learners an opportunity to "notice the gap" between what they said and contingent utterances from more competent interlocutors.

Nowhere did Doughty suggest that learning from other sources or by alternative means cannot occur, or occur over a longer period than 40 seconds. As Swan suggested, it would be patently absurd to suppose that explicit metalinguistic knowledge provided at Time 1 could never be used by learners at Time 2—to use Swan's example, that conscious knowledge that the particle *ka* can be attached to

statements to form questions in Japanese could not be used to form questions that way when someone starts learning Japanese a year later—without first acquiring the rule “naturalistically in the prescribed way” (Swan, 2005, p. 379).

The switch from incidental (what Swan calls “online”) to intentional learning, as facilitated by focus on form, recasts and other devices, has always been part of the interaction hypothesis (Long, 1996) and is a key component of the psycholinguistic rationale for TBLT (Long, 2015, pp. 49–57, 316–321). Leaving aside the rather unusual nature of the long-term benefit postulated in Swan’s imagined example, and while usually of very limited value beyond the simplest rules accessed during offline performance, provision of explicit metalinguistic information is one of many possible PPs by means of which MP6 (focus on form) can be realized in the short term. A key characteristic of focus on form, negative feedback, and expansions of various kinds, however, is that they are *reactive*, supplied in harmony with a learner’s current developmental readiness to learn, not that they must be or cannot be provided in a particular way, as Swan’s complaint implies. The value of explicit knowledge for language learning was irrelevant for the issue Doughty was discussing and Swan’s charge is therefore a *nonissue* both there and where TBLT is concerned.

2.4. L2 Knowledge Acquired Incidentally Does Not Last

A focus of considerable research interest in SLA is the relative durability of L2 knowledge acquired incidentally or intentionally. Swan claims that few studies have demonstrated lasting retention and availability for spontaneous use of forms acquired incidentally (Swan, 2005, p. 379). Again, empirical findings tell a different story. Results show that explicit procedures often do as well as, and sometimes slightly better than, incidental, online focus on form, but as noted above, usually only with simple linguistic targets, only on immediate posttests, and only using discrete-point measures, and that improvements achieved that way tend to deteriorate over time (Doughty, 2003). In contrast, while the jury is still out, results to date suggest that incidentally and/or implicitly learned L2 knowledge is more durable and tends to increase over time (probably because of the initial greater depth of processing involved). In a statistical meta-analysis of 33 studies, Li (2010) found a medium overall effect for oral corrective feedback, that the effect was maintained over time, and that while the immediate and short-term effect of explicit feedback was greater, the longer-term effect size for recasts (a prime example of “incidental, on-line focus on form”) was slightly larger than the short-term effect, more effective than explicit feedback on delayed posttests, and more enduring, even increasing over time (Li, 2010, p. 343).

2.5. TBLT Relies on the Noticing Hypothesis, With Which There Are Problems

Another concern for Swan is TBLT’s alleged reliance on Schmidt’s noticing hypothesis. Swan claimed (2005, pp. 379–380) that limitations of the noticing hypothesis constitute a problem for TBLT, notably the fact that many items are acquired

without conscious awareness by both native and nonnative speakers. This is good to hear, given his previously expressed skepticism about incidental learning: If Swan were to be believed, TBLT assumes that all learning is online, that is, incidental, and now, that all learning requires conscious awareness. Neither is correct, of course, and TBLT claims neither.

Like others before him (e.g., Truscott, 1998), Swan (2005) pointed out that knowledge of some constraints on rule application cannot be the product of noticing because there are no examples to notice. For instance, some verbs (give, offer, promise, etc.) can figure in the double-object construction (e.g., The Minister promised Eubanks-Smythe a knighthood), but others (donate, present, explain, etc.) cannot (*Eubanks-Smythe donated the Party a million pounds), requiring a prepositional phrase, instead (Eubanks-Smythe donated a million pounds to the Party). Schmidt (2010, and elsewhere) readily acknowledged that learning in such cases indicates alternative learning mechanisms at work (theoretical accounts run the gamut from innate linguistic knowledge to statistical learning), but maintained that the noticing hypothesis still applies to the great majority of language learning. The noticing hypothesis holds that learning requires attention and awareness, but not intention or understanding. Understanding is facilitative, but not required. He has never denied the possibility of incidental learning, for example, of vocabulary and collocations, but remains deeply skeptical of all claims of learning without awareness (a position with which I disagree). However, while instances have to be attended to and noticed, he recognizes that *generalizations* from those instances can be either explicit or implicit, implicit learning being a basic human learning mechanism that automatically detects regularities across instances. Schmidt's position can now be summarized (briefly, for lack of space) not as saying noticing is required (noticing of surface features, at least) for all aspects of language, but as "more noticing means more acquisition," and "more attention and more awareness, means more acquisition" (R. W. Schmidt, personal communication, July 26, 2015).

Absence of items in the input to notice is a problem for TBLT, Swan asserted, because the noticing hypothesis provides "a psycholinguistic justification for the observationally motivated need for focus on form" (Swan, 2005, p. 380). In fact, there is no assumption in TBLT that conscious awareness is required for acquisition, which would anyway only be an issue with "hidden" constraints of the kind illustrated. Noticing can be facilitative, especially in speeding up learning of perceptually nonsalient features or in establishing a new representation in long-term memory, after which subsequent examples of the same feature in the input can be parsed implicitly (N. Ellis, 2005), but it is neither a necessary component nor a necessary result of focus on form. What TBLT relies on is students' capacity for both *incidental and intentional learning*, and teachers and materials writers' use of both. Teachers or task designers (materials writers) may find it advantageous deliberately to direct students' attention to language as object in context during meaning-focused, receptive, or productive L2 use, but conscious awareness at the moment the learner detects the targeted item is not a requirement for uptake.

