30 Task-Based Teaching and Testing

JOHN NORRIS

Introduction

Task-based language teaching (TBLT) is an approach to second or foreign language education that integrates theoretical and empirical foundations for good pedagogy with a focus on tangible learning outcomes in the form of "tasks" – that is, what learners are able to *do* with the language. Task-based practice draws on diverse sources, including philosophy of education, theories of second language acquisition, and research-based evidence about effective instruction. Equally important, TBLT acts on the exigencies of language learning in human endeavors, and the often ineffectual responses of language education to date, by providing a framework within which educators can construct effective programs that meet the language use needs of learners and society.

Though there is global interest in the value of TBLT to foster worthwhile language learning, there is also diversity in the educational scope, practical applications, and research associated with the name. Certainly, TBLT remains a contested domain of inquiry and practice, though much of the debate surrounding TBLT results from incomplete understandings of precisely what this educational approach comprises. In the following, I review key underpinnings of task-based instruction, its emergence within language education, and its component parts. I then highlight the fundamental processes of teaching and testing, outlining a task-based approach to each and posing questions in need of inquiry. I conclude by forecasting several challenges that will condition the ultimate contribution of TBLT to language education.

Task-Based Language Education

At its most basic, task-based instruction rejects the notion that knowledge can be learned independently of its application and embraces instead the value of learning by doing, or "experiential learning" (Dewey, 1933). In Dewey's terms, principal

elements around which instruction should be built are "activities worthwhile for their own sake" (p. 87), and it is by engaging learners in doing valued activities that relevant declarative and procedural knowledge is developed, learners are motivated to engage with instructional content, and learners develop deep linkages between what they learn and how that learning can be put to use beyond the classroom. Crucial cognitive and emotional mechanisms are triggered through learning by doing things holistically, including in particular the essential feedback loop (James, 1899) whereby "we receive sensible news of our behavior and its results" (p. 41) in the context of, and relatable to, the activity that we are immediately focused on doing. Furthermore, holistic activities provide learners the opportunity to analyze what they do, what works, and what doesn't, thereby constructing their own explanations (Dewey, 1938) and rendering such understandings "available under actual conditions of life" (p. 48).

These and related tenets of experiential learning have grown into diverse models of practice, including cognitive psychological learning theories (e.g., Sternberg, 2003), apprenticeship and socialization frameworks (e.g., Lave and Wenger, 1991), and others (e.g., Kolb, 1984). Key is the idea that holistic activity structures, such as tasks, offer an ideal frame within which knowledge use can be experienced and understood, and from which learning opportunities should be developed. For example, Sternberg (2003) advocated:

> For starters, this means having students do tasks, or at least meaningful simulations, that experts do in the various disciplines. Second, it means teaching them to think in ways that experts do when they perform these tasks. (p. 5)

Following these models, "task" has proved to be an effective organizing principle for the implementation of experiential learning across diverse disciplines, including medicine (e.g., Virjo, Holmberg-Mattila, & Mattila, 2001), environmental studies (e.g., Wright, 2000), and social work (e.g., Reid, 1997). In language education, too, "task" has emerged as a heuristic for encapsulating the many benefits of experiential learning, as summarized in Samuda and Bygate (2008): "What we are calling 'tasks' can thus be seen as a means of creating experience-based opportunities for language learning" (p. 36).

In the recent history of language teaching (cf. Musumeci, this volume), the need for an organizing principle like "task" materialized in the 1980s, responding to dissatisfaction with dominant traditions and in consonance with findings from research into how second language acquisition occurs in instructed and naturalistic settings (Long, 1985). Prevailing language teaching methods of the post-war era were found lacking on several scores: (1) they presented language forms (grammatical rules, vocabulary words, etc.) in an arbitrary order and disarticulated from their communicative functions; (2) they posited rapid and complete development of accuracy in response to rule- or pattern-based learning; and (3) they presumed the transfer of accumulated bits of discrete knowledge about the language into the ability to utilize the language for communication. By adopting a synthetic approach to syllabus design (i.e., where learners synthesize discrete

facts into holistic understandings; Wilkins, 1976) and a focus-on-forms methodology (i.e., teaching language forms disconnected from their functional uses; Long & Robinson, 1998), these approaches ran counter to how, and how quickly, language is actually acquired, and they generally met with disappointing outcomes and disenchanted learners (Widdowson, 1978).

Communicative language teaching (CLT) represented the opposite swing of the pendulum, countering that language learning in the classroom must mimic the naturalistic acquisition of communicative abilities outside of the classroom. The natural approach (Krashen & Terrell, 1983), immersion education, and other strong forms of CLT adopted an analytic syllabus (whereby learners experience language in holistic chunks and analyze the parts as needed) and a focus-on-meaning methodology (stressing comprehensible input, communication, and respect for interlanguage development). However, while communicative classrooms proved more interesting and motivating than their predecessors, CLT fell short in several regards: (1) learners achieved high levels of communicative success without concomitant levels of grammatical accuracy (accuracy not necessary for meaningful communication); (2) even very long periods of exposure to rich input in the target language, as in immersion settings, were insufficient for achieving nativelike ultimate attainment (Swain & Lapkin, 1982); and (3) most formal language education contexts could ill afford the substantial time and resources for learners to benefit maximally from learning through naturalistic processes. Simultaneously, research had begun to identify clear benefits of planned exposure to acquisitional processes including input, interaction, and output; that is, it did seem that learning experiences could be manipulated intentionally to bring about the acquisition of language forms through a variety of instructional techniques (Norris & Ortega, 2000).

