What does the Business Processing Outsourcing industry want from English language assessment?

JANE LOCKWOOD – Hong Kong Institute of Education

ABSTRACT

This article reports on a consultancy project carried out in a large third-party call centre in Manila in 2004. The project aimed to improve English language communications assessment and training for the call centre representatives on the telephones to the United States, the United Kingdom and Australian customers. This article looks particularly at the use of English language assessment tools and processes used at this call centre. First, there is a description of what the business requirements were for language assessment at recruitment, at training and at quality assurance. The article evaluates current assessment tools and practices used in this call centre against current language assessment research and best practice. Finally there is a discussion on how the Business Processing Outsourcing industry and the language assessment fraternity, including the large-scale testing companies, may provide improved solutions for this industry.

Introduction

In many countries, most notably India and the Philippines, the Business Processing Outsourcing (BPO) industry is rapidly expanding. In these cheaper destinations, back offices, call centres, shared services centres and professional services (such as legal, financial, medical and human resource (development HRD functions) are rapidly being outsourced and operated offshore from the United States, the United Kingdom, Europe, Canada, Australia and New Zealand.

It is estimated that by 2015, 15 million jobs will be generated offshore in this industry, and this represents only 10% of possible jobs (NASSCOM-McKinsey Report 2005; Tuchman 2004). While the savings are significant (Friginal 2004; Magellan Alliance 2005), these cheaper destinations employ staff who speak English as a second language, not as a first language. The consequences of employing non-native speaker (NNS) staff as customer service representatives (CSRs) are significant in terms of language and acculturisation. There is currently great concern regarding the quality of the communication with the customers. Poor recruitment rates ranging from 1–5% are reported by HRD departments in India and the Philippines. Another major problem is that the customer satisfaction (CSAT) scores, which are carried out by independent parties overseas and which form the basis of the service level agreements between the United States companies and the local call centres, are often negative about the quality of communication.

A range of communication problems with customers has already been identified in the research (Forey and Lockwood 2007; Forey and Hood forthcoming), ranging from an inability to explain products and services clearly, lack of lexicogrammatical choices in the soft skills and a lack of intercultural understanding of what the customer really means. This all puts pressure on the HR development, training and quality assurance (QA) processes in the local call centres to improve services. While these QA problems may be attributed to weak English language communication skills of the CSRs on the telephones, I argue that the low language proficiency abilities and the lack of TESOL assessment skills of the HR department, the training department and QA personnel themselves exacerbate the problem when recruiting, diagnosing, training and coaching for language.

Thus, the English language communication and intercultural problems in the new offshore and outsourced centres may not have been fully appreciated by multinational companies domiciled in the English-speaking West when selecting their call centre sites in NNS destinations. The issues related to English communications, intercultural awareness and English language assessment have been brought into sharp focus over the past ten years in which India and the Philippines have
seen huge growth in their call centre operations. Anecdotally, everyone has a story to tell about the difficulties of communicating with call centre staff in India and other parts of Asia. However, I only consider the particular problems call centres have with their English language assessment practices in their workplaces by describing and evaluating them.

**Overview of the study**

In this article I describe the need for English language assessments in the BPO industry and the way, particularly in call centres, that these assessments are carried out for a variety of business purposes. This study is based on a consulting project carried out in a third-party call centre (i.e. a call centre that provides outsourced support services to a number of clients) in Manila in 2004. This call centre, which will henceforth be referred to as Call Centre X, was one of the first to be established in the Philippines and houses about 35 client accounts and approximately 3000 call centre seats across three sites in metropolitan Manila.

The business requirements for English language assessment in call centres are complex. Based on observations of the business processes, on analyses of developed language assessment and curriculum documents, on observations of language-training sessions, and on focus group and individual interviews with a range of stakeholder groups within the call centres, it became clear that the key stakeholders in the call centres need language assessments that are responsive to their different business requirements, such as recruitment and promotion/migration, training, QA and coaching, and assessments that can be embedded into their processes. This is because the purposes for the language assessments are different; for example, recruitment needs to provide a yes/no answer to gaining employment, training needs a diagnosis for effective instruction and a score for readiness at the end of training, and QA needs evidence of customer satisfaction and further diagnosis for coaching. As well, each of these stakeholders needs to operate and own these language assessment processes in an ongoing way themselves. They therefore require training in order to develop and use language assessments on a routine basis. Unfortunately, many different kinds of language assessments have already been developed and installed by non-specialists into the call centres and, not surprisingly, they are failing to do a valid and reliable job. This is exacerbating the recruitment team, where they complain of low recruitment rates; training, where few are successfully endorsed into client accounts as a result of training; and QA, where there is often a discrepancy between the CSAT scores done overseas and the client account score-carding done locally.

