

This is an Author's Accepted Manuscript of an article published in *Language Assessment Quarterly*, 21 May 2014, © Taylor and Francis.

Harding, L. (2014). Communicative language testing: Current issues and future research. *Language Assessment Quarterly*, 11(2), 186–197. DOI: 10.1080/15434303.2014.895829

Available online: <http://dx.doi.org/10.1080/15434303.2014.895829>

Communicative language testing: Current issues and future research

Luke Harding
Lancaster University

Abstract

This article discusses a range of current issues and future research possibilities in Communicative Language Testing (CLT) using, as its departure point, the key questions which emerged during the CLT symposium at the 2010 Language Testing Forum (LTF). The paper begins with a summary of the 2010 symposium discussion in which three main issues related to CLT are identified: (1) the “mainstreaming” of CLT since 1980, (2) the difficulty for practitioners in utilising and operationalising models of communicative ability, and (3) the challenge of theorising a sufficiently rich communicative construct. These issues are each discussed and elaborated in turn, with the conclusion drawn that, while the communicative approach lies dormant in many test constructs, there is scope for a reinvigorated communicative approach which focuses on “adaptability”. A number of future research directions with adaptability at the forefront are proposed.

1. Introduction

The Communicative Language Testing (CLT) symposium within the 2010 Language Testing Forum was notable for the lack of explicit focus on its intended topic.¹ The session was originally designed to reflect the symposium held in 1980 at the very first Language Testing Forum, where discussion centred on a position paper written by Keith Morrow (1979) as well as several responses to that paper which had been circulated prior to the meeting (see Alderson & Hughes, 1981).² In 2010, the structure was slightly different; discussion followed papers presented by Dianne Wall and Cathy

¹ The plenary discussion of the Language Testing Forum 2010 Communicative Language Testing symposium was recorded and summarised, and this summary is drawn on in the current paper.

² The Alderson & Hughes (1981) publication is available for download at:
<http://www.ling.lancs.ac.uk/groups/ltrg/ltf2010.htm>

Taylor, who provided an exam board perspective on CLT (this issue), and by Cyril Weir who provided an academic review. The plenary discussion lasted for around one hour, and the bulk of this concerned the socio-cognitive framework for test development and validation which formed the core of Weir's presentation, but which is not exclusively associated with a communicative approach (see Weir, 2005; Weir & O'Sullivan, 2011). There were questions from the participants about the distinction between construct validity and Weir's notion of "cognitive validity"; whether learners all progress through the same developmental stages; and whether process should be considered as important as product in a test setting. These more general issues all garnered debate and there was little consensus by the end of the session.

There were, however, some questions which were more clearly oriented towards the "communicative" focus of the symposium. For example, a general question was raised in relation to Weir's framework concerning how practitioners might make use of various models of communicative ability when these are often highly sophisticated and require interpretation. There was also a discussion of how "unpredictability" was dealt with in the Trinity GESE (see Wall & Taylor, this issue), and whether a candidate's charm or self-assurance might impact on test scores. Finally, in one of the concluding remarks of the symposium, there was an observation by Keith Morrow that the very general nature of the discussion might signal that the term "communicative" has now "faded gracefully into history: that it is no longer a burning issue, and that this is a good thing".

Viewed more broadly, these three episodes of the discussion can be understood as representing important questions and challenges that remain for CLT. In the case of Morrow's statement, it is worth considering whether CLT has become the unremarkable norm since the late 1970s/early 1980s when it was considered an innovative approach to language assessment, and whether the term "communicative" still has any particular resonance in the field. The question regarding the application of Weir's framework points to a more general difficulty experienced by exam boards or smaller testing enterprises in interpreting and applying complex communicative frameworks in test design and validation. Finally, the discussions of unpredictability in the testing situation, and the impact of charm and personality on judgements of performance, point to a concern over whether current communicative constructs are sufficiently rich – a question which includes within its scope whether theories of communicative ability have kept up to date with the rapidly changing nature of communication in the early 21st century.

These three broad issues will provide the focus of this paper. I will firstly chart developments in communicative language testing from 1980 until recently, with a view to establishing the foundations of a "mainstream" communicative approach. I will then review literature related to two current challenges for communicative language testing: (a) utilising and operationalising models of

communicative ability, and (b) theorising a sufficiently rich construct. Finally, a number of future research directions are proposed. Throughout, specific reference will be made to relevant parts of the CLT symposium discussion at the Language Testing Forum (LTF) 2010.