3. ALLEGED PROBLEMS WITH TBLT AT THE CLASSROOM LEVEL

3.1. TBLT Neglects Grammar

A number of critics (e.g., Swan, 2005; Widdowson, 2003) have alleged that TBLT pays insufficient attention to the teaching of grammar. The source of the confusion seems to be the belief that if the purpose of sequences of pedagogic tasks is gradually to develop learners' ability to perform target tasks (which it is), not to practice particular grammatical structures, then grammar is ignored. This is simply a misunderstanding. It is true that so-called unfocused pedagogic tasks in genuine TBLT are not designed to teach particular grammatical structures, but that certainly does not mean that grammar (or any other linguistic subsystem) is not taught. The difference is that attention to grammar (or phonology, lexis, collocations, pragmatics, etc.) is not carried out as a separate activity, as an end in itself (focus on forms), but *during* (and if necessary after, but not before) task work, as part of the *methodology* of TBLT.

A lot of grammar is learned from positive evidence, that is, from repeated exposure to instances of target grammatical features and their authentic uses encountered primarily in task-based materials. Most problems are dealt with reactively, usually by the teacher, but also by other students, and, if pedagogic tasks are designed cleverly enough, by corrective feedback loops in the tasks themselves. The *reactive* quality of MP6 (focus on form) and MP7 (provide negative feedback) means that the timing of attention to grammar is more likely to be developmentally appropriate and occur at the most propitious moment for the learner(s) concerned, not arbitrarily, when predetermined by an unseen textbook writer. In sum, linguistic items are dealt with, and dealt with in a more scientifically defensible manner than by the traditional synthetic syllabus.

3.2. TBLT Neglects Vocabulary

While some critics allege that TBLT neglects grammar, others, and a few of the same ones, claim that it neglects vocabulary:

As course designers discover, even at beginner's level a surprising amount of basic vocabulary is likely to fall through the net of a syllabus based purely on a communicative needs analysis; at higher levels ensuring coverage of the 'most significant lexis' in this way becomes increasingly problematic. (Swan, 2005, p. 393)

This is a straw man. TBLT devotes far *more* attention to lexis and to collocations—and to *relevant* lexis and collocations, moreover—than do the majority of commercial textbooks. Advanced learners, especially, have usually mastered most of the grammar they need, but typically still lack the vocabulary and collocations required to operate effectively in the specialized discourse domains (academic, occupational, vocational, etc.) that usually motivated their need for advanced L2 abilities in the first place. The English for nebulous purposes (ENP)

courses commercial publishers favor do not teach such things because they are not designed to cater to the communicative needs of particular learners or groups of learners. ENP textbook sales depend on consumers' continued belief in the myth that the same course and the same language are appropriate for all learners, regardless of where and how they will put their L2 to use. In comments that, even if true, could only apply to beginners, Swan maintained that this is not a problem:

The most basic structures are needed all the time as learners struggle to talk about themselves, their surroundings, and their experiences; a well-planned traditional structural syllabus is therefore, very precisely, an expression of a needs analysis. (Swan, 2005, pp. 394–395)

In TBLT course design, once target tasks for learners have been identified, the second stage of a (real) NA involves the analysis of representative samples of target discourse surrounding successful completion of those tasks. This is precisely so that the materials based on the findings will expose students to the language used in the target discourse domains relevant for them—and that most certainly includes the lexis and collocations appropriate for those domains.

One of several advantages of the *task*-based *elaborated input* employed in TBLT materials is precisely that it retains authentic models of the use of lexical items and collocations typically absent in the *text*-based *linguistically simplified input* of spoken and written materials in commercial course books. Many such items are removed through the simplification process because they fall outside the frequency range for the textbook and/or because they are domain-specific. Needless to say, removing them from the input makes their acquisition difficult for learners.

At least a dozen empirical studies have demonstrated that elaborated spoken and written input achieves almost the same level of comprehensibility as simplified input, with various forms of redundancy in the elaboration process compensating for the higher linguistic burden that elaborated materials present. (For a review of research findings, see Long, 2015, pp. 248–258). The thousands of words advanced learners require (Nation, 2006), and the even greater numbers of collocations, are far too many to be handled explicitly through PPP pedagogy in the limited classroom time available. Attention (with or without conscious awareness) and repeated exposure to target items will need to suffice for *gradual uptake* (hence, input frequency effects). This is another reason for the major research focus in recent years on relationships between incidental and intentional learning of lexis and collocations, moderated by frequency and intensity of exposure, and on so-called chunk learning (for reviews and discussion of some pedagogic options, see Boers & Lindstromberg, 2012; Long, 2015, pp. 307–316).

In Swan's view (2005, p. 393), a lexical syllabus based on word frequency lists is the only way of ensuring adequate lexical coverage. Unfortunately, most word frequency lists are based on corpora unconnected to the discourse domains relevant for particular groups or types of learners, especially advanced learners. Meanwhile, Swan was confident in his own ability to intuit "the most significant lexis" learners require (most significant for *all* learners, presumably):

While, clearly, in a task-based program language relevant to the performance of the chosen tasks will be foregrounded, other important items that fall outside this framework cannot be guaranteed to occur naturally under normal time constraints, and may not therefore become available for learning. (What selection of tasks would reliably generate the occurrence of, or the need for, the following common lexical items: *control, rough, move, calm, noisy, swallow, take trouble, lane, genuine, out of sight, purple?*). (Swan, 2005, p. 393)

What selection of tasks, indeed? I confess, “purple,” for example, has yet to appear in any task-based materials with which I am familiar. There must be L2 speakers, therefore, faced with a lifetime flummoxed by Alice Walker and struggling to express themselves whenever they visit exhibitions of impressionist paintings, victims of the task syllabus.

The purpose of a NA is precisely to identify both the target tasks and the language students need, including vocabulary and collocations. Language shown to be important through data-based research, not stipulated as important on the basis of word frequency lists or a textbook writer’s intuitions concerning a field with which he or she is unfamiliar, is then included in the task-based materials that follow. This does not preclude students adding to their L2 repertoire later, of course. Why should items that fall outside this framework be a concern? If they fall outside, it is because the NA has shown they are not yet needed by the learners in question. It should not be forgotten, moreover, that studies have often found LT materials writers’ intuitions about language use to differ considerably from authentic language use, even in familiar everyday activities for which their textbooks are supposedly intended to prepare students (see, e.g., Auerbach & Burgess, 1985; Bartlett, 2005; Cathcart, 1989).