Call Centre X had reported a lack of success with using commercially available business English tests for these different purposes within the industry. Most commercially available business English tests provide a score for proficiency and do not break down the result in a diagnostic profile. Furthermore, they are also costly and the long turnaround time does not fit in with urgent business requirements. However, most importantly, Call Centre X reported that such tests lacked credibility in terms of what they were testing for. For example, many speaking tests available commercially do not test for interactive capability. This is critical for call centres. They also wanted an assessment process that could be administered by their own employees. However, Call Centre X, at the time of the consultancy, was very concerned about the veracity of its existing language assessment processes because they did not appear to be working. This was adding to the problem of the dislocation between the CSAT scores done overseas and QA scores done locally. Another worrying factor was the widespread evidence of poor levels of English proficiency of those administering communications assessments. Despite the fact that Cambridge ESOL has recently abolished the requirement for NNSs to have evidence to IELTS (International English Language Testing System) 9, there is an assumption that language assessors need very strong language skills. Furthermore, they require language assessment training and calibration if they are to be effective.

First, I describe the call centre end-to-end business requirements for English language assessment and evaluate their current practices. I propose that the BPO industry context requires a shift in thinking about language assessment beyond the use of large-scale generic business English tests and conventional approaches to language assessment. I suggest that the industry will not find complete answers to their problems by using
commercial business English tests and that the call centre companies need to be more analytical about their own business requirements, more judicious when choosing commercial business English tests and more skilled when developing their own assessments. Furthermore, the call centres, themselves, need to think carefully about who owns the internal assessment processes when they are transferred into the company itself via recruiters, trainers and QA personnel; they also need to think about the mechanisms they will use to ensure sustainable validity and reliability in their processes.

Implications of this study provide a challenge to the business English test providers and the language-testing fraternity in general. Providers of large-scale commercial tests are currently seeking to promote their highly centralised and scalable testing products as easy and lucrative answers to the highly complex business requirements in the call centres. Unfortunately there is great disappointment in the BPO industry with the existing large-scale business tests because language testers are not aware of the BPO stakeholders’ needs. As one Communications Manager in a large, well-known Asian regional bank, who was seeking a comprehensive internal language assessment solution regionally, said recently:

_The only time language testers come and talk to us is when they want to sell their business English tests off the shelf. They are not interested in finding out what we do and whilst we know what we want in terms of the business requirement, we do not know what we need nor do we know how to articulate all this. You are the experts. We’ve wasted a lot of money on business tests that promise the earth and deliver nothing … we need good assessment consultants helping us develop our own_ (regional South-East Asian bank Communications Manager, conference call, June 2008).

**What does the literature say?**

There is much written in the applied linguistic literature on the processes of needs analysis (Brindley 1984; Willing 1988; Nunan 1988; Brown 1995; Dudley Evans and St John 1998) for English language curriculum design, although interestingly these writers are mostly concerned with learner backgrounds, learner styles and strategies and program task needs. Prince (1992) and Joyce (1992) take a situational analysis and they look at the network and processes of language needs in Australian workplace settings. Reeves and Wright (1996) take a broad-angled ethnographic view of needs analysis in the workplace and encourage language providers to _tailor_ courses specifically to the needs of the workplace through what they term _a language audit_. Language auditing (Reeves and Wright 1996) assists the organisation to understand its own language needs across different business requirements, accounts and post-holders. Once these language needs and thresholds are understood, planning for language improvement and measurement can take place.

_A language audit therefore should strive to help a company’s management make the right strategic decision in recruitment, in modifying the organisation and the behaviour of some departments, as well as allocating resources for training and quality assurance_ (Reeves and Wright 1996: 2).

In the language assessment arena, Bachman and Palmer (1996) provide a framework for language assessment based on current research findings for the design and implementation of language tests across different teaching environments. Brindley (1994) also has brought together a summary of assessment and testing initiatives and documented the issues and problems that have arisen, particularly in his work on language competency. The language-testing literature is rich in discussions about the relative merits of using different applied linguistic frameworks to achieve construct validity and reliability in the design and use of any test. Much of the discussion has revolved around pushing criteria for assessment beyond the traditional preoccupations with pronunciation and grammatical accuracy to a consideration of broader communicative domains such as discourse and interactive capabilities (Canale and Swain 1980; Bachman and Palmer 1982; Bachman and Savignon 1986; Davies 1988; Hughes 1989; Bachman 1990; Weir 1993). Assessment practices that reflect communicative approaches to language training in Language for Specific Purposes (LSP) contexts are also well covered (Lumley and Brown 1996; McNamara 1997; Douglas 2000; Elder 2001), and well-documented accounts of test development for different groups of occupations and professionals abound. McNamara (1990, 1996) and McDowell
What does the Business Processing Outsourcing industry want from English language assessment?

(1995) have both developed standardised and competency-based tests for a range of teachers and health professionals in Australia, and Douglas (2000) looks at specific language use situations to derive test content and test methods for highly specific LSP, such as English for air-traffic control. Such frameworks, however, have yet to be fully understood in an industry context and incorporated into business language assessment practices onsite.