2. Foundations of the “mainstream” communicative approach

Morrow’s observation that the term “communicative” has faded gracefully into history suggests that CLT has become the unremarkable norm in test design. On face value, this seems to be true. A “communicative” language test is often understood to be a test where language is assessed in context, often involving authentic tasks and interaction (see, for example, Davies et al., 1999), and there are numerous tests both within institutions and in the global marketplace which subscribe to this approach. In fact, Fulcher (2000) has suggested that it would be difficult to market large-scale tests which did not profess a communicative orientation (p.493). TOEFL iBT, for example, describes its construct as “communicative competence—the ability to put language knowledge to use in relevant contexts” (2011, p. 4³). Communicative language competence also forms a fundamental basis for the Common European Framework of Reference (see Council of Europe, 2001). A communicative approach towards language testing is the implicit orientation for many well-known texts in the field used on training courses such as Alderson, Clapham and Wall, (1995), Hughes (2003) and those books which feature in the Cambridge Language Assessment Series. In short, there is strong evidence to suggest that the communicative approach has become the dominant paradigm in modern language testing. This section will outline the development of CLT over the past 30 years with the aim of demonstrating how CLT has been “mainstreamed”, and in what form, before returning to evaluate Morrow’s observation.

Morrow’s (1979) paper “Communicative language testing: Evolution or revolution”, which was a focus of discussion at the 1980 LTF, represented one of the first detailed outlines of CLT.⁴ In this paper, Morrow described the characteristics of a particular type of communicative test: (1) that it be criterion-referenced (and linked to authentic tasks), (2) that it be validated against the abilities it claims to measure; (3) that it not emphasise too heavily quantitative/numerical scores, and (4) that it embrace face validity, possibly at the expense of reliability. Points (3) and (4) were bold and radical claims then, and they still appear to be so today (although it is worth noting that the depiction of communicative testing here may fit most closely with current approaches to task-based testing [see, for example, Norris, 2009]). The most enduring points, however, seem to have emerged from the list of features which Morrow perceived as missing from conventional tests of the day, and

³ http://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v6.pdf

⁴ It is acknowledged that communicative language testing did not begin in 1980; however the publication of the Morrow paper, and its centrality to the LTF forum, make it a useful departure point.

which form the organising framework of Wall and Taylor's paper in this issue. Fulcher (2000) summarises these, suggesting that the "buzzwords of early communicative language testing" were: "real-life tasks", "face validity", "authenticity" and "performance" (p.484).

Morrow's own contribution was soon to be eclipsed, however, by the landmark publication of Canale and Swain's (1980) "Theoretical bases of communicative approaches to second language teaching and testing", which appeared in the first issue of *Applied Linguistics*. In this article, a definition of communicative competence was put forward, drawing heavily on the early work of linguistic anthropologist Dell Hymes (though it has been described as a much reduced version of this original model by McNamara, 1996 – see discussion below). Under Canale and Swain's model, communicative competence involved three sub-competences: grammatical competence, sociolinguistic competence and strategic competence. In articulating this definition, Canale and Swain effectively put flesh to the bones of the communicative approach by providing the whole enterprise with a theoretical framework. Canale (1983) built on this model and revised it, before the publication of Bachman's (1990) *Fundamental considerations in language testing* – which contained Bachman's model of communicative language ability – provided a watershed moment for the narrative of communicative language testing. Bachman's model provided a more detailed picture of the skills which underlie communicative language use based around three main components: language competence, strategic competence and psychophysiological mechanisms. It subsumed sociolinguistic competence under language competence (via "pragmatic competence"), and provided a more comprehensive description of strategic competence than was found in Canale and Swain (1980).