The concept of "task" presented an opportunity to consolidate these ideas into an integrated approach which might accomplish a variety of instructional ends. Early recourse to tasks in language teaching focused on ways of bringing "real" communication and learner-centered processes into the classroom, by respecting learners' interests and attending to interlanguage development as it unfolded in the use of language (Breen, 1987; Long, 1985; Prabhu, 1987). Tasks provided motivating communication activities that bore some relevance to language use beyond the classroom. Tasks also had the advantage of offering learners some reason for communicating, beyond practicing to do so, in that they came replete with actual outcomes, criteria for success or failure, even tangible results (getting what you ordered at the restaurant, winning the debate, etc.). In the classroom, tasks enabled teachers and learners to see language development as it unfolded in communication trial and error, thereby establishing a linkage between functions, the language forms that realized them, and the meaning-bearing uses to which they were put.

Building upon these first principles, subsequent attention to tasks incorporated available research findings and extended the scope of considerations (Crookes & Gass, 1993a, 1993b; Skehan, 1996, 1998). It became apparent that certain tasks behaved in more or less predictable ways that could direct learners' attention at

particular aspects of the language in use. Task features (amount of information, interlocutor relationships, etc.), conditions (planning time, stress, etc.), and other characteristics could be designed such that the likelihood of certain kinds of language development was enhanced. Similarly, by challenging learners to use the target language – to negotiate for meaning – teachers were afforded the opportunity to apply a variety of feedback techniques for increasing learners' noticing and awareness of how the target language might realize communicative functions. Critically, this "focus-on-form" methodology (Long, 1991; Long & Robinson, 1998) was intended to occur within the otherwise task-driven communication that was taking place, through brief attention to aspects of communicative need, thereby enhancing the probability that learners would associate the form-focus with the meaning to which it was related.

Ultimately, drawing on several decades of discussion and research, TBLT emerged as much more than a language teaching "method," fun communicative techniques, or the kinds of clever activities that good teachers have always done. At its most complete, TBLT applies available understandings and evidence to the comprehensive design of entire language education programs. Though several task-based architectures have been proposed (e.g, Ellis, 2003; Skehan, 1998; Prabhu, 1987; Willis, 1996), most subscribe to the following elements (Long & Crookes, 1993; Long & Norris, 2000):

- *Needs analysis*: Following fundamental principles of program design (Patton, 1997), the needs that an L2 educational program will meet are first specified, ideally on the basis of a thorough-going analysis of the kinds of language use tasks that learners should be able to accomplish upon completion of a program (Long, 2005b). Needs analysis (Brown, this volume) integrates multiple sources of information from multiple methodological perspectives – including, but not limited to, learners' goals, the values of program constituents, occupational or societal demands, observations of language use situations, etc. – in identifying exactly what learning outcomes will be targeted by the program.
- *Task selection and sequencing*: Based on needs, relevant target tasks and/or task types are articulated to unit, course, and syllabus sequences as befits the program's theory of learner development. Several possible frames may be adopted (e.g., Estaire & Zanón, 1994; Robinson, 2001; Skehan, 1998), though most task-based approaches incorporate both a content rationale (i.e., grouping of tasks by thematic relationships) and a complexity rationale (i.e., locating tasks along a progression from least to most complex). At their most effective, sequencing decisions draw upon intimate knowledge of the learner population, as well as understanding of the communication demands and acquisition opportunities comprised by diverse language use tasks.
- *Materials and instruction development*: Once sequenced, tasks are didacticized through learning experiences that maintain a vision of the target task throughout while incorporating language, procedural, and content knowledge via learners' engagement with pedagogic activities (Chaudron, et al., 2005). Materials and instruction minimally provide: (1) substantial language/content

input through examples of authentic tasks; (2) sustained analytic work on portions of tasks that have been elaborated to facilitate a focus on form; (3) interactive activities, structured and scaffolded in ways that maximize noticing and awareness of form–function–meaning relationships; and (4) target-task performances, emphasizing the transfer of learning to non-instructional settings and providing opportunities for feedback.

- *Teaching*: Teachers facilitate the essential link between instructional materials and their use with learners in classrooms. In task-based teaching, learners must be schematized to the expected procedures and outcomes of tasks, monitored in their performance of tasks, and offered opportunities for enhanced understandings of language use throughout. It is up to the teacher to initiate these task processes flexibly and adjust them as necessary, such that maximal learning may be realized with sensitivity to the ways in which learners actually engage with tasks (Samuda, 2001).
- *Assessment*: Regular assessment of students' task-based learning takes place throughout the delivery of courses and programs to meet diverse purposes (Norris, 2002). Though a variety of instruments and procedures may be necessary, assessment within task-based programs emphasizes the performance of target tasks (as opposed to the demonstration of knowledge about the language), primarily as a mechanism for providing meaningful feedback to learners and teachers, for determining students' abilities with target tasks, and for ensuring an overall focus on target-task learning throughout the program.
- *Program evaluation*: Essential to the ultimate effectiveness of task-based education, and drawing on assessment, as well as other sources of evidence, evaluation enables language educators to understand, improve, judge, and otherwise ascertain that all of the above elements are functioning conjointly in support of targeted outcomes. In task-based programs, evaluation focuses on the relevance of target tasks for learners, the appropriateness of sequencing decisions and L2 acquisition expectations, the effectiveness of materials, the preparation and support of teachers, and the validity of assessments vis-à-vis the interpretations and uses to which they are put. Evaluation also provides a programmatic frame of reference within which observations about task-based teaching and learning may be thoroughly understood.