Biggs (1991) provides an interesting description of the paradigmatic shift in the area of language assessment and considers the links and effects of different modes of assessment and different contexts for assessment incorporating workplace stakeholders into the process. Furthermore, Revell (1994) talks specifically about language testing in the Hong Kong corporate context and makes an interesting and relevant link between language assessment and quality management practices in Hong Kong companies:

*Internal and external customer requirements are first priority, but other sources of information may be needed and should be found to make sure training and assessment are aligned with the company’s quality requirements. These may be found in training criteria, in job descriptions, or in external quality requirements …* (Revell 1994: 333).

While Revell has articulated the problem, he offers no solutions in the form of practical guidance or prevailing theoretical frameworks that may prove useful to the English language workplace assessment practitioner.

The current testing and assessment literature effectively makes the case for context and purpose sensitivity, but is silent on how this might be achieved in such a dynamic workplace as call centres. What is required, therefore, is a multidimensional model for language assessment incorporating different stakeholder views and requirements for English language assessment gatekeeping and reporting. In 2008 little still exists in the international literature about the systematic inclusion of various workplace stakeholders into the curriculum and assessment planning process for workplace English language training and assessment reporting. At the 30th Annual Language Testing and Research Conference (LTRC), the keynote address, entitled *Why don’t the stakeholders in language assessment just cooperate?*, was delivered by James Dean Brown. He opened his keynote address by saying:

*The LTRC theme this year is Focusing on the core: justifying the use of language assessments to stakeholders. That theme leads me to my title question – one answer to which is that stakeholders don’t cooperate because, arrogantly or inadvertently, we fail to consider them, or properly inform them* (Brown 2008).

Lockwood (2002), in an analysis of workplace training and evaluation processes in Hong Kong workplaces, has argued for an interdisciplinary approach to curriculum and program evaluation incorporating applied linguistic and business management models (Goldstein 1993; Kirkpatrick 1994; Bramley and Pahl 1996) to ensure business stakeholder involvement in these educational processes.

Because commercial business English language tests need to have market scalability to be profitable, complex stakeholder requirements would add costs and complexity. Therefore, large-scale testing companies have not addressed this issue. However, aside from commercial business test development, few language testing academics and practitioners have immersed themselves in business requirements to establish the precise needs of LSP testing, and this remains a very under-researched area. These conspicuous absences of research-based solutions have resulted in companies developing their own assessment solutions, with little or no knowledge of language assessment theory and practice.

Currently the BPO industry is being let down and limited by traditional testing paradigms of test development and practice. This general malaise has been articulated by McNamara and Roever (2006: 247), who describe this as a ‘difficult juncture’ in language testing:

*A discipline whose intellectual sources are fundamentally asocial, in measurement, psychology and a linguistics that focused on individual cognition, has found itself lacking effective tools to conceptualise, acknowledge, and confront the social dimension of assessment either at the micro level (of face-to-face interaction) or at the macro level, the social context in which test use occurs.*
This study considers the business requirements of the different stakeholders in the call centres. It describes and evaluates current practices in language assessment and provides a discussion on possible ways forward for research and better-informed practice.

The study and approach

This study was carried out during a six-week consultancy within one of the largest third-party call centres in the Philippines. There were approximately 35 different industry client accounts in this call centre across five verticals (industry types), including finance, insurance, retail, information technology, and travel and tourism. Ninety per cent of the calls were inbound (i.e., customers call the call centre with specific problems and requests), as opposed to outbound calls that upsell and market new products to customers. Ninety per cent of the accounts at the time of the consultancy were American.

Call Centre X employed a permanent team of 12 language trainers, two QA communications assessors (who did post-course language training assessment) and at least one QA specialist per account (who did regular QA checks and coaching on site). They were all overseen by a senior communications manager. The investment in communications development in Call Centre X was therefore substantial. The aims of the consultancy were to investigate the sources of communication breakdown within a sample of United States client accounts, evaluate training and assessment processes, and make recommendations for improvement. The following methodology, to analyse the problems and to establish the training and assessment requirements of the call centre, was agreed upon.

It was agreed that the consultant would spend time researching ten accounts – two per vertical, with one that was supposedly an easy account and one that was supposedly difficult. Specifically, the following methodologies were agreed.

1. Collection of documents relevant to the study (e.g., CSAT data, scorecards, training documents and assessment tools).
2. Observation on the floor in each account and within HR development and training (e.g., barging into calls, watching screening interviews, and watching training, mentoring and QA processes in action).
3. Focus group interviews with the key stakeholder groups (e.g., account managers, HR personnel, trainers, QA monitors and specialists).
4. Individual interviews with a sample of CSRs.
5. Survey of client account managers regarding the English language needs within their call centre accounts.

A final report was written after the six weeks and recommendations for future development were made. This article only discusses those aspects of the study that relate to the needs and processes of English language assessment.