It is beyond the scope of this paper to discuss the Bachman model at length; this has been covered well elsewhere (see, for example, McNamara, 1996; Fulcher & Davidson, 2007). However Bachman (1990) made another key contribution in his text, delineating the difference between a real-life (RL), task-driven approach and an interactional authenticity (IA) approach to CLT. Bachman defines interactional authenticity as "a function of the interaction between the test taker and the test task" (1990, p.317). Importantly, he contrasts this with the "real-life authenticity" approach – which was the basis of Morrow's "Promised Land" vision – in the way in which these concepts are operationalised in test design: "rather than attempting to replicate actual instances of non-test language use, the IA approach uses a theoretical framework of factors affecting test performance to construct tests that include features of language use that are relevant to both the interpretations and uses to be made of test scores" (p.317). In other words, there is a separation here between the

desire to measure communicative language ability, and the need for real-life tasks, high in face validity, in order to elicit evidence of this ability.⁵

The idea that simply selecting tasks which mimic real-life language activities is not sufficient to form the basis for a test had, in fact, been rehearsed for many years prior to this in the literature on performance assessment (closely related to the communicative tradition – see, for example, the needs analysis approach to communicative syllabus design of the 1970s [e.g., Munby 1978], EAP testing [Weir, 1983], Carroll’s [1981] paper, discussed at the original LTF, and Brunfaut’s article on the 2010 Language for Specific Purposes symposium in this volume). McNamara (1996) distinguishes two traditions of performance assessment: a work-sample tradition (which was atheoretical, but focused on careful specification of authentic contexts in task design), and a cognitive/psycholinguistic tradition which emphasises the measurement of underlying ability and knowledge with less emphasis on task. There are clear parallels here in Bachman’s characterisation of the RL approach (atheoretical but authentic), and the IA approach (focusing on defining and measuring underlying traits of communicative ability). But even if such ideas had their basis in earlier discussions, Bachman (1990) usefully articulates what can be understood as two separate forms of CLT which functioned under the same banner.

Both approaches still form the basis for test design, although Fulcher (2010) maintains that “performance models [that is, real-life, task-driven approaches] have flourished” (p.113). Fulcher is particularly critical of the CEFR in these terms, arguing that the scales are not rooted in a theory of performance or second language acquisition (2010; 2004), and that the approach of the framework as a whole is “firmly behavioural” (p.114). Yet while performance models may have proliferated, it is not clear that “atheoretical” approaches to CLT have necessarily held sway to the extent that Fulcher suggests. It seems more likely that the current “mainstream” communicative approach is a synthesis of the two traditions. A clear example of this type of synthesis can be seen in Buck’s (2001) book *Assessing Listening* where he discusses how one might go about defining a listening construct. He begins by discussing the development of a competence-based listening construct (theoretical), but notes the limitation that there is no way of knowing which competences are of interest without recourse to TLU tasks. He then discusses a task-based listening construct (real-life, task-driven), but notes the impossibility of replicating TLU tasks with true verisimilitude in a testing situation. Buck therefore recommends a construct definition which is based on the interaction between test task and hypothesised competences, suggesting that we see (listening) performance “in terms of underlying language competence, but this is influenced by contextual factors” (Buck, 2001, p.109).

⁵ These ideas were developed further in Bachman and Palmer (1996) where a distinction was made between “authenticity” (a measure of how well the characteristics of a test task match those of tasks in the target-language use domain) and interactiveness, which was described in a similar way to Interactional Authenticity.

These contextual factors are determined by the characteristics of the task, which is based on real-life language tasks that occur in the TLU domain.

A consequence of advances in theories of communicative language ability, and the shift towards a synthesis approach, is that the “communicative” in CLT is no longer the contested site of meaning it once was in the 1980s and 1990s. Much mainstream language testing is now “communicative” in the sense that it draws on existing theories of communicative language ability and utilises “real-life” tasks, paying heed to authenticity and often including interactive performance elements. As Morrow points out, these aspects of test design are no longer burning issues. We might, however, characterise the communicative in language testing as lying dormant rather than having “faded away”. As some questions raised during the discussion at LTF 2010 CLT symposium demonstrate, there still exist important challenges for CLT, and in meeting these challenges a reinvigorated type of CLT might begin to take shape. These challenges are outlined below.