While introducing tasks in existing language education programs may offer some benefits (see notions of "task-supported" teaching in Samuda & Bygate, 2008; Ellis, 2003), the full advantages of task-based learning are not likely to be realized outside of a programmatic commitment to task as the basis for educational design. Where any one of the elements above occurs in the absence of the others the ultimate effectiveness of task-based learning will likely break down, a consideration often missed by TBLT critics. However, where a long-term commitment is made to the development of task-based language teaching *programs* that incorporate these elements (e.g., Van den Branden, 2006), evidence suggests that language and task abilities advance in tandem, learners achieve expectations, and language teaching evolves into a potentially more meaningful endeavor.

Task-Based Teaching

While vital planning and deployment of resources occur during curriculum design, the actual implementation of language learning occurs during task-based teaching, that is, what teachers and learners do with tasks in the classroom (or other venue). Highly structured approaches (e.g., Willis, 1996) may provide immediately useful ways of framing some classroom work, but task-based teaching calls upon more than just pre-, during-, and post-task activities or other formulae. With respect to the actual range of tasks that may be targeted for instruction, it is likely that flexible implementation of general methodological principles (Long, this volume) will be required to meet the needs of diverse contexts and learners. In general, though, several phases of classroom work highlight what teachers and learners do during task-based teaching (see the complete example in Chaudron, et al., 2005).

A *task input* phase typically initiates the teaching sequence, preceding any pedagogic activities. Building from the idea that exposure to language in use can facilitate considerable amounts of L2 acquisition, the input phase introduces the target task as it is realized in actual communication. Repeated viewing of video segments, observing live performances, reading texts, and other techniques enable the presentation of a full-fledged target task without manipulation. By seeing what they will do in its entirety, learners become motivated and begin to establish essential linkages to the contexts in which the target task occurs. By engaging receptively with the task, learners begin to focus their attention on trying to understand what is being said or written, thereby initiating their noticing of what forms are used in what ways. In addition to the incidental acquisition of some forms, through association with the physical and linguistic context of the task, learners also begin to identify gaps in their individual L2 repertoires. During this input phase, the teacher acts as provider of input, either by performing tasks or by presenting various recorded task exemplars (and teachers play an important role in making sure the tasks are appropriate to the learners). Teachers avoid didactic treatment of the tasks at this point, however, so that learners are afforded maximal opportunity to attend to the contextual, content, linguistic, functional, and other features that combine to form the holistic target task.

Building from thorough exposure to the target task, *pedagogic task work* ensues. It is during this phase that tasks are segmented, elaborated, and otherwise manipulated by the teacher, with the objective of raising learners' awareness of new language forms and their use for particular functions. Typically, this phase features multiple iterations of work on increasingly complete versions of the target task, often moving from comprehension to production. Early stages of task work emphasize important form–function relationships, through input enhancement (e.g., textual or oral emphasis, focused listening) or through learner analysis of task discourse (e.g., identification of known and unknown vocabulary). Interactive tasks in pairs or small groups then require learners to utilize existing and new language. Such tasks present learners with information gaps to be overcome,

problems to be solved, decisions to be made, or otherwise meaningful reasons to interact with each other, that negotiation of meaning often leading to communication breakdown and the opportunity for self-, other-, or teacher-initiated feedback. Feedback episodes, then, play a key role during this phase, as a primary mechanism for focus-on-form. Different kinds of feedback, such as recasts, models, brief grammatical explanations, and others can be deployed as a way of fostering greater awareness and incorporation of language forms where they are particularly needed. Subsequent introduction of distinct task types or conditions also may help learners to focus differentially on features of their language use (e.g., accuracy, complexity, fluency; Skehan, 1998). For example, the introduction of planning time prior to a story-telling task (Ortega, 2005) may allow learners the cognitive space to incorporate newly acquired forms into a syntactically more complex narration.

The didactic possibilities available for this phase are considerable (Ellis, 2003), as are the language forms to which they may be directed (vocabulary, syntax, pronunciation, pragmatics, discourse structure, etc.). A key responsibility of teachers is to select a variety of pedagogic tasks that respond to potentially diverse learner types (e.g., with different aptitudes or motivational complexes, see Robinson, 2001). Furthermore, teachers must ensure that learners understand the purposive, procedural, and goal-oriented nature of pedagogic tasks; teacher framing and scaffolding is essential in this regard, and without it, learners will do unrelated things. In addition, perhaps most critically, teachers monitor both task processes and learner language use, and they must ultimately arbitrate the appropriate occasions for feedback interventions.