A needs analysis for call centre language assessment

One of the biggest problems in call centre English language assessment relates to the complex and multipurpose nature of the language assessments required by the business. This complexity was established by doing a needs analysis that involved spending time watching what went on in the HR development, training and QA departments, what went on within the client accounts themselves and by talking extensively to all stakeholders. Prior to the consultancy, the communications manager of this call centre was insistent that the call centre required a systematic, valid and reliable end-to-end solution to the English language training and assessment practices. Furthermore, Call Centre X required an English language assessment system that could be owned and controlled by the different stakeholders, and could be embedded into each of the business processes as captured in Figure 1.

Who are the stakeholders?

At Call Centre X each point of the business requirement had developed highly idiosyncratic and fragmented processes to scoring the CSRs. This multiplication of language-scoring tools and processes based only on layman knowledge was wasteful, confusing, invalid and unreliable. Recommendation 7.1 in the subsequent report (Lockwood 2004: 12) proposed that:
Given the volume of applicants to Call Centre X and the need for valid and reliable tools and sets of processes to get the right recruits into the right accounts, there is an immediate need to revamp the tool and streamline the process. This will also ensure Call Centre X starts using a common language to talk about language level and gain.

Each of the business requirement points for language assessment shown in Figure 1 will be considered individually below.

**Did Call Centre X carry out a language audit?**

In a study of Hong Kong workplace language training design and evaluation processes (Lockwood 2002), it was noted that few workplaces were aware of the language requirements of their organisations. The idea of a language audit and the value it could bring to streamlining an integrated approach to communications training, recruitment and QA was unknown. The Manila call centres, at the time of the consultancy, were no exception.

Carrying out a language audit requires language assessment experts to understand the company organisation and processes and to assess accounts and post-holders for language levels commensurate with the language complexity required to do the job. Typically, standardised scales and descriptors are used for these purposes and typically time is spent observing the post-holders at work in the various accounts. Sample numbers of post-holders undergo a detailed language assessment, and these processes and results enable the language assessment expert to benchmark or tag a post for language assessment. In the case of Call Centre X, sample post-holders and the client accounts needed to be benchmarked to demonstrate the value of this knowledge to the organisation.

This language auditing approach had not yet been considered by Call Centre X at the time of the consultancy, although plenty of anecdotal evidence existed about the relative difficulty or ease of the 35 separate client accounts and who were the good and poor CSRs in terms of language level. This lack of language auditing and client account benchmarking meant that there was no systematic way for placing, nor migrating, CSR recruits into appropriate client accounts commensurate to their levels of English. In order to carry out audits in the BPO in Manila, the author developed a set of speaking and writing scales and descriptors called the Business Processing Language Assessment Scales (BUPLAS1).

**What happens in the HR department at Call Centre X?**

A common problem currently facing HR departments in call centres is knowing how to assess, with accuracy and in a granulated way, language proficiency that best leads to employment and client account placement. Call centres typically screen hundreds, sometimes thousands, of applicants per month. They channel them into the new business client accounts when they are ramping up, and, as attrition runs very high, they are also constantly replacing CSRs who are leaving.
Call Centre X was screening over a thousand applicants a week at the time of the consultancy. Given that the recruitment success rate was 1–5% of applicants, depending on the account, the HR department was devoting many resources to simply screening CSR applicants. In this company more than three floors of one of the Manila business sites was devoted to the recruitment and screening process, and the HR department employed approximately 15–20 staff members for this purpose. Casual staff members were called in from time to time during very busy periods. Of the hundreds of applications that were received weekly in Call Centre X, many were weeded out before they presented for the interview and computerised language assessment. This initial weeding out occurred because the company simply did not have the resources to interview everyone.

This first process was carried out by casual staff employed to conduct telephone interviews of all prospective CSRs. The staff had no qualifications for what they were required to do and were not assessed for language proficiency. No training or rating information was given to these telephone screeners regarding criteria, processes and reporting for language assessment. In fact, in follow-up interviews with these staff, they were generally uncertain about the main purpose of the telephone interview. Some telephone interview staff members thought they were screening for aptitude, while others thought the purpose of the interview was to elicit and provide information about Call Centre X, as well as, or instead of, assessing language.

Typically these casual staff members tried to hit high targets for each day (eg telephone interview 30 applicants) and they generally worked fast and in isolation. There was little sharing of information and processes with co-workers and little guidance from the HR department.

Observation of the casual staff at work showed that they would typically first skim and scan the applicants’ résumés before telephoning them. Applicants were not contacted beforehand about the interview and were not told, when they answered the telephone, that the purpose of the interview was to assess language proficiency and competency. Consequently, the applicants were caught completely off guard; for example, some were in the middle of feeding children at meal times and some had just rushed out of the shower to answer the telephone.