3.3. TBLT Provides Students With an Inadequate Supply of New Language

This allegation comes in different flavors. Bruton (2002a, p. 286), for example, repeatedly equated TBLT with oral communication among students working together in pairs or small groups on a miscellany of “task types” (see section 1). He wrote that “in TBLT, the mono-episodic interactive oral communication task, or the task of communicating orally, is central” (Bruton, 2005, p. 56), and that “the most obvious limitation of TBLT is its virtually exclusive focus on oral expression” (p. 57). This would come as news to the many designers, materials writers, and end users of task-based courses for listening and/or reading skills. Bruton denied that students can learn much from one another, especially new language, through working on such tasks, particularly if they are beginners and/or at roughly the same proficiency level. Seedhouse (2005) argued that because it is hard, and sometimes impossible, to predict language use from tasks, tasks cannot be relied upon for planning a language course.

The “lack of new language from peers” charge ignores the fact that small group work is typically just one component in a TBLT course, when it is a component at all, and also, as noted in section 1, that it is the task-based *materials* students work

on—alone, in pairs, in small groups, at the computer, or in lockstep mode with the teacher—which constitute the major source of new language, just as they do in conventional PPP courses. Moreover, contra Bruton, students, including those of the same L2 proficiency, can and do learn from one another when working in pairs or small groups (Eckerth, 2008; Swain, Brooks, & Tocalli-Beller, 2002). As for the difficulty of predicting linguistic forms from task, this would be a relevant concern if a structural syllabus were the goal, as in task-supported LT, but not in TBLT, which employs a task syllabus.

Bruton and Seedhouse are not alone. Swan claimed that TBLT may successfully develop learners' command of what they already know, but that the use of tasks, coupled with TBLT's supposed emphasis on spoken production, is less successful for teaching new language. Where does new language come from, he asked?

TBI [task-based instruction] puts great emphasis on output ...has less time available per class hour for new language input than a conventional text-heavy course with its battery of reading and listening material ...[and] provides learners with substantially less new language than "traditional" approaches. This seems a serious weakness. (Swan, 2005, p. 392)

Like Bruton, Swan provided no evidence for these assertions, but there is plenty of counterevidence. Many of the 80-plus studies included in two aforementioned statistical meta-analyses (Goo et al., 2015; Norris & Ortega, 2000), for example, involved new language for the learners concerned. Other researchers have compared the linguistic and/or informational content of genuine, linguistically simplified, and elaborated versions of texts, and their effects on listening and reading comprehension, and sometimes on uptake of new linguistic items, as well (see, e.g., Kim, 2006; Yano, Long, & Ross, 1994; Oh, 2001; and for review, Long, 2015, pp. 248–259). In sum, in my own approach to TBLT, (a) the collection and analysis of genuine samples of target discourse for specific groups of learners (the second part of a task-based needs analysis), (b) the use of *modified elaborated* versions of those samples in the resulting task-based materials, and (c) MP3 (elaborate input) and MP4 (provide rich input) collectively address the "new language" issue. Detailed examples of the whole process are available, along with illustrative modules of task-based materials for learners at various proficiency levels (see, e.g., Bartlett, 2005; Long, 2005, 2015, pp. 169–204, 300–328; O'Connell, 2014).

The asserted emphasis on output is a fallacy. Depending on why students are learning the language concerned, some TBLT courses devote little or no time at all to output. Contrary to the implicit assumption of most commercial textbooks, which usually deal in all four skills, students whose academic or occupational needs are such that they exclusively require listening and/or reading abilities may have little or no interest in learning to speak or write the L2. Some classroom studies, moreover (e.g., Shintani, 2011, 2013), have compared input-based and output-based instruction; the results, together with those of a statistical meta-analysis of the same issues in teaching grammar in general (Shintani, Li, & Ellis, 2013), favor input-based approaches. Almost all sequences of pedagogic tasks with which I

am familiar begin with input-based tasks, and many end there. Individual modules of task-based materials may prioritize oral output, but if they do so, it is because speech is relevant for the target tasks at hand, for example, some designed for tourists or for U.S. diplomats required to deliver celebratory speeches in Japanese (Hillman, 2015), but it is simply untrue of TBLT as a whole.

Swan also had this to say about new language':

Pre-teaching, texts and other materials in fact seem to have a largely ancillary function in TBI: they are primarily a source of *task-relevant language or information*, rather than a vehicle for presenting *new language that has been selected in accordance with task-external criteria*. (Swan, 2005, p. 391, emphasis added)

It is true that TBLT devotes most of its attention to equipping learners with the abilities they need, including familiarity with the new *task-relevant language* they need, to perform their target tasks successfully. That is what TBLT is about. The intriguing question is whence this other more important "task-external" language Swan claimed all learners should know, regardless of their purposes in learning the language.

3.4. Task-Focused Peer-Peer Interaction Is Restricted and Often Ungrammatical

Despite abundant evidence to the contrary (e.g., Long & Porter, 1985; Pica, Lincoln-Porter, Paninos, & Linnell, 1996), a fear persists in some LT circles that student talk in small groups inevitably leads to errors, potentially perpetuated by deviant peer input and the absence of prompt teacher "correction." Task-based peer-peer interaction, in particular, is supposed to be especially vulnerable. Seedhouse (1999) based his critique of TBLT along these lines, using one short excerpt of 10–13 lines from the conversation of each of three pairs of beginning English as a foreign language (EFL) learners performing Spot the Difference tasks. There is no information about how typical the excerpts were of those conversations, or how typical or representative they might be of similar task-based conversation among other students, especially more advanced students. Seedhouse seems simply to have assumed that they were, and are, typical. The interaction in each case is marked by a tendency toward indexicality and minimalization (Seedhouse, 1999, pp. 152–154). The former means interaction that is context-embedded and inexplicit, with talk interpretable only by outsiders with knowledge of the task at hand. This is of less concern to Seedhouse, as he recognized that real-world interaction among native speakers is often like that. Minimalization means that learners use as much language as needed to complete a task, and no more, and language that, at least where beginners are concerned, may be stripped of obligatory grammatical items, for example, copulas. This, Seedhouse (1999) argued, is more serious because "teachers want to see some evidence of the learners' linguistic competence being stretched and challenged and upgraded" (p. 153), not language that "resembles a pidgin" (p. 154). Copula omission, he noted, is a feature of pidgins and creoles.