3. Challenges for modern communicative testing

3.1 Utilising and operationalising models in test design

There are on-going challenges for designers of tests with a communicative orientation both in terms of mapping-out the precise nature of the constructs of communicative competence which underlie test performance, and in generating validity evidence that these constructs are being operationalised through a particular set of test tasks. In the LTF discussion, this issue was raised with respect to Weir’s socio-cognitive model (see Weir, 2005; Weir & O’Sullivan, 2011). This model had been discussed at some length in the forum because Weir had presented his model and elaborated on its elements. In light of this, a general question was asked of how, given the comprehensiveness and depth of various models of communicative language ability, exam boards and test developers more generally could make practical use of such scholarship; whether the sophisticated constructs that had been developed could feed back into testing “products”. In response to this, Lynda Taylor spoke about the early years of IELTS (International English Language Testing System) where she recalled trying to apply the Bachman and Palmer 1990 model, attempting to turn the model into practical checklists that would enable more standardised IELTS tests in an operational sense. However, she reported that she had found it extremely difficult to use the CLA model in a practical context, noting that Cambridge ESOL’s current use of Weir’s socio-cognitive framework had been much more manageable. McNamara (2003) has made a similar observation in his review of Bachman (1990) and Bachman and Palmer (1996), noting that the test method facets approach advocated by Bachman had been little used in test development projects by that point (although aspects of the CLA model itself have been validated - see Phakiti, 2008).

We can see here a problem: increasingly complex frameworks of communicative language ability may hinder as much as enable good test design. Although this comment can be applied to the complexity involved in operationalising almost any framework of language ability, it is particularly problematic in terms of a theory of communicative language ability, because this is a theory that is so closely tied with what teachers do in the classroom. Intuitively, it should be more straightforward to draw on a communicative model than, say, a purely psycholinguistic theory of language ability, but this is not necessarily the case. Test developers – particularly in smaller scale situations, such as a person in charge of test development at a higher education institution in an EFL context – tend to be reliant on frameworks which have been designed to “unpack” existing models of communicative language ability. The Common European Framework of Reference (CEFR) is currently playing this role across many contexts as an accessible de facto theory of communicative language ability (and see Harsch, this issue, for further information on the CEFR). Other resources may also be drawn upon, such as the freely available research reports made available through large-scale testing programs. Educational Testing Service (ETS), for example, has published a number of valuable resources in the form of working papers detailing the construct definitions which underlie reading, writing, listening and speaking in the TOEFL iBT (see Bejar et al., 2000; Butler et al., 2000; Cumming et al., 2000; Enright, et al., 2000). A very useful published resource has been Chapelle, Grabe and Berns’ (1997) report on the development of the definition of communicative language proficiency at the beginning of the TOEFL revision process; the first appendix of which – “Chronological Development of the TOEFL 2000 Model at COE Meetings” (pp.48-50) – provides a detailed insight into how theoretical notions of CLA were transformed into a workable blueprint for test design which may be useful for practitioners in other contexts.

3.2 The richness of the construct

While there may exist difficulties in applying existing theories of communicative competence such as Canale and Swain (1980), Canale (1983), Bachman (1990) and Bachman and Palmer (1996) in test development projects because of their complexity, these models have also been criticised on the grounds that they are not complex enough; or at least that they are too narrow, and not rich enough, to provide adequate models of what it means to communicate competently in an additional language. McNamara has made this critique pointedly in his (1996) book *Measuring second language performance*. His main critique centred on the absence of “ability for use” in existing models of CLA. Ability for use was a component of Hymes’s original model which took into account not just an individual’s underlying ability for communication, but also their *capacity* for using this

across various contexts. Attempting to model this, however, presents enormous challenges, as McNamara writes:

Ability for use ... is more difficult to grasp, because we need to consider here a range of underlying language-relevant but not language-exclusive cognitive and affective factors (including general reasoning powers, emotional states and personality factors) which are involved in performance or communicative tasks (1996, p.59).

This question of ability for use was at the heart of the question raised during the LTF discussion where a question was asked about whether aspects of a test-taker's charm or their self-assurance were playing into the perception of their language performance on the Trinity GESE. The fact that the question was asked, and that it was responded to by a Trinity representative with the answer that "language requirements" are the focus of assessment, suggests that ability for use is still on the outside of conceptualisations of what is deemed important in the assessment of communicative language performance.