The questions the telephone screeners used were set by the HR department and probed general employment suitability questions; for example, did the applicants have any problems with working through the night, how did they think the call centre work might impact on family life or how might they deal with an irate customer? Once the short telephone interview was over, the résumé was either endorsed and put in a pile for further consideration, or rejected. In discussions with the telephone screeners after the interviews, many said they had decided before they talked to the applicant whether or not they would endorse them on the basis of the secondary schools and tertiary institutions they had attended and the academic results they had attained. Others commented on whether they liked the content and manner of the responses. A few talked about applicants’ heavy first language accents and therefore deemed them unsuitable. None appeared to make any principled decisions based on accepted criteria for making language assessments. In the view of the consultant, this process was very wasteful and ironically served, in many cases, to advance weaker English language candidates to the second stage of the interview process and to screen out those with good language levels or real potential. About 50% of all applicants were screened out at this stage.

Once applicants had been endorsed, they were invited to come into the company for a computerised test of grammar, reading and writing. This second language assessment stage needed to be passed before the applicant finally got through for a speaking assessment. The target set by the HR department was to screen out a further 50% of applicants at this stage. The computerised reading test was essentially a collection of newspaper articles and literary passages (including poems), with comprehension questions constructed internally by recruiters in the HR department. The items probed mostly content and interpretations of poetry and literary passages. The grammar segment comprised a large number of discrete items based on common first language mistakes made by Tagalog speakers; for example, subject–verb agreement, missing articles and prepositions. Out of 50 items analysed, 30 subject–verb agreement sentences were contained in the grammar test. Interestingly, listening
comprehension was not measured at this stage of the interview. Good listening skills are key to good CSR communication, and this is an area where CSRs have difficulty (Forey and Lockwood 2007; Forey and Hood forthcoming; Lockwood, Forey and Price 2008). When asked about the absence of a listening test, HR personnel said this was covered in the final stage of the oral speaking test. Although a writing test was administered at this point, the author did not observe anybody in the HR department mark the test. Those who were successful at this stage went through to final interviews for speaking assessment. Those conducting the final speaking tests were client account managers and QA specialists who were now selecting potential agents for their accounts.

The client account managers and QA specialists typically conducted their 15-minute interviews with criteria that were developed in-house by the communications trainers and called the American English Enhancement (AEE) Test. When talking to interviewers at the end of each speaking assessment, most cited accent and pronunciation problems as reasons for accepting or rejecting applicants. Some interviewers simply counted the mistakes made in grammar and pronunciation, and one reported that if the applicant made 36 mistakes in the 15-minute interview she would not accept that person into the account she was interviewing for, but might endorse the applicant to an easier account. The approach of looking at mistakes and counting them was a typical one for the lay language assessors at Call Centre X. Knowledge of interlanguage mistakes and looking for lexicogrammatical range (even if there were mistakes) as evidence of level of proficiency was not understood by this group.

Accuracy trumped range when the interviewers were making language assessments. The interviewers reported that they were constantly responding to the client account pressure to have virtual native speakers sounding like Americans for the accounts. During the time of the consultancy, one very large United States insurance account was experiencing very poor CSAT scores. The insurance company responded to this quality problem by sending out its own HR native-speaker team to Manila to control the final account endorsement speaking tests at recruitment. None of these United States native-speaking recruiters had TESOL backgrounds, nor tools or processes to carry out speaking assessments. In discussions with these United States recruiters when they were in Manila, they were unanimous in their request for recruitees with good American accents and evidence of good American idiom use. Needless to say, they went home mostly disappointed and unable to assist in this language assessment task at recruitment. The need for good language assessors at this recruitment stage is critical, as the HR department is now looking for near-hire CSRs (ie those, who with four to six weeks of intensive language training, can become CSRs) and for CSRs ready to go on to the telephones with minimal language training.

A concluding remark about the recruitment assessment process in the consultancy report (Lockwood 2004: 7) stated:

The language screening process at recruitment is flawed because of invalid, incomplete and unreliable tools and processes. There is no principled differentiation being made regarding the language proficiency levels of the recruitees. It appears that not only are good people being rejected at the early stages of this process, but those who are getting through are not being successfully sorted into the right accounts at the recruitment stage. Valid and reliable assessment tools and processes are required.

What happens in training support at Call Centre X?

Once CSRs have been endorsed to accounts, they follow a general two-week communications and culture program where they are rated again, pre-course and post-course, to determine final suitability for the account. All scores are entered into a database and CSRs are deemed to be at one of three levels: Level 1 (63–79%), Level 2 (80–90%) or Level 3 (91–100%). All agents must reach Level 2 before they are allowed to start work on-site. The scores are arrived at by the final speaking assessment, which is recorded and then checked by two independent QA specialists. The final scores are then averaged and entered into the database. If the CSRs show no language proficiency gain, they are recycled through the two-week program until they are deemed to have improved enough to move on-site. Many are rejected, too, during this process.

The company-developed AEE Test measured four main domains:
What does the Business Processing Outsourcing industry want from English language assessment?

1. Pronunciation – articulation of segmentals, vowel and consonant sounds, final plosives (e.g., /f/ versus /p/ sound differentiation, syllable stress).