Bruton echoed this charge, professing an “overriding concern” that unsupervised peer interaction will lead to the development of a classroom pidgin (Bruton, 2002a, p. 286). He feared that “if oral communication tasks do stimulate creative use of the TL [target language] and what Swain (1985) calls ‘pushed’ output with no immediate interactive constraints, there is no guarantee that the outcome will not be pidgin-like,” and that if too much peer–peer interaction on communication tasks by less proficient students is allowed, “pidgin-like language has every likelihood of emerging” (Bruton, 2005, p. 60). While this sudden interest in pidgins (and creoles?) is welcome, it reveals an unfortunate ignorance of their true nature, functions, and linguistic characteristics, and of the social and sociolinguistic circumstances required for pidginization. Bruton (2005, p. 63), for example, was especially worried about the potential for a pidgin to emerge in *monolingual* secondary school foreign language classrooms—a possibility that would intrigue linguists who have devoted their careers to the study of pidgins and pidginization.

Seedhouse’s conclusion is that “task-based interaction is a particularly narrow and restricted variety of communication ... [and that] it remains to be proven that task-based interaction is more effective than other varieties of classroom interaction” (Seedhouse, 1999, p. 155). Leaving aside the sweeping nature of these generalizations based on flimsy evidence and their refutation by many studies documenting the effectiveness of task-based conversation for language development (again, for statistical meta-analyses of research findings, see Li, 2010; Mackey & Goo, 2007), it is certainly the case that teachers’ skillful use of tightly controlled mechanical or meaningful production opportunities in PPP lessons can temporarily elicit limited examples of native-like performance, complete with copulas and other otherwise potentially vulnerable morphology. But as many researchers have demonstrated, it is precisely that sort of *traditional code-focused instruction*—language-like behavior—not task-based interaction, that generates “a narrow and restricted variety of communication,” especially at the elementary level, typically couched in long sequences of initiation–response–feedback (IRF) exchanges:

T: Is Paris the capital of France?

S1: Yes, it is.

T: Good. Yes it is. Is New York the capital of the USA?

S2: No, it isn’t.

T: No, it isn’t. Good.

And so on.

For decades, studies have found the repetitive IRF exchange structure to be typical of both content and language classrooms (see, e.g., Dinsmore, 1985; Hoetker & Ahlbrand, 1969; Long & Sato, 1983; Nunan, 1987). Unlike TBLT, the input and output opportunities *that* sort of interaction provides certainly are narrow and restricted. The fact that the same learners usually still lack implicit knowledge of the structures du jour (or for those of the skill-building persuasion, sufficient control of them, due to unfinished proceduralization or automatization) is usually concealed by the fact that code-focused LT means they rarely have to try to *communicate*

using the forms. Task-based interaction can be seen as valuable precisely because it both reveals what learners really know and provides them with opportunities to develop and use their L2 in ways more akin to conditions they will face outside the classroom.

For his claims to be taken seriously, Seedhouse would need to produce longitudinal data. (Bruton offered no data of any kind.) Did the students in Seedhouse's (1999) study stabilize as he described, or did they gradually improve, as most elementary classroom learners do? The kind of speech he reported is what is sometimes produced initially (not permanently) by the comparatively few beginners allowed opportunities to *communicate* in the L2—the opportunities often arising precisely because of a teacher's use of tasks. In other words, copula omission (for example) is neither peculiar to task performance, nor task performance its cause, but rather, reflective of the current state of the learner's interim interlanguage grammar. Seedhouse's three excerpts are all sampled from elementary learners performing pedagogic tasks of very limited scope. It is easy to show more target-like, even near-native, task-based L2 use by more advanced learners performing more complex tasks—not *because* they are performing tasks, any more than the beginners' non-target-like use is *because* of tasks.

By way of illustration, the late Johannes Eckerth (2008) reported a methodologically rigorous longitudinal study of peer–peer interaction on three text-repair, and two text-reconstruction, consciousness-raising tasks—interaction that resulted not only in gains in explicit knowledge of the targeted structures, but also of additional linguistic features outside the explicit focus of the tasks. Participants were 31 young adult learners in two intact German classes, 14 of lower-intermediate and 17 of higher-intermediate proficiency. The pretest, posttest, delayed posttest study lasted 8 weeks, with 12 hours of instruction per week. As measured by a discrete-item sentence-assembly task, statistically significant short- and medium-term gains by both lower-intermediate and higher-intermediate groups were observed for the targeted structures and, through use of additional tailor-made test items, for non-targeted structures, as well.

Of particular interest in the present context, detailed analyses by Eckerth (2008) of 50 hours of audiorecorded peer–peer interaction revealed 254 instances of individual learner L2 hypotheses during 127 language-related episodes. Less than 1% led to non-target-like revisions of originally correct learner hypotheses, meaning that there was little empirical support for the idea that learners reinforce one another's incorrect L2 assumptions. On the contrary, in 38% of cases, non-target-like hypotheses were replaced by L2-conforming representations and integrated into the learners' interlanguage grammars, as reflected in the medium-term posttest, with the majority of such changes traceable to learning opportunities arising from the task-based interactions. In 22% of cases, neither learner possessed the relevant L2 knowledge at the outset, yet the collaborative dialogue led to learning gains after task completion. Eckerth's verdict was that the task work by pairs of learners indicated “significant gains in the short and medium term. Thus, it can be concluded that learners are able to provide each other with feedback rich in acquisitional potential” (Eckerth, 2008, p. 133).

Holding communicative task-based work culpable for, and likely to foster, restricted and/or ungrammatical L2 use would be analogous to holding a physician's prescription of a treadmill stress test responsible for what it reveals about a patient's physical health, problems that would have remained hidden without use of the treadmill. The relevant question is whether the patient's health improves over time if he or she undertakes a program of regular physical exercise. Research on language learning and teaching shows that where regular linguistic exercise is concerned, it does. The idea that task-based interaction is restricted, and when ungrammatical, likely to remain so, is unfounded—another *nonissue*.