In communicative English language testing, one way in which ability for use might potentially be operationalised is through interaction with test-takers from other L1 backgrounds in paired- or group-speaking assessments (Taylor and Wigglesworth, 2009).⁶ For example, competencies such as "dealing with affective response" may be activated if a candidate holds negative attitudes towards his/her interlocutor's variety of spoken English (cf. Lindemann, 2002): the candidate must overcome this affective response in order to maintain successful communication. This potential for modifying existing notions of communication competence through the introduction of different varieties of English in tests is extended in Canagarajah's (2006) paper where he calls for a new paradigm which emphasises language awareness, sociolinguistic skills and negotiation skills in order to assess a candidate's ability to "shuttle" between different linguistic communities with different norms of language use (cf. Byram, 1997; Leung, 2005). Ability for use might also be operationalized through innovative measures, such as written tests of sociopragmatic knowledge and awareness (see Roever, 2011; Timpe, 2012). The emphasis in these newer reconceptualisations of communicative competence is on adaptability – how a candidate copes in a novel or challenging language situation in real time – which is itself closely connected to the older conceptions of strategic and sociolinguistic competence, but may be as close as we can come to incorporating ability for use.

⁶ It should be noted that much English language testing in EFL contexts is performed with groups of students who share an L1. It could be difficult to operationalize ability for use in this way within homogeneous learner contexts.

From a different perspective, communicative constructs have also been criticised as being deficient in terms of a psycholinguistic theory of how language is processed (see Van Moere, 2012), and this relates more broadly to a need for greater engagement with newer theories of second language acquisition which are challenging long held beliefs about how language develops and what makes language learning difficult. As a case in point, in the LTF 2010 discussion the issue of unpredictability in communicative language tests was raised in a question to Dianne Wall and Cathy Taylor on the nature of the interaction between candidate and examiner in the Trinity GESE: “how can any test cope with unpredictability?” Communicative speaking tests will usually involve interaction of some kind, and it is the irregularity of interaction, and its potential conflict with the requirement for some degree of control in the testing situation, that is at the heart of this question. Wall responded to this query by explaining that unpredictable situations arise in the interactive exam when, for example, the candidate’s speech moves in an unanticipated direction, and it becomes the duty of the Trinity examiner to deal with the change of topic. The candidate is also expected to deal with that kind of unpredictability in the examiner’s speech using the communicative strategies at his/her disposal. The ability to deal with unpredictability as an unexpected “twist” in an interaction was one of the cornerstones of Morrow’s original (1979) paper, in which he stated: “the apparently trivial observation that the development of an interaction is unpredictable is in fact extremely significant for the language user. The processing of unpredictable data in real time is a vital aspect of using a language” (p.16). However, this emphasis on the unpredictability of communicative interactions needs to be balanced with an awareness of the prevalence of patterns in language use. A good deal of recent research has provided evidence that many elements of language use are highly predictable, from formulaic chunks to routines which span many turns over discourse (Sinclair, 2004; Hoey, 2005). This has led to a reconceptualisation in some areas of linguistics about the nature of language itself: that it is essentially an assemblage of patterns which hold currency within a given discourse community (Macqueen, 2012). It follows that it is this high level of predictability – known to native speakers, or to those highly socialised within a given discourse community – which makes learning a second language so difficult; as Ellis and Larson-Freeman (2006, p. 568) say, “the very things that make a known language easy make a new language hard”. If this is the case, it becomes equally important to ask “how can any communicative test adequately model predictability in language use?”

Finally, there is a great challenge ahead in revising theoretical constructs in light of changes to communicative practices which have been brought about by technology – particularly the rapid developments in online and mobile communication. Although this important point was not discussed at the 2010 LTF, it is being grappled with in other areas of language testing research such

as PISA (Program for International Student Assessment) reading assessment. The 2009 PISA Reading Framework stated that, “the new demands on reading proficiency created by the digital world have led to the ... inclusion of electronic reading, an inclusion that has in turn resulted in some redefinition both of texts and of the mental processes that readers use to approach texts” (OECD, 2009, p.20). This redefinition of text meant the inclusion of both print reading texts – “fixed text with defined boundaries” (OECD, 2009, p.28) – and electronic reading texts, characterised by multiple texts, blurred boundaries between texts, and tools/features for navigating texts (e.g., hypertext links). The electronic texts lead to mental processes that are quite distinct from traditional print reading. For example, the need for the reader to orient him/herself within a complex electronic reading task, and to navigate pages in order to locate information requires knowledge of the “access structures and typical site structures” of webpages. This also means, however, that the test-taker has more control over how they read compared with a more traditional print reading assessment task where a single text is organised for the reader.