2. Application of speech techniques – intonation, pitch, volume, rates (i.e., suprasegmentals).

3. Vocabulary and grammar – verb–subject agreement, pronouns, prepositions etc.

4. Listening and comprehension (Call Centre X, AEE Parameters document).

Interestingly, there were no domains or parameters probing the interactive or discourse capability of the prospective agents. Recent studies (Forey and Lockwood 2007) show that these are the two domains where, tellingly, communication breaks down on the telephones. A communicative framework for language assessment had not been considered at Call Centre X because the communications team was not TESOL educated nor familiar with the basic principles of good language assessment practices. In most call centres in Manila, language is simply viewed as grammar and pronunciation, and any other problems fall into a generic soft skills/communications bundle, training for which uses United States soft skills materials designed for native speakers. These are all that are available and not surprisingly do not have an impact, as described by Friginal (2007) later in this article.

The process of administering the 15-minute AEE Test was observed. The test entailed:

1. a face-to-face interview with the trainer
2. an activity in giving directions
3. a narration based on a cartoon strip/storyboard.

Tasks 2 and 3 were taken from a well-known test downloaded from the Internet.

Candidates were extremely anxious during this post-course test because of the high-stakes decisions being taken. They would either successfully go on-site and start earning money, or they would be recycled back into training for at least another two weeks. If they did not show language improvement in two to four successive training programs, they were finally rejected. In interviews with the trainers and the QA personnel at this pre-course and post-course testing, they said they looked primarily for first language interference in pronunciation and grammar mistakes typical of Filipinos, such as subject–verb agreement, prepositions and lack of plural endings.

The AEE Test measures carried out by the trainers and the QA personnel were not tested for reliability, nor was the assessment checked for validity at the time of this consultancy.

The following two recommendations (Lockwood 2004: 13) were made for pre-hire training:

1. Do not use the pre-hire training program as a language proficiency booster program, as no proficiency gain is realistic in two weeks.
2. Do not do pre-hire and post-hire training proficiency assessments, as it will show no gain.

What goes on in quality assurance processes?

Perhaps one of the most critical functions for the call centre business is the demonstration that the CSRs are providing high-quality service. This will ensure good customer service feedback and high scorecard results. Service level agreements revolve around CSAT surveys that must meet the agreed benchmark levels to fulfil the terms of the contract. CSAT surveys are carried out by independent third parties in the United States and are based on a series of questions answered by customers. CSAT scores are generated regularly as a quality check.Locally, the call centre also has an elaborate QA mechanism, called scorecarding, carried out by QA specialists. One problem reported by call centres is the discrepancy between the CSAT scores and the locally assessed scorecard scores as a QA measure. Discrepancies between CSAT scores and local QA scores was a concern in some of the client accounts in Call Centre X at the time of the consultancy.

A team of QA monitors was attached to specific accounts in Call Centre X for the purposes of rating the CSR recorded calls against these client account scorecards. The scorecards, which are account specific, are mostly generated by clients in the United States or the United Kingdom, and are generally problematic, as they have been designed for native, rather than non-native, speakers of English and used in the United Kingdom and the United States as quality measures. Although the scorecards aim to measure...
communicative competence, they have not been informed (nor adapted for NNS) by existing frameworks in language testing and assessment. There are no language domains, no descriptors and no scales. They do not measure the domains of pronunciation, language accuracy and range, discourse and interactive competency with validity and reliability.

The QA monitor simply ticks yes, no or not applicable to a series of questions (35–40) about how effective the communication has been with the customer. The series of questions on the scorecard relate to product knowledge as well as communication skills, mostly soft skills. This practice is highly problematic in terms of fairly and accurately assessing language skills, as the questions have been designed for native-speaker CSRs. The QA monitors, themselves, in Call Centre X agreed that this process was highly subjective. (For a full discussion of the veracity of the scorecard used as a QA process in call centres, see Lockwood, Forey and Elias forthcoming.) Unfortunately, these scorecards remain a non-negotiable cornerstone of the QA process.

Apart from the validity issues related to the construct of the scorecard, when observing the quality monitoring process there were a number of problems that threatened the reliability of the scoring process itself. First, many of the QA monitors had very low levels of English. In many cases the QA monitor had a lower level of English than the CSR being assessed. This made any meaningful assessment of language and communication effectiveness highly problematic, as suggested earlier. There was also a second problem related to the lack of training and calibration given to the QA monitors to assess communication. Additionally, the scorecard process was very hard to carry out because of the format of the tool they are required to use. On the scorecard, the number of fragmented items that the QA specialist is supposed to score, listening only once to the call,