3.5. TBLT Views Form-Based and Meaning-Based Instruction as Mutually Exclusive

Swan claimed that TBLT unconstructively polarizes meaning-based and form-based instruction, adding that “any teaching philosophy which deliberately excludes one of these complementary elements in favor of the other has arguably forfeited its right to be taken seriously” (Swan, 2005, p. 396). This is a particularly egregious straw man. There is no such polarization in the rationale for TBLT or in its implementation, and no exclusion of either form or meaning. On the contrary, through MP1 (use task, not text, as the unit of analysis), MP2 (promote learning by doing), MP6 (focus on form), MP7 (provide implicit negative feedback), and other MPs, attention to form and meaning are much *better* integrated in TBLT than in traditional LT. Attention to grammar and other code features in TBLT occurs in context, embedded in meaning-based activities, not in separate drill-and-kill sessions.

3.6. The Role of the Teacher Is Downgraded in TBLT

Critics often assert, wrongly, that the role of the teacher in TBLT is downgraded. Swan (2005, p. 391) wrote that “the thrust of TBI is to cast the teacher in the role of a manager and facilitator of communicative activity rather than an important source of new language.” Bruton included downgrading of the teacher's role in a list of increasingly bizarre assertions:

Probably the most notable defects of strong TBLT are the lack of importance given to input, the lack of clarity on language development and the exclusion of the communicative role of the teacher. These inherent defects are largely the result of avoiding clarity on design and intended procedures and resorting to on-line responsive implementational gambits, which are extremely unreliable. (Bruton, 2005, p. 65)

Again, this is a significant misunderstanding of TBLT. In fact, quite the opposite is true. The teacher's role in TBLT requires greater expertise, and is more important, more demanding, and certainly more communicative than in PPP. MP6 (focus on form) and MP7 (provide corrective feedback), for example, can each often require an appropriate diagnosis and quick-thinking reaction, triggered by unforeseen learner errors or by the surprising direction in which learners sometimes take a task. This is harder (but also more intellectually stimulating and rewarding)

than working through tedious drills and exercises one item at a time, knowing what the learners are supposed to say or write before they say or write it. Untrained teachers with poor command of the L2 can do PPP; they cannot do TBLT. The PPP teacher is relegated to the job of implementing lessons planned down to the last drill and exercise by an unseen textbook writer and assumed appropriate for all students in a group on the same predetermined day. As Skehan (2002) pointed out in a response to an earlier flurry of charges from Bruton (2002a), it is Bruton's favored language-based, lockstep PPP, delivered following the dictates of a one-size-fits-all commercial textbook, which really undermines the teacher's creativity and decision-making role. In contrast, TBLT tailors input and feedback to individual learners. "This," Skehan pointed out, "requires a skilful, responsive, knowledgeable teacher who is able to cope with groups of learners and access relevant material as the need arises" (Skehan, 2002, p. 295).

Bruton, however, had not only claimed that TBLT downgraded the language teacher, but that empirical research on tasks was used "to lobby particular academic interests" (Bruton, 2002a, p. 288). I can do no better than quote Skehan's response:

If therefore an assessment is made about whose interests are served, it seems to me that the anti-task approach that Bruton takes implicitly serves the interests of publishers and of conventional approaches to instruction. It does a disservice to the teacher. It does an even greater disservice to the learner. (Skehan, 2002, p. 295)

4. ALLEGED PROBLEMS WITH IMPLEMENTING TBLT IN SPECIFIC CONTEXTS

4.1. TBLT Is Too Limited for Use in State Secondary Schools

Bruton (2005) equated TBLT with the use of what he called "communicative oral peerwork tasks (COPTs)," and then attacks (imaginary) courses based on their exclusive use. A more obvious case of straw man argumentation would be hard to find. His conclusion:

TBLT for more initial levels of formal language teaching has serious limitations, especially because students need to receive the necessary input and relevant feedback from somewhere on a consistent basis. On these counts, for teachers in (state) secondary schools with large numbers of students at pre-intermediate levels, COPTs have very little to offer as core elements in the curriculum ... they are very limited in scope and focus on one skill, namely speaking. (Bruton, 2005, p. 66)

All this may or may not be true, but what does it have to do with TBLT?

4.2. TBLT Will Not Work in 3-Hours-per-Week Foreign Language Situations

Several critics (e.g., Bruton, 2005; Fotos, 2002), asserted that TBLT can only be successful, if anywhere, in acquisition-rich L2 settings, not in typical foreign

language situations. Swan (2005, pp. 393–394) concurred. “Heavily task-based programs” are especially inappropriate in his view where instructional time is limited to 3 hours per week, and input is unavailable outside the classroom, the situation for most of the world’s language learners.

Inadequate instructional time and lack of L2 exposure outside the classroom are real problems for all kinds of LT, not just TBLT, and not just in foreign language settings. The question is how best to utilize what time there is. Studies like those cited earlier on focus on form show that TBLT makes better use of it than the diet of PPP most critics advocate. Students in TBLT courses learn as much grammar and vocabulary, even at the beginners level, with instruction geared to learners’ processing capacity, and with the additional crucial advantage that course content is relevant to their communicative needs.

In my view, a more serious problem (for any communicative LT, not just TBLT) in many foreign language programs for high school students, especially those that form part of centralized, state-controlled education systems, is washback from what is often a grammar-based final exam that both teachers and students know lies in wait at the end of the year. In some countries, the exam looks more like an IQ test and lacks validity for language assessment, but it can have serious implications for university entrance, so tends to exert a huge influence on what is taught and how. A functional command of the L2 is all very well, and much more useful in the real world, but will the students be able to convert active declarative sentences into passives on the test, or remember the uses and verb endings for the Spanish subjunctive? Studies to date (for review, see Long, 2015, pp. 350–364) suggest that students who learn a language in communicative courses of some kind, including TBLT courses, do as well on discrete-point tests and measures of “general” language proficiency as graduates of PPP-type programs, and better than them on measures of communicative abilities. However, there is a need for research on relationships between students’ communicative abilities and their performance on the state-mandated type of discrete-point tests favored by some ministries of education. Meanwhile, in a change that may eventually obviate the problem, there is a growing recognition in the language policy of some east Asian countries of the importance of communicative language abilities, and a wish to foster more positive washback.