Following what may be multiple iterations of pedagogic task work, *target task performance* calls upon learners to deploy what they have learned in doing the target task. Depending on learners and curriculum, of course, this doing may involve anything from a brief communicative transaction (e.g., ordering a meal, writing an e-mail message, following directions on a map) to longer pieces of work (e.g., taking accurate lecture notes, following a manual to repair an engine part) to extended performances (e.g., participating in a debate, teaching a class). The target task performance also may be staged to facilitate learner attention to outcomes that aggregate along the way to accomplishing complex tasks. For example, in preparing for an academic presentation, students may engage in title and abstract writing, assembling a handout, preparing audiovisual materials, drafting presentation notes, and then putting the elements together into the full presentation. The teacher's job at this point is to replicate the conditions under which the target task will be performed, providing an audience (simulated or real), reproducing constraints (physical surroundings, available resources), introducing authentic complications (interlocutor questions, disturbances, criteria to be met), and so on. By "going it alone" with a target task *in situ*, learners practice the use of language for meaningful purposes, thereby engaging their developing cognitive, motivational, linguistic, content-knowledge, and other resources under conditions that vary from the safety of structured pedagogic tasks. In doing so, they also come to understand the range of competing factors (stress, interlocutor

reactions, unexpected interruptions) that constitute purposeful language use, and they extend their language and content learning to incorporate strategies for dealing with actual communication (e.g., circumlocution, questioning, repetition).

Often, a *task follow-up* phase can enhance the learning that has taken place over previous phases of the task-based cycle. Thus, the target task performance establishes a key watershed point at which teachers and learners can reflect on task success or failure, performance strengths, gaps that remain in language/content/ task knowledge, and related concerns, all leading to instructional decisions regarding what features are in need of subsequent repetition or expansion. Here, teacher guidance is again essential for identifying salient gaps and providing relevant feedback; follow-up activities offer another chance for focus on form to take place with the added cognitive consonance provided by doing the full task. The principle to keep in mind is that focus-on-form works best when there is a "focus"; an essay covered in red ink identifying all "errors" may not prove as effective as the careful highlighting of selected form–function issues that bear relevance to communicative success and the features targeted during preceding pedagogic task phases. Procedurally, the follow-up phase may involve a variety of activities, such as teacher presentation of common patterns that emerged across students, learner analysis of their own and others' performances (e.g., in video-taped or written format), and individual or group work on refining and repeating task performances.

Several aspects of this task-based teaching cycle are in need of inquiry. As TBLT is applied across diverse learner populations, it will be important to investigate what among the available task-based techniques prove the most effective with particular learner types, including, for example, young learners, learners with differing motivations, learners with little or no literacy, and learners at differing L2 proficiency levels. At a minimum, researchers will need to describe or measure learner types to enable accurate generalizations/limitations regarding their findings (Norris & Ortega, 2003). Additionally, along with diverse learners come diverse tasks, and it is as yet uncertain to what extent available techniques enable acquisition of abilities that underlie a variety of potential target tasks. For example, many of the researched pedagogic tasks involve spoken interaction, but whether this mode of teaching leads to efficient learning of highly literate target L2 uses remains little understood. Finally, it should be clear that much more goes on in the task-based classroom than simply turning learners loose on tasks (or vice versa). Teachers play an essential role throughout the task lesson cycle, motivating, schematizing, scaffolding, monitoring, intervening, and so on. It is critical that research examine exactly what training, resources, and support teachers need in order to meet this array of expectations.

Task-Based Testing

Though often narrowly construed as "testing," task-based assessment incorporates a variety of instruments and procedures (not just tests) for gathering data

about student learning, in order to provide an empirical basis for decisions and actions that must occur in education. Tasks have captured the attention of testers for some time (e.g., Cureton, 1951), because they present goal-oriented, contextualized challenges that prompt examinees to deploy cognitive skills and domain-related knowledge in authentic performance rather than merely displaying what they know in selected-response and other discrete forms of tests (Kane, 2001; Wiggins, 1998). For language testing, in particular, recent interest in task reflects the need to incorporate language use into assessments, such that interpretations about learners' abilities to communicate are warranted (Brindley, 1994; Norris et al., 1998). While language testers have concerned themselves primarily with psychometric issues of what gets tested in task-based performance, much less attention has been afforded more fundamental questions regarding why and how task-based assessments are being used in language education settings (Norris, 2002). At stake for educators is not whether task-based assessments accurately represent theoretical language ability constructs; rather, as Cureton (1951) pointed out, it is "how well a test does the job it was employed to do" (p. 621).

How are assessments used in task-based programs and related settings? While intended uses vary in the details, as befits the distinct users of assessment information, several common purposes for assessment have emerged over the past decade or so of experiences with TBLT. On the one hand, *summative* assessments incorporate specific target tasks or task types for making often high-stakes judgments about learners and the programs that educate them (e.g., Gysen & Van Avermaet, 2005). For example, certification of learners' task abilities is essential for jobs that require extensive communication, such as air traffic controllers, interpreters/translators, or international teaching assistants. Additionally, certification of task abilities may play a broader role in providing learners and the public with information on the basic social domains within which a language user can function. Judgments regarding whether learners can meet minimum task requirements also are used for making program acceptance or placement decisions, as in second language vocational training settings. Finally, tasks may be incorporated into assessments for the purpose of holding programs accountable to educational standards or for washing back onto the ways in which language teaching is occurring. Common to these summative uses for tasks-based assessment are particular qualities: (1) dependence on representative tasks that can be trusted to reflect language use in actual targeted domains (general or specific); (2) replication of authentic task performance conditions and criteria; and (3) consistency in administration and reliability in rating, scoring, or otherwise judging task performances.