The approach taken in PISA is an important step forward, particularly in an era where print reading (or single text on-screen reading) still dominates many communicative-oriented tests, and where writing tasks might still include formulating a letter of complaint rather than posting a critical message on a company’s social networking page (where, in real-life, the perlocutionary effect would be much more immediate). PISA shows, as well, that it is not just the texts and associated tasks of non-electronic tests which need updating, but the underlying constructs as well. Yet there is still the sense that these developments are only scratching the surface of what communicative language ability in a digital environment really entails. The digital revolution extends far beyond navigating websites and writing emails, and spans communication in personal as much as public domains across new technologies such as blogs, microblogs (e.g., Twitter), wikis and social networking sites. Importantly, digital communication is in a process of constant change, as forms and functions are determined by technological innovations and the social practices of the users of technology (Coiro, Knobel, Lankshear & Leu, 2008). While some pedagogically-focused research has touched on the challenges of digital communication for assessment (see Godwin-Jones, 2008; Kol & Scholnik, 2008) in approaching these profound issues, and considering the implications for language tests, it would be useful for there to be greater engagement with the fields of “New Literacies” or “Digital Literacies” (see, for example, Coiro, et al., 2008; Gillen & Barton, 2010), where research is being conducted which overlaps substantially with the concerns of testers for enriching theoretical constructs. Chapelle and Douglas (2006) note that the approaches of the New Literacies movement – with its focus on contextually-bound social practices rather than the development of skills/abilities – may be “incommensurable” with the concerns of language testers (p.109; see also Teasdale and

Leung [2000] for a similar example of “paradigm crossing”). It is certainly the case that assessment is a looming issue for the New/Digital Literacies movement (see Gillen & Barton, 2010), however this only serves to emphasise the need for engagement between these two fields.

4. A research agenda

Departing from the broad issues raised during the LTF 2010 Communicative Language Testing symposium, this paper has argued that the current state of communicative language testing might best be described as dormant. CLT still remains the theoretical construct underlying many tests, and informs a focus on authentic language tasks which involve performance, but it is now accepted as a conventional approach across many testing situations. However there remain key challenges to CLT in utilising and operationalising models of communicative ability, and articulating a construct which is rich enough to reflect current communicative needs. Addressing these challenges would represent a reinvigoration of CLT, embodying a re-evaluation of the primary constructs that have supported the approach for more than 30 years. It would mean turning to the field of sociolinguistics, to new and emerging fields of communication such as digital literacies, and to psycholinguistics in order to enhance our understanding of what is involved in modern communication, to compare this with the constructs of communicative competence that are commonly operationalized, and to revise constructs in light of these findings. At the same time, a specific research program with a practitioner focus needs to be instituted in order to create opportunities for knowledge sharing between researchers and practitioners, and to ensure that the advances in research are of benefit to those who actually create tests across various contexts, particularly those practitioners who do not have the resources to fund their own research programs.

In terms of specific research plans, based on the challenges to CLT outlined above, a guiding principle of new research on a communicative competence construct must be a focus on “adaptability”. The notion of adaptability is the common denominator in a test-taker’s need to deal with different varieties of English, to use and understand appropriate pragmatics, to cope with the fluid communication practices of digital environments, and to notice and adapt to the formulaic linguistic patterns associated with different domains of language use (and the need to move between these domains with ease). With adaptability at the forefront, the following research areas might be addressed:

1. The development and validation of language tests which specifically assess a test-taker’s ability to deal with diverse, and potentially unfamiliar, varieties of English. These tests would use as their basis a different range of skills and abilities including:

- Ability to tolerate different varieties of English
- Ability to rapidly attune to new patterns of phonology, syntax, discourse style, etc.
- Ability to negotiate meaning
- Ability to accommodate
- Ability to use phonological features crucial for intelligibility with a given interlocutor
- Ability to ascertain and deploy appropriate pragmatics
- Ability to notice and repair breakdowns in communication

It has been noted above that this sort of construct may be operationalized – if inadvertently – by paired- and group-speaking tasks where test-takers interact with others from different L1 or cultural backgrounds. Discourse data yielded from tasks of this kind (complemented by stimulated recall performed by test-takers) could be analysed with a view to locating points at which these abilities are tapped in these interactions.