4.3. Evidence Is Lacking of Successful Long-Term Classroom Implementation of TBLT.

Of the many charges Swan leveled, this one (2005, p. 379) had some merit in 2005, when his original article was published, although less so when it was reprinted in 2011 and today in 2016. Reports of the first 20 years of successful implementation of TBLT in over 200 state primary and secondary schools in Flanders appeared in 2006 (Van den Branden, 2006a). Additional formative and summative evaluations of TBLT for children and adults have been published since, all favorable to a greater or lesser extent—for example, Gonzalez-Lloret & Nielson (2015; Spanish in the United States), Markee (2007; English as a second language [ESL] in the United

States), Nielson (2015; distance EFL and ESL), and Shintani (2011, 2013; EFL for children in Japan). (For a review of these and other studies, see Long, 2015, pp. 341–366). While results so far are positive, more evaluations of TBLT are needed, especially of complete programs in situ. The scarcity of evaluations is mostly due to the advent of TBLT having been relatively recent. Meanwhile, critics would have more credibility if they could point to a single published program evaluation supporting the kind of LT they advocate, which has been around a lot longer.

5. TASKS AND TBLT: SOME REAL ISSUES

While many criticisms of tasks and TBLT are unwarranted—*nonissues*—this is emphatically *not* to say that TBLT is a panacea. Several problems remain—*real issues*. For space limitations, I can do no more than mention three and very briefly reference work under way to deal with them.

5.1. Task Complexity Criteria

Course designers and materials writers need reliable criteria by which to measure task complexity, the basis for sequencing pedagogic tasks in a task syllabus. Indicators of complexity will ideally be low inference, such as the number and distinctiveness of elements or steps a task involves, or whether it is context-embedded/here-and-now or context-reduced/there-and-then. Current practice often calls upon more abstract, less readily operationalized factors, such as the type (spatial, causal, intentional, etc.) and severity of “reasoning demands” involved.

Much good work has been published on task complexity and on task syllabus design (see, e.g., Gilabert, Barón, & Levkina, 2011; Révész, Michel, & Gilabert, 2015; Robinson, 2009, 2011; Skehan, 2014), and more is under way. The research effort has been intense, with over 300 studies to date in journal articles alone. The overall yield has been disappointing, however, largely due to the noncomparability of what are sometimes conflicting findings, mostly the result of researchers having used a wide variety of tasks, task types, moderator variables, and outcome measures (Sasayama, Malicka, & Norris, 2015). Recognizing the need to systematize future work, an international collaborative research network (CRN) on task complexity was established in 2015, which should help make findings cumulative, encourage replication studies, increase productivity, and generally speed up progress on this issue.

5.2. Task-Based Assessment and the Transferability of Task-Based Abilities

Views differ on the best way(s) to assess task-based abilities, although the process always involves use of task-based, criterion-referenced performance tests. (Speculations using scores on discrete-point grammar tests or achievement on a proficiency scale clearly will not do.) Things are fairly straightforward as long as unambiguous behavioral variables exist to indicate success or failure at task

completion. Were learners able or unable to reach their destination by following street directions? Did they score 80% or better on a multiple-choice test of the information (identified by a subject-matter expert) in an undergraduate physics lecture? But when, if at all, is assessment improved by the addition of a linguistic caboose? And when do results on one target task allow prediction of the same student's performance on a related task without the need to test the second task or tasks like it? How "distant" can two tasks be before predictions break down, and how can "distance" be measured? Work on these issues has begun, with valuable contributions from, among others, Benson (2015); Brown, Hudson, Norris, and Bonk (2002); Nielson (2015); Nielson, Masters, Rhoades, and Freynik (2009); and Norris (2009), but the transferability of task-based abilities remains a real issue.

5.3. *In-Service Teacher Education for TBLT*

Far from being downgraded, the teacher's role in TBLT is more important and requires more expertise than that of the teacher in focus on forms/PPP. This is potentially a problem, as the literature on the diffusion of education (see, e.g., Markee, 1993) has long identified large perceived differences between current performance and that required by new developments as a major factor mitigating against teacher buy-in and adoption of new ideas. This can be offset by the increased student motivation teachers report that TBLT brings, but it is a real issue, nonetheless. Useful work on teachers' classroom performance and in-service education for TBLT has been carried out (e.g., East, 2012; Foster, 2009; Newton, 2015; Samuda, 2015; Van den Branden, 2009). The Belgian experience, in particular (see Van den Branden, 2006b), has highlighted the importance of adequate in-service (re-)training and continued teacher support.

6. CONCLUSION

TBLT is still a relatively recent innovation—one whose adoption requires expertise on the part of course designers and classroom teachers, and a considerable investment of time and effort if it is to be successful. Success will be more likely in programs housed in institutions with a stable, qualified, and experienced teaching staff and a relatively stable student population (thereby making an NA and materials development worth the effort), located in societies where language learning is a serious matter and recognized and supported as such. Increasing numbers of school systems, universities, and government language programs in many countries meet those requirements. Programs for refugees and migrant workers would be very well served by TBLT, too, but often lack adequate staffing, funding, and other resources, and so require the assistance of volunteers.

Many criticisms of TBLT lack merit, but some raise important matters. No approach to LT has proven "correct" to date, and there should be no illusion as to TBLT's chances of breaking the tradition. Real issues remain. Advances in theory and research, coupled with further field trials, will assuredly refine current models,

and quite probably identify needed changes. The goal is for researchers and practitioners to move forward together systematically in what must be a collaborative endeavor.

Acknowledgments

I thank the editor, Alison Mackey, and two anonymous reviewers for their careful reading and valuable feedback on the original version of this article.

NOTE

1. A detailed point-by-point response to Wilkins's full critique is available elsewhere (Long, 2007, pp. 160–163), so will not be repeated here.