Assessment may also provide constituents in task-based classrooms and programs with the kinds of rich information needed to support experiential learning and foster students' abilities to do things with language. These *formative* uses for assessment call upon tasks to enable a close understanding of learner development and facilitate the provision of relevant feedback to teachers, learners, curriculum developers, and others (e.g., Byrnes, 2002). Thus, performance of curriculum- and lesson-embedded tasks throughout teaching sequences provide learners the

opportunity to demonstrate what language and content features they are acquiring and how well they can put that knowledge to use. These assessments serve an important motivational and awareness-raising function, giving learners a clear target for learning and pushing them to do so. They also offer a frame of reference within which teachers and learners can observe, reflect on, judge, and otherwise understand the effectiveness of pedagogic activities and learning processes. Within the task frame, specific feedback can be provided to individual learners about particular aspects of task performance and L2 learning, thereby helping them to help themselves, and patterns of performance across students offer important feedback to teachers for the adaptation of materials and instruction. Furthermore, the accumulation of task-based assessment data over longer term instructional sequences offers an invaluable basis for illuminating actual patterns of learner language development in comparison with curricular expectations. It is this combination of process (L2 development) and product (task performance) data that most distinguishes formative from summative task-based assessments. Formative uses for assessment prioritize certain characteristics: (1) close articulation of tasks to curricular sequences and learning expectations; (2) thorough and accessible guidelines for task performance expectations and assessment criteria; (3) the use of language profiles, analytic rating scales, and other information-rich mechanisms for providing meaningful feedback about diverse L2 and task features; (4) multiple iterations of teacher (and potentially learner self-) assessment at different stages of task completion; and (5) frameworks for tracking and interpreting important aspects of learner development over time.

Several concerns at the interface of tasks, TBLT, and assessment would benefit from further empirical attention. To date, related inquiry has been driven largely by the question of "what constitutes a task-based language assessment?", adhering to a monolithic model of testing that seeks primarily to inform interpretations about theorized language proficiency constructs. A more educationally relevant approach might be to adopt a "validity evaluation" model (Norris, 2008), which derives research priorities from the actual uses to which assessments are put, the contexts in which they are used, and the individuals or groups who are using them. In conjunction with summative uses for assessment, key areas in need of attention include the extent to which performance on assessment tasks can be assumed to extrapolate to other tasks beyond the test setting, the relevance of rating scales and criteria for the kinds of high-stakes interpretations that are being made about learners, and the impact that introduction of tasks into summative tests may be having on educational systems (including learner, teacher, and public attitudes toward language education). With more formative uses of assessment, inquiry will illuminate which tasks help reveal learners' emerging accuracy, complexity, and fluency with various language forms, and it will disentangle a variety of task-setting and learner individual-differences factors from language features of interest. Research is also much needed into classroom diagnostic and feedback frameworks that can help teachers to efficiently analyze task performances, turn those analyses into maximally useful understandings of learner development, and communicate with students in terms that help them to learn.

Challenges and Opportunities for Task-Based Language Teaching

Though contested at times, task-based principles have begun to gain purchase in a variety of educational settings, and TBLT shows signs of crystallizing into a robust domain of inquiry, with an international conference series (www.tblt.org), a book series (*Task-based language teaching: Issues, research, and practice*), and a healthy literature. However, the extent to which TBLT offers sustainable solutions – in an increasingly volatile policy and practice environment – will be conditioned by the ways in which researchers, educationalists, and practitioners meet several critical challenges to the application of task-based ideas.

A top priority must be the transformation of task-based research into an educationally relevant endeavor. While indirect value may be gained from research that focuses on the generation or testing of theories about instruction and acquisition, it is not the case that primary studies of a few task features or learner conversations will, on their own, provide much in the way of warranted implications for task-based teaching in practice. To be sure, theory-driven primary studies can contribute isolated bits of information which, once sufficient studies have accumulated (Norris & Ortega, 2006), may shed light on fundamental questions about tasks and instructed L2 acquisition (e.g., the effectiveness of certain task design features; see Keck et al., 2006). However, alternative epistemologies are required where research seeks to illuminate organic questions of interest to those who are responsible for making language learning happen (teachers, administrators, curriculum developers). Language education in practice, rather than theory, is an extremely complex and multivariate undertaking, and research that informs education must adopt methodologies that reflect the full scope of what is going on; to do otherwise is to sell short the task-based proposal from the outset by focusing on only selected parts of what is intended to be a holistic enterprise.