<table>
<thead>
<tr>
<th>#</th>
<th>Components</th>
<th>Criteria for Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Greeting – offer welcoming words</td>
<td>No ‘Hi’, ‘Hello’, ‘can’, ‘speaking’</td>
</tr>
<tr>
<td>2</td>
<td>Greeting – maintain upbeat tone</td>
<td>Does not drop in tone at end, ends in an upbeat</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(form of a question)</td>
</tr>
<tr>
<td>3</td>
<td>Greeting – use unhurried pace</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Really listen – don’t interrupt</td>
<td>Let caller vent</td>
</tr>
<tr>
<td>5</td>
<td>Express empathy through words</td>
<td>Can have without #6</td>
</tr>
<tr>
<td>6</td>
<td>Express empathy through tone</td>
<td>Can’t have without #5</td>
</tr>
<tr>
<td>7</td>
<td>Use caller’s name as soon as you hear it</td>
<td>First opportunity to use name – ask for it if not given</td>
</tr>
</tbody>
</table>

Table 1: The scorecard: Part 1 Greeting

<table>
<thead>
<tr>
<th>#</th>
<th>Components</th>
<th>Criteria for Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Tell caller you will help</td>
<td>Must have ‘I’, ‘you’ and a verb – must be clear statement</td>
</tr>
<tr>
<td>9</td>
<td>Ask permission to gain more information</td>
<td>Must be clear question (when 3 or more questions are to be asked)</td>
</tr>
<tr>
<td>10</td>
<td>Use ‘I’ not ‘we’ when appropriate</td>
<td>‘We’ used in reference to next steps, ownership</td>
</tr>
<tr>
<td>11</td>
<td>Be courteous, use ‘please’ and ‘thank you’</td>
<td>Be polite and friendly</td>
</tr>
<tr>
<td>12</td>
<td>Express sincere, helpful attitude with tone</td>
<td>86% – majority of the call</td>
</tr>
<tr>
<td>13</td>
<td>Remain calm</td>
<td>Not defensive in words or tone</td>
</tr>
</tbody>
</table>

Table 2: Gathering information (Act positively: tell them you will help)

(Tables 1 and 2 from Forey, Lockwood and Elias forthcoming)
is extremely difficult. Tables 1 and 2 show a small sample of 13 out of 35 items that require yes, no or not applicable responses from the QA monitors in real-time listening.

What was notable in this consultancy was the amount of hard work and resources that Call Centre X had put into the quality assessment processes. Call centres generally use one of three rating systems to monitor quality:

1. scores are based on a yes/no assessment
2. scores are based on weighted values for each item
3. scores are based on the proportion or number correct or incorrect.

Two call centre researchers (Cleveland and Mayben 2004: 211) recommend that:

There is room for opinion, and we’ve seen all three methods work well or poorly. But based on our findings, we generally recommend that score results be based on the proportion of defects. The pass/fail method doesn’t provide enough information about calls and trends to guide you in making improvements at the individual and process levels. And methods that assign a value to each time are all too often overly subjective.

This consultancy was one of the first commissioned to look at how applied linguistic and language assessment practices could improve, and indeed simplify, processes by bringing applied linguistic frameworks and processes to language assessment. Call centres are constantly seeking a competitive edge in their QA processes. When there is negative feedback on the communication skills of the CSRs, American native-speaker trainers and QA personnel arrive on the doorstep and appear to further complicate the processes by adding more questions to the scorecard, by taking control of the recruitment interviews, looking for the best accents and by putting pressure on the training departments to work miracles. None of this is helpful, nor is it informed by good language teaching and assessment practices.

The report on the QA processes (Lockwood 2004: 13) made a set of final recommendations to Call Centre X:

1. Ensure all QA personnel are at/or exceed the English language levels of the accounts that they are attached to.
2. Ensure that all QA personnel are trained in diagnostic English language training assessment and coaching skills.
3. Introduce one English language proficiency scorecard across all accounts for QA purposes that mirror the English communication domains and criteria used by the English language training specialists.
4. Ensure there is an interface between the English language specialist team and the QA staff.

Discussion
Clearly there are solutions that language assessment frameworks and practices can offer call centres in their endeavours to improve and systematise their processes and reporting for English language communication quality. However, this relies on applied linguistic practitioners and researchers spending time in the call centres to understand the business requirement needs. It also relies on the BPO industry recognising that there is something to learn from applied linguistics and language assessment best practice.

The absence of needs analysis and training materials
Until recently, and before the trend to outsource and offshore services, the call centre industry back home was staffed by native speakers of English. It is not surprising, therefore, that attention has not been paid, until recently, to communication problems that relate to language proficiency and intercultural awareness in the call centre industry. With the newly found BPO destinations of India and the Philippines, Western English-speaking companies are still discovering that they cannot simply transport communications training solutions and QA processes offshore without considering the language needs of the new NNS cohorts of CSRs. India was one the first call centre destinations that responded to problems of communication breakdown by running accent neutralisation programs (Cowie 2007) and grammar classes that focused on the contrastive differences between English and the Indian languages. In fact, a whole
training industry has sprung up in India to meet this perceived need. Unfortunately, these training solutions have now been transplanted into the Philippines as the BPO industry expands into new Asian destinations. However, there has been little needs analysis of the cause of call centre communication breakdown in NNS destinations. Phonological problems are not a root cause of why call centre communication is breaking down in the Philippines (Lockwood and Forey 2007). Friginal (2007: 335) further comments on this problem in the research he carried out in the call centre industry in the Philippines with particular reference to the lack of appropriate English as a Second Language training materials:

Because of limited training materials designed for outsourced CSRs, many language training programs in call centres in the Philippines use materials from the US and those that are available on the market. These references and activity manuals on call-handling practices and mock transactions are primarily written for native speakers of English or those with high-level language proficiency.