REFERENCES

- Auerbach, E. R., & Burgess, D. (1985). The hidden curriculum of survival ESL. *TESOL Quarterly*, 19(3), 475–495.
- Bartlett, N. D. (2005). A double shot 2% mocha latte, please, with whip: Service encounters in two coffee shops and at a coffee cart. In M.H. Long (Ed.), *Second language needs analysis* (pp. 305–43). Cambridge, UK: Cambridge University Press.
- Benson, S. (2015, February 9). Task-based language teaching: An empirical study of task transfer. *Language Teaching Research*. Advance online publication. doi:10.1177/1362168815569829
- Bettoni, C., & Di Biase, B. (Eds.). (2015). *Grammatical development in second languages: Exploring the boundaries of processability theory* (EUROSLA Monographs Series No 3). Retrieved from EUROSLA website: <http://www.eurosla.org/eurosla-monograph-series-2/eurosla-monographs-03/>
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, 32, 83–110.
- Bonilla, C.L. (2015). From number agreement to the subjunctive: Evidence for processability theory in L2 Spanish. *Second Language Research*, 31(1), 53–74.
- Brown, J. D., Hudson, T., Norris, J., & Bonk, W. J. (2002). *An investigation of second language task-based performance assessments*. Honolulu, HI: Second Language Teaching and Curriculum Center.
- Bruton, A. (2002a). From tasking purposes to purposing tasks. *ELT Journal*, 56(3), 280–288.
- Bruton, A. (2002b). When and how the language development in TBI? *ELT Journal*, 56(3), 296–297.
- Bruton, A. (2005). TBLT for the state secondary school classroom? *Language Learning Journal*, 31, 55–68.
- Cathcart, R. (1989). Authentic discourse and the survival English curriculum. *TESOL Quarterly*, 23(1), 105–126.
- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161–170.
- DeKeyser, R. (1998). Beyond focus on form: cognitive perspectives on learning and practicing second language grammar. In C. J. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition*, (2nd ed., pp. 94–112). Mahwah, NJ: Erlbaum.
- Dinsmore, D. (1985). Waiting for Godot in the EFL classroom. *ELT Journal*, 39(4), 225–234.
- Doughty, C. J. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge, UK: Cambridge University Press.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 256–310). New York, NY: Basil Blackwell.
- East, M. (2012). *Task-based language teaching from the teachers' perspective: Insights from New Zealand*. Amsterdam, the Netherlands: John Benjamins.
- Eckerth, J. (2008). Investigating consciousness-raising tasks: Pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics*, 19(3), 119–145.

- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–52.
- Ellis, N. C., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition. An introduction*. (2nd ed., pp. 75–93). Mahwah, NJ: Erlbaum.
- Ellis, R. (1989). Are classroom and naturalistic acquisition the same? A study of the classroom acquisition of German word order rules. *Studies in Second Language Acquisition*, 11(3), 305–328.
- Ellis, R. (1994). The structural syllabus and second language acquisition. *TESOL Quarterly*, 27(1), 91–113.
- Ellis, R. (1997). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.
- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- Foster, P. (2009). Task-based language learning research: expecting too much or too little? *International Journal of Applied Linguistics*, 19(3), 247–263.
- Fotos, S. (2002). Structure-based interactive tasks for the EFL grammar learner. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms*. Mahwah, NJ: Erlbaum.
- Gilabert, R., Barón, J., & Levkina, M. (2011). Manipulating task complexity across task types and modes. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 105–140). Amsterdam, the Netherlands: John Benjamins.
- Gonzalez-Lloret, M., & Nielson, K. (2015). Evaluating TBLT: The case of a task-based Spanish program. *Language Teaching Research*, 19(3), 525–549.
- Goo, J., Granena, G., Novella, M., & Yilmaz, Y. (2015). Implicit and explicit instruction in L2 learning: Norris and Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam, The Netherlands: John Benjamins.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35(1), 127–165.
- Håkansson, G., & Norrby, C. (2010). Environmental influence on language acquisition: Comparing second and foreign language acquisition of Swedish. *Language Learning*, 60(4), 628–650.
- Hillman, K. K. (2015, September 16–18). *Needs analysis and development of a TBLT module for US foreign service officers*. Paper presented at the Sixth International Conference on Task-Based Language Teaching, Katolieke Universiteit Leuven, Belgium.
- Hoetker, J., & Ahlbrand, W.P. (1969). The persistence of the recitation. *American Educational Research Journal*, 6(1), 145–167.
- Johnston, M. (1985). *Syntactic and morphological progressions in learner English*. Canberra, Australia: Commonwealth Department of Immigration and Ethnic Affairs.
- Johnston, M. (1997). *Development and variation in learner language (Unpublished doctoral dissertation)*. Australian National University, Canberra, Australia.
- Kim, Y. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as foreign language. *TESOL Quarterly*, 40(2), 341–373.
- Klapper, J. (2003). Taking communication to task? A critical review of recent trends in language teaching. *Language Learning Journal*, 27, 33–42.
- Klapper, J., & Rees, S. (2003). Reviewing the case for explicit grammar instruction in the university foreign language learning context. *Language Teaching Research*, 7(3), 285–314.
- Lenzing, A. (2015). Relating regularities and dynamical systems in L2 development. *Language Learning*, 65(1), 89–112.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365.
- Long, M. H. (1985). A role for instruction in second language acquisition: task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and assessing second language development* (pp. 77–99). Clevedon, UK: Multilingual Matters.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bahtia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York, NY: Academic Press.
- Long, M. H. (Ed.). (2005). *Second language needs analysis*. Cambridge, UK: Cambridge University Press.