What methods might work? Certainly, mixed-methods designs (Cresswell & Clark, 2007) will prove more effective than methodologically constrained designs (e.g., conversation analysis, which only derives evidence from observable discourse), in that they account for a complex array of factors that interact in any task-based episode (e.g., teacher, learner, task, and setting variables). Similarly, longitudinal studies (Ortega & Iberri-Shea, 2005) of task-based instruction will be essential for capturing just how teaching and learning happen; given that task-based proposals are intended for the level of multiple lessons, units, full courses, and cross-curricular sequences, it seems clear that *only* longitudinal studies looking across these periods of instruction will illuminate task-based development. Finally, most directly relevant for educational decision-making will be the application of program evaluation to the pragmatic resolution of questions as they arise in the delivery of full TBLT programs. While theoretical research informs *conclusion*-oriented inquiry, evaluation operationalizes *decision*-oriented inquiry (Cronbach & Suppes, 1969). Adopting ongoing evaluation will enable decision-makers to address priority questions about elements of task-based programs,

gather information that is immediately relevant, turn that information into decisions and actions, and then pose new questions. An excellent example of this iterative approach can be found in ten years of evaluative work on task-based teacher development programs in Dutch language education throughout Belgium (Van den Branden, 2006).

A second major challenge is identifying meaningful starting and ending points to language education classes and programs. From a task-based perspective, needs analysis provides an empirical response, offering triangulated understandings of what target tasks are most relevant for particular groups of learners and what the associated language/content expectations may be. Unfortunately, task-based needs analysis has suffered frequent mis-understanding among practitioners and mis-representation in the literature. Common reactions include claims that learners cannot be expected to know their own needs, learners do not have any apparent uses for the language, the instrumental focus of needs undermines the broad value of language education, and similar.

To rectify these debilitating perceptions, it behooves the TBLT community to encourage greater consistency vis-à-vis the nature of needs analysis and the language tasks that may ensue. First, it is absolutely critical to clarify that needs are what educational and social programs are designed to meet; without needs of some sort, programs are essentially pointless. Second, needs analysis does not rely on learners to identify their own language use situations and tasks, though there is a didactic role to be played by learners thinking about how they might use language. Good educational needs analysis does take the learners as the point of embarkation, but it incorporates multiple sources of information and methodologies of data collection to elicit insights into the kinds of language use that learners will confront (Long, 2005a). Third, needs might issue equally from the target tasks that a group of language educators hold valuable for their program and from analysis of how those tasks are accomplished in language use situations. The point is that some effort is made to justify what the language program is targeting – where ability to do things with the language is a major expectation (as it surely must be in language education), then it is essential to establish just what those things are. Fourth, it must be clear that *any* of the valued things that people do with language can be identified as the needs addressed by a language program, including survival tasks, mundane job tasks, tasks that enable access to education, and creative artistic tasks, to name a few (see diverse needs in Van Avermaet & Gysen, 2006). It will be particularly important to help the language teaching community understand this last point, as it is these truly valued, yet varied, goals of doing things with language that serve to define the potential contribution of task-based programs. In the end, just as TBLT polices its own discourse, it also will be worth posing the same question to all language programs: What are the needs to which your efforts are targeted, and to what extent do they offer defensible goals for language education?

A third challenge issues from the simple fact that innovation – planned and executed change – will be required if the purported benefits of TBLT are to be realized. Of particular salience is how TBLT can be implemented in light of both

the cultural traditions and the teaching practice dimensions of innovation (Carless, 2004; Markee, 1996). Though little researched, much has been asserted about the incongruence of task-based ideas with non-Western societies, foreign language instruction, and other educational "cultures." Large class sizes, novice or young learners, minimal contact hours, classroom power hierarchies, non-interactiveness of students, exam-driven learning, teacher intransigence or apathy, and other impediments are regularly cited as justification for not engaging in TBLT. Certainly, cultural values, institutional constraints, and histories of practice will shape the possibilities that any approach to language teaching may realize, and it is important to enter into innovation gradually and with sensitivity to what is possible and what is not (Carless, 2007). Nevertheless, several countervailing realities speak to the viability, even the necessity, of TBLT-oriented innovation.

For one, recent educational policies imply that innovation will be necessary if the value of language learning is to be realized. For example, countries in East Asia and the European Union have witnessed the rise of policies emphasizing the need for learners to develop communicative abilities that enable them to do things with language, and "task" has figured explicitly into related documents (e.g., Council of Europe, 2001; Curriculum Development Council, 2002). Likewise, in re-structuring language curriculum to target communicative abilities in these contexts, it has become apparent to educationalists that L2 acquisition does generally call upon predictable cognitive processes and that these may be maximized through a context that features rich input, purposeful interaction, pushed output, and related techniques characteristic of task-based instruction (e.g., Izumi, 2003). Furthermore, for a variety of reasons, language teachers in supposedly unlikely contexts are incorporating tasks to bring about change. For example, teachers in Hong Kong and Japan (e.g., Flowerdew, 2005; Watanabe, Konoeda, & Mochizuki, 2007) have introduced tasks into the language classroom to raise learners' critical awareness of their own language needs in spite of institutionalized traditions. In foreign language education (e.g., Byrnes et al., 2006; González-Lloret, 2003; Leaver & Willis, 2004) teachers have turned to tasks as a way of operationalizing language acquisition in tertiary classes that both respect adult learners and push them to advanced abilities. Even with young learners (e.g., Van Gorp & Bogaert, 2006), teachers have found tasks to be effective ways of developing learners' language and cognitive abilities while maintaining interesting and motivating classroom environments.