Trying to use such materials in a different culture and with CSRs with problematic levels of English proficiency will not show results.

The problem of untrained staff

Current high-stakes assessment practices within BPO companies, and particularly in call centres, are being carried out by lay people with no background in language assessment and applied linguistic practice. These practices are characterised by nothing more than a commonsense approach to the business requirement for communication information and measurement. These practices are computerised and regular communications metrics are generated as part of the business requirement. While the language assessment tools the author evaluated and the processes observed were developed in good faith and with a practical approach, they lacked validity, reliability and, in the end, an overall sense of fairness and justice to the CSR. Furthermore, these flawed language assessments were more expensive and more time-consuming than was necessary and, tellingly, were not getting the numbers of new CSRs into the company at time of ramp up. The ability to carry out these language assessment functions is threatened by a lack of TESOL trained personnel and consulting services available to call centres. The majority of the language training team members in Call Centre X had degrees but no formal TESOL qualifications.

The problem of staff with poor levels of English

Hand in hand with the general paucity and quality of data generated by Call Centre X on the language levels of the CSRs, there was also the issue of the language level of those tasked with improving the CSR communication skills. There was evidence during the consultancy in Call Centre X that the QA specialists and the HR department interviewers had proficiency language levels lower than the levels needed for recruitment and to diagnose the need for communications support. This is highly problematic and call centre staff members in charge of language assessment were like ‘the blind leading the blind’. It was extremely wasteful of both time and resources. A final recommendation was made to assess the language proficiency of all key communications assessment personnel in Call Centre X.

The problem of a fragmented approach and no shared metalanguage

Different departments within Call Centre X invented different language assessment solutions to cope with their specific needs and this led to serious fragmentation and no shared metalanguage to talk to each other about the English communication levels of the CSRs. The advantage of having a systematic language assessment across all departments is that the business is using only one way of measuring communication skills. In doing this, the business is ‘talking the same language’ when assessing for communication skills and making business decisions and reporting outcomes.

A systematic language assessment approach will ensure that auditing, screening and interviewing, training outcomes, coaching support, evaluating quality and migrating to different accounts can all be made consistently and with tools and processes informed by theory and best practice. This finding was reinforced in a report commissioned by a large publishing house into the industry need for English language assessment in 2006. This study focused, among other industry sectors, on the language
What does the Business Processing Outsourcing industry want from English language assessment?

The level of satisfaction with the current commercially available business English tests was very low. All but two of the call centres and BPOs surveyed in India and the Philippines had developed their own tests. One used BULATS [Business Language Testing Service] and one a test from a commercial assessment company MeriTrac. The main complaints were that the tests did not assess the areas of performance necessary, that they took too long to get results and that the test results were unreliable given the job requirement for communication. This led to almost all call centres using interviews as the main language test (with highly subjective outcomes) (From a market research report for a publishing company).

Constraints of existing business English tests available commercially

The constraints of the existing business tests in the BPO sector are very real. Centralised and large-scale business English tests such as TOEIC (Test of English for International Communication), BULATS and BEC (Business English Certificate) are run at designated times, are charged on a per-head basis, have a turnaround time of a minimum two weeks, give a one-point-in-time score that is not diagnostic, and are controlled by trained language assessors calibrated by the centralised examination body. A per-head cost, typical of the commercial business testing model, is prohibitively expensive (even if the cost is modest) for the call centre industry that accepts only 1–5% of the applicants that they test at interview.

The BPO industry will take responsibility for language training, so they require diagnostic information as part of their training needs analysis and this is a requirement for any language assessment. Given the high cost of training, the BPO industry needs to be able to measure outcomes of this investment, which is currently done on post-training assessments. Call centre managers need to know the communications demands of all their accounts and they need to provide quality scores on the communication success of the calls to meet their contractual obligations. No commercial testing organisation can provide this breadth of service, nor can any one-point-in-time proficiency test meet their needs. As a consequence, the industry is doing most of this itself with, in some cases, disastrous results.

The centralised model of language assessment, typical of Education Testing Services (ETS), Cambridge ESOL and other language testing bodies seeking high scalability and financial returns, is coming under serious challenge from the BPO companies. The BPO industry wants to control costs and wants to meet the business requirements by owning the language assessment tools and processes itself. This requires a new way of approaching language assessment in industry. At present there is a situation where language-testing vendors are unaware of the business needs, and businesses do not