- Long, M. H. (2007). SLA: Breaking the siege. In M. H Long, *Problems in SLA* (pp. 139–168). Mahwah, NJ: Erlbaum.
- Long, M. H. (2009). Methodological principles in language teaching. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 373–94). Oxford, UK: Blackwell.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Oxford, UK: Wiley-Blackwell.
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based language teaching. *TESOL Quarterly*, 26(1), 27–56. Reprinted in K. Van den Branden, M. Bygate, & J. M. Norris (Eds.), *Task Based Language Teaching: A reader* (pp. 57–81). Amsterdam, the Netherlands: John Benjamins, 2009.
- Long, M. H., & Porter, P. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207–227.
- Long, M. H., & Sato, C. J. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. W. Seliger & M. H. Long (Eds.), *Classroom-oriented research on second language acquisition* (pp. 268–85). Rowley, MA: Newbury House.
- Mackey, A. (1999). Input, interaction and second language development. *Studies in Second Language Acquisition*, 21(4), 557–587.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407–452). Oxford, UK: Oxford University Press.
- Markee, N. (1993). The diffusion of innovation in language teaching. *Annual Review of Applied Linguistics*, 13, 229–243.
- Markee, N. (2007). *Managing curricular innovation*. Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82.
- Newton, J. (2015, November 30). *Teachers, tasks and textbooks*. Plenary address, AALA conference, Adelaide, Australia.
- Nielson, K. (2015, September 16–18). *Evaluating the impact of online practice on real-world task performance of second language learners*. Paper presented at the Sixth International Conference on Task-Based Language Teaching, Katolieke Universiteit Leuven, Belgium.
- Nielson, K. B., Masters, M. C., Rhoades, E., & Freynik, S. (2009). *Prototype implementation of an online Chinese course: An analysis of course implementation and learner performance*. College Park, MD: University of Maryland Center for Advanced Study of Language.
- Norris, J. M. (2009). Task-based teaching and testing. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 578–594). Oxford, UK: Blackwell.
- Norris, J. M., Brown, J. D., Hudson, T. D., & Bonk, W. (2002). Examinee abilities and task difficulty in task-based second language performance assessment. *Language Testing*, 19(4), 395–418.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- Nunan, D. (1987). Communicative language teaching: Making it work. *English Language Teaching Journal*, 41(2), 136–145.
- Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL Quarterly*, 25(2), 279–295.
- O'Connell, S. (2014). A task-based language teaching approach to the police traffic stop. *TESL Canada Journal*, 31(Special issue 8), 116–131.
- Oh, S.-Y. (2001). Two types of input modification and EFL reading comprehension: simplification versus elaboration. *TESOL Quarterly*, 35(1), 69–96.
- Ortega, L. (2009). Sequences and processes in language learning. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 81–105). Oxford, UK: Blackwell.
- Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of language learners? *TESOL Quarterly*, 30(1), 59–84.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, 6(2), 186–214.
- Pienemann, M. (2005). *Cross-linguistic aspects of processability theory*. Amsterdam, the Netherlands: John Benjamins.
- Pienemann, M., & Kessler, J.-U. (2012). Processability theory. In S. M. Gass & A. Mackey (Eds.) *The Routledge handbook of second language acquisition* (pp. 228–246). New York, NY: Routledge.
- Rastelli, S. (2014). *Discontinuity in second language acquisition. The switch between statistical and grammatical learning*. Bristol, UK: Multilingual Matters.

- Révész, A., Michel, M., & Gilabert, R. (2015, September 14). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*. Advance online publication. doi:10.1017/S0272263115000339
- Robinson, P. (2009). Syllabus design. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 294–310). Oxford, UK: Blackwell.
- Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3–37). Amsterdam, the Netherlands: John Benjamins.
- Samuda, V. (2015). Tasks, design, and the architecture of pedagogical spaces. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 271–302). Amsterdam, the Netherlands: John Benjamins.
- Sasayama, S., Malicka, A., & Norris, J. M. (2015, September 16–18). *Primary challenges in cognitive task complexity research: Results of a comprehensive research synthesis*. Paper presented at the colloquium, An International Collaborative Research Network (CRN) on Task Complexity, at the Sixth International Conference on Task-Based Language Teaching (TBLT), Katolieke Universiteit Leuven, Belgium.
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24, 139–161.
- Schmidt, R. W. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, ... & I. Walker (Eds.), *Proceedings of CLASIC 2010*, Singapore, December 2–4 (pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.
- Seedhouse, P. (1999). Task-based interaction. *ELT Journal*, 53(3), 149–156.
- Seedhouse, P. (2005). “Task” as research construct. *Language Learning*, 55(3), 533–570.
- Serafini, E. J., Lake, J., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, 40, 11–26.
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15, 137–158.
- Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginner learners. *TESOL Quarterly*, 47(1), 36–62.
- Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, 63(2), 296–329.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- Skehan, P. (2002). A non-marginal role for tasks. *ELT Journal*, 56(3), 289–295.
- Skehan, P. (2014). *Processing perspectives on task performance (task-based language teaching)*. Amsterdam, the Netherlands: John Benjamins.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass, & C. M. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). Peer-peer dialog as a means of second language learning? *Annual Review of Applied Linguistics*, 22, 171–185.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3), 376–401.
- Swan, M. (2011). Legislation by hypothesis: The case of task-based instruction. In M. Swan, *Thinking about language teaching: Selected articles 1982–2011* (pp. 90–113). Oxford, UK: Oxford University Press.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14, 103–135.
- Van den Branden, K. (Ed.), (2006a). *Task-based language education: From theory to practice*. Cambridge, UK: Cambridge University Press.
- Van den Branden, K. (2006b). Training teachers: Task-based as well? In K. Van den Branden (Ed.), *Task-based language education: From theory to practice* (pp. 217–248). Cambridge, UK: Cambridge University Press.
- Van den Branden, K. (2009). Mediating between predetermined order and chaos: The role of the teacher in task-based language education. *International Journal of Applied Linguistics*, 19(3), 264–285.
- Van Gorp, K., & Deygers, B. (2013). Task-based language assessment. In A. Kunan (Ed.), *The companion to language assessment: Vol. 2. Approaches and development* (pp. 578–593). Oxford, UK: Wiley-Blackwell.

- Widdowson, H. G. (1993). Aspects of syllabus design. In M. L. Tickoo (Ed.), *Language syllabuses: State of the art* (pp. 65–89). Singapore: SEAMEO Regional Language Center.
- Widdowson, H. G. (2003). *Defining issues in English language teaching*. Oxford, UK: Oxford University Press.
- Wilkins, D. (1994). Language, language acquisition and syllabus design: Some recent issues. *English Teaching (Korea)*, 49, 41–56.
- Willis, D. (1994). Comments on Michael H. Long and Graham Crookes' "Three approaches to task-based syllabus design." *A reader reacts. TESOL Quarterly*, 27(4), 726–729.
- Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, 44(2), 189–219.