While these examples testify to the viability of task-based innovation, it remains an empirical question as to why and how innovation occurs and the requirements for educational systems and actors. Large-scale inquiry (e.g., policy analysis, program evaluation) should address the interrelationships between educational policy, resource demands, institutional structures, and cultural patterns that encourage or constrain innovation. Smaller-scale teacher-driven inquiry (e.g., action research) will shed needed light on the central role played by teachers in the success or failure of any innovation, providing grounded evidence regarding how teachers continue to learn, professional demands for doing so, values that motivate teachers, and processes that facilitate change.

As we consider the possibilities and limitations of task-based innovation, and as we move ahead with the evolution of language teaching in response to an ever-changing social and educational environment, it will pay to keep a few basics in mind. Communicating is a fundamental fact of human cultures, and learning to communicate – learning to use language to do things – in another language offers considerable value from instrumental, aesthetic, moral, cognitive, and other perspectives. A basic question that must be answered by all language educators, then, is "what do our learners learn how to do, and how do they learn to do it"? Task offers a helpful way of encapsulating the things humans tend to do with language, in particular because it emphasizes the functional sense of language use. Without that impetus toward use, language reverts to a body of knowledge to be apprehended, a canon of great words, but not a particularly functional (or essential?) ability. Task also provides a useful frame of reference, directing teachers and learners toward a purpose for communicating, affording contextualized meaning to language forms, and indicating starting and ending points to the communicative effort. It is within this frame that task also helps to operationalize the fundamental point at which language learning occurs, as learners become aware of the language forms that provide particular meanings appropriate for achieving particular functions.

At the same time, task comes with a price. It counters our traditions of practice, requires rethinking the outcomes of our programs, and implies an overhaul of the teaching and testing that is going on in many language classrooms. Task challenges us to respect (and investigate) the actual uses to which language is put in diverse cultures and discourse domains, rather than adopting a one-size-fits-all approach to language. Task also demands that we think about language use and language learning in holistic terms. To achieve the benefits of task-based practice, we must first accept that language develops not as accretion of discrete bits of knowledge but through a series of holistic experiences. That holism also translates into the development of language education programs. Thus, just as task instantiates and facilitates the organic interaction of form, meaning, and function, so too do task-based programs enable learning on the basis of a holistic educational ecology wherein curriculum, instruction, assessment, and other processes interact to bring about learning. Though challenging, this vision of task-based language learning offers one comprehensive alternative to the status quo. That alternative may help us realize a language education practice that is valuable to a variety of discourse communities, the language learners who need and want to interact with them, and the language teaching profession whose job it is to facilitate their doing so.

REFERENCES

Breen, M. (1987). Learner contributions to task design. In C. Candlin & D. Murphy (eds.), *Language learning tasks* (pp. 23–46). Englewood Cliffs, NJ: Prentice-Hall.

Brindley, G. (1994). Task-centred assessment in language learning: The promise and the challenge. In N. Bird, et al. (eds.), *Language and learning: Papers presented at the Annual International Language in Education Conference* (Hong Kong, 1993) (pp. 73–94). Hong Kong: Hong Kong Education Department.

Byrnes, H. (2002). The role of task and task-based assessment in a content-oriented collegiate foreign language curriculum. *Language Testing* 19, 419–37.

Byrnes, H., Crane, C., Maxim, H., & Sprang, K. (2006). Taking text to task: Issues and choices in curriculum construction. *ITL – International Journal of Applied Linguistics* 152, 85–109.

Carless, D. (2004). Issues in teachers' re-interpretation of a task-based innovation in primary schools. *TESOL Quarterly* 38, 639–62.

Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System* 35, 4, 595–608.

Chaudron, C., Doughty, C., Kim, Y. et al. (2005). A task-based needs analysis of a tertiary Korean as a foreign language program. In M. H. Long (ed.), *Second language needs analysis* (pp. 225–61). Cambridge: Cambridge University Press.

Council of Europe (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cresswell, J. & Clark, V. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Cronbach, L. & Suppes, P. (eds.) (1969). *Research for tomorrow's schools: Disciplined inquiry of education*. New York: Macmillan.

Crookes, G. & Gass, S. (eds.) (1993a). *Tasks and language learning: Integrating theory and practice*. Clevedon, UK: Multilingual Matters.

Crookes, G. & Gass, S. (eds.) (1993b). *Tasks in a pedagogical context: Integrating theory and practice*. Clevedon, UK: Multilingual Matters.

Cureton, E. E. (1951). Validity. In E. F. Lingquist (ed.), *Educational measurement* (pp. 621–94). Washington, DC: American Council on Education.

Curriculum Development Council (2002). *English language education: Key learning area curriculum guide (primary 1–secondary 3).* Hong Kong SAR, China: Government Printer.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: Henry Holt.

Dewey, J. (1938). *Experience and education*. New York: Macmillan/Collier.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Estaire, S. & Zanón, J. (1994). *Planning classwork: A task-based approach*. Oxford: Oxford University Press.