2. Some of the skills and abilities listed above which are not necessarily confined to intercultural communicative settings (e.g., ability to negotiate meaning, ability to accommodate, deployment of appropriate pragmatics) might also be assessed through carefully designed role-play tasks where an interlocutor is deliberately vague or inconsistent; or through integrated reading-to-write tasks where the initial input is deliberately challenging (e.g., responding to a letter of complaint where the complaint is unreasonable). Research into the types of strategic and pragmatic knowledge at play in tasks of this nature would be valuable.
3. The further use of corpus approaches could be used to establish the degree to which predictability in certain domains of language use can be modelled in test tasks (see also Flowerdew, 2012).
4. Piloting of new and innovative test tasks which utilise social networks, wikis, blogging and other modes of digital written communication. These tasks would ideally be based on an elaborated theory of what it means to communicate in the 21st century, with a focus on (perhaps) collaborative writing, asynchronous discourse, multimodal comprehension, and electronic reading. Such a theory would ideally be based on a melding of the social practices view utilised by the New Literacies field and a cognitive/psycholinguistic view of the abilities which underlie language performance in these environments.

As with existing models of communicative language ability, newer conceptualisations are likely to be complex and unwieldy for straightforward operationalization in language test design. This will always be a problem: if a theoretical model of communicative competence is to be sufficient, it will necessarily be multifaceted and detailed. From a test developer's perspective, however, this presents an opportunity for engagement with these issues through practitioner-led research. In translating complex models into workable testing blueprints, smaller-scale testing development teams might benefit from sharing knowledge about how conceptual frameworks have been interpreted and operationalized in different contexts, and how successful prototype assessment tasks have been in trial. If teams of practitioners were able to disseminate their practice-related findings more widely the field would be greatly enriched. Other practitioners would benefit from seeing decisions made about test design across different settings, and academic researchers (who may not be directly involved in test development) would in turn gain insight into the constraints imposed by practicality in operationalizing models of communicative language ability. A dedicated, practice-centred journal – an assessment equivalent of *ELT journal* – would be a useful first step.

In conclusion, although the nature of discussion in the CLT symposium at the 2010 LTF might have suggested that the communicative approach has become conventional, the term “communicative” still has symbolic currency as a way of describing language use that is authentic and interactive. This paper has demonstrated that there are still significant challenges in exploring the potential of CLT, particularly the shift in focus to adaptability. Rather than allowing the communicative approach to “fade away gracefully”, it should instead function as a dynamic paradigm from which to re-evaluate the status quo and account for on-going developments in communication practices. As long as there is change in these communicative practices, there will be a need for a determinedly communicative approach to language testing.

References

- Alderson, J. C. (1981). Report on the discussion on Communicative Language Testing. In J. C. Alderson & A. Hughes (eds.), *Issues in Language Testing* (pp.55-65). London: The British Council.
- Alderson, J. C., Clapham, C. & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. and Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

- Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL® 2000 listening framework: A working paper* (TOEFL® Monograph No. MS-19). Princeton, NJ: ETS.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Butler, F. A., Eignor, D., Jones, S., McNamara, T., & Suomi, B. K. (2000). *TOEFL® 2000 speaking framework: A working paper* (TOEFL® Monograph No. MS-20). Princeton, NJ: ETS.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters Ltd.
- Canagarajah, A. S. (2006). Changing communicative needs, revised assessment objectives: Testing English as an International Language. *Language Assessment Quarterly*, 3(3), 229-242.
- Canale, M. (1983). On some dimensions of language proficiency. In J. W. Oller (ed.), *Issues in language testing research* (pp. 333-342). Rowley, MA: Newbury House.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Carroll, J. B. (1981). Specifications for an English Language Testing Service. In J. C. Alderson and A. Hughes (eds.), *Issues in Language Testing* (pp. 66-110). London: The British Council.
- Chapelle, C. & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge: Cambridge University Press.
- Chapelle, C., Grabe, W., & Berns, M. (1997). *Communicative language proficiency: Definition and implications for TOEFL® 2000* (TOEFL® Monograph No. 10). Princeton, NJ: ETS.
- Coiro, J., Knobel, M., Lankshear, C., & Leu, D. (2008). Central issues in new literacies and new literacies research. In J. Coiro, M. Knobel, C. Lankshear, & D. Leu (eds.), *Handbook of research on new literacies* (pp. 1–21). New York: Lawrence Erlbaum Associates.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cumming, A., Kantor, R., Powers, D. E., Santos, T., & Taylor, C. (2000). *TOEFL® 2000 writing framework: A working paper* (TOEFL® Monograph No. MS-18). Princeton, NJ: ETS.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., McNamara, T. (1999). *Dictionary of language testing*. Cambridge: Cambridge University Press.
- Ellis and Larson-Freeman (2006). Language emergence: Implications for Applied Linguistics – Introduction to the special issue. *Applied Linguistics*, 27(4), 558-589.
- Enright, M. K., Grabe, W., Koda, K., Mosenthal, P., Mulcahy-Ernt, P., & Schedl, M. (2000). *TOEFL® 2000 reading framework: A working paper* (TOEFL® Monograph No. MS-17). Princeton, NJ: ETS.
- Fulcher, G. (2000). The “communicative” legacy in language testing. *System*, 28(4), 483-497.

- Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 20(4), 384-408.
- Fulcher, G. (2010). *Practical language testing*. London: Hodder Education.
- Fulcher, G. & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. London and New York: Routledge.
- Gillen, J. & Barton, D. (2010). *Digital literacies: A research briefing by the Technology Enhanced Learning phase of the Teaching and Learning Research Programme*. London: London Knowledge Lab, University of London.
- Godwin-Jones, R. (2008). Emerging technologies web-writing 2.0: Enabling, documenting, and assessing writing online. *Language Learning & Technology*, 12(2), 7-13.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Oxon: Routledge.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hymes, D. (1972). On communicative competence. In J. B. Pride and J. Holmes (eds.), *Sociolinguistics* (pp.53-73). Harmondsworth: Penguin Books.
- Leung, C. (2005). Convivial communication: recontextualizing communicative competence. *International Journal of Applied Linguistics*, 15(2), 119-144.
- Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society* 31(3), 419-441.
- McNamara, T. (1996). *Measuring second language performance*. London and New York: Longman.
- McNamara, T. F. (2003). Looking back, looking forward: Rethinking Bachman. *Language Testing*, 20(4), 466-473.
- Macqueen, S. (2012). *The emergence of patterns in second language writing*. Frankfurt am Main: Peter Lang.
- Morrow, K. (1979). Communicative language testing: Revolution or evolution? In C. J. Brumfit and K. Johnson (eds.), *The communicative approach to language teaching* (pp.143-159). Oxford: Oxford University Press.
- Munby, J. (1978). *Communicative syllabus design*. Cambridge: Cambridge University Press.
- Norris, J. M. (2009). Task-Based Teaching and Testing. In M. H. Long and C. J. Doughty (eds.), *The Handbook of Language Teaching* (pp. 578-594). Oxford: Wiley-Blackwell.
- Organisation for Economic Co-operation and Development (OECD) (2009). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. Paris: OECD.
- Phakiti, A. (2008). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language Testing*, 25(2), 237-272.

- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4), 463-481.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. New York: Routledge.
- Kol, S., & Schcolnik, M. (2008). Asynchronous forums in EAP: Assessment issues. *Language Learning & Technology*, 12(2), 49-70.
- Taylor, L. & Wigglesworth, G. (2009). Are two heads better than one? Pair work in L2 assessment contexts. *Language Testing*, 26(3), 325-339.
- Teasdale, A. & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*, 17(2), 163-184.
- Timpe, V. (2012). Strategic decoding of sociopragmatic assessment tasks – an exploratory think-aloud validation study. *Second Language Studies*, 30(2), 109-246.
- Van Moere, A. (2012). A psycholinguistic approach to oral language assessment. *Language Testing*, 29(3), 325-344.
- Weir, C. (1983). *Identifying the language problems of overseas students in tertiary education in the UK*. Unpublished PhD thesis: University of London.
- Weir, C. (2005) *Language testing and validation: An evidence-based approach*. Palgrave MacMillan, Basingstoke.
- Weir C. J. & O'Sullivan, B. (2011). Test development and validation. In B. O'Sullivan (ed.), *Language Testing: Theories and Practices* (pp.13-32). Basingstoke: Palgrave.