Flowerdew, L. (2005). Integrating traditional and critical approaches to syllabus design: The "what", the "how", and the "why?". *Journal of English for Academic Purposes* 4, 135–47.

González-Lloret, M. (2003). Designing task-based call to promote interaction: En busca de esmeraldas. *Language Learning & Technology* 7, 86–104.

Gysen, S. & Van Avermaet, P. (2005). Issues in functional language performance assessment: The case of the Certificate Dutch as a Foreign Language. *Language Assessment Quarterly* 2, 51–68.

Izumi, S. (2003). EFL education in Japan from the perspective of second language acquisition research. *Sophia Linguistica* 50, 3–14.

James, W. (1899/1958) *Talks to teachers*. New York: W. W. Norton and Company.

Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement* 38, 319–42.

Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Philadelphia: John Benjamins.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Krashen, S. & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Leaver, B. L. & Willis, J. R. (eds.) (2004). *Task-based instruction in foreign language education: Practices and programs.* Washington, DC: Georgetown University Press.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language training. In K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition* (pp. 77–99). Clevedon, UK: Multilingual Matters.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamins.

Long, M. H. (2005a). Methodological issues in learner needs analysis. In M. H. Long (ed.), *Second language needs analysis* (pp. 19–76). Cambridge: Cambridge University Press.

Long, M. H. (ed.) (2005b). *Second language needs analysis*. Cambridge: Cambridge University Press.

Long, M. H. & Crookes, G. (1993). Units of analysis in syllabus design: The case for task. In G. Crookes & S. Gass (eds.), *Tasks in a pedagogical context: Integrating theory and practice* (pp. 9–54). Clevedon, UK: Multilingual Matters.

Long, M. H. & Norris, J. M. (2000). Task-based teaching and assessment. In M. Byram (ed.), *Encyclopedia of language teaching* (pp. 597–603). London: Routledge.

Long, M. H. & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. J. Doughty & J. Williams (eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge: Cambridge University Press.

Markee, N. (1996). *Managing curricular innovation*. Cambridge: Cambridge University Press.

Norris, J. M. (2002). Interpretations, intended uses and designs in task-based language assessment. *Language Testing* 19, 4, 337–46.

Norris, J. M. (2008). *Validity evaluation in language assessment*. New York: Peter Lang.

Norris, J. M., Brown, J. D., Hudson, T. D., & Yoshioka, J. (1998). *Designing second language performance assessment*. Honolulu: University of Hawai'i Press.

Norris, J. M. & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50, 417–528.

Norris, J. M. & Ortega, L. (2003). Defining and measuring SLA. In C. J. Doughty & M. H. Long (eds.), *Handbook of second language acquisition* (pp. 716–61). Oxford: Blackwell.

Norris, J. M. & Ortega, L. (eds.) (2006). *Synthesizing research on language learning and teaching*. Philadelphia: John Benjamins.

Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (ed.), *Planning and task performance in a second language* (pp. 77–109). Philadelphia: John Benjamins.

Ortega, L. & Iberri-Shea, G. (2005). Longitudinal research in SLA: Recent trends and future directions. *Annual Review of Applied Linguistics* 25, 26–45.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Reid, W. J. (1997). Research on task-centered practice. *Social Work Research* 21, 3, 132–7.

Robinson, P. (2001). Task complexity, cognition and second language syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge: Cambridge University Press.

Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan, & M. Swain (eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 119–34). Harlow, UK: Longman.

Samuda, V. & Bygate, M. (2008). *Tasks in second language learning*. New York: Palgrave Macmillan.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics* 17, 38–62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Sternberg, R. J. (2003). What is an "expert student?". *Educational Researcher* 32, 8, 5–9.

Swain, M. & Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study*. Clevedon, UK: Multilingual Matters.

Van Avermaet, P. & Gysen, S. (2006). From needs to tasks: Language learning needs in a task-based approach. In K. Van den Branden (ed.), *Task-based language education: From theory to practice* (pp. 17–46). Cambridge: Cambridge University Press.

Van den Branden, K. (2006). Training teachers: Task-based as well? In K. Van den Branden (ed.), *Task-based language education: From theory to practice* (pp. 217–73). Cambridge: Cambridge University Press.

Van Gorp, K. & Bogaert, N. (2006). Developing language tasks for primary and secondary education. In K. Van den Branden (ed.), *Task-based language education: From theory to practice* (pp. 76–105). Cambridge: Cambridge University Press.

Virjo, I., Holmberg-Mattila, D., & Mattila, K. (2001). Task-based learning (TBL) in undergraduate medical education. *Medical Teacher* 23, 1, 55–8.

Watanabe, Y., Konoeda, K., & Mochizuki, N. (2007). Task-based critical pedagogy in JHS EFL. Paper presented at the JALT International Conference on Language Teaching and Learning, Tokyo, Japan.

Widdowson, H. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.

Wilkins, D. (1976). *Notional syllabuses. A taxonomy and its relevance to foreign language curriculum development*. Oxford: Oxford University Press.

Willis, J. (1996). *A framework for task-based learning*. Harlow, UK: Longman.

Wright, T. (2000). No more pencils. No more books? Arguing for the use of experiential learning in post secondary environmental studies classrooms. *Electronic Green Journal*, Dec 2000, n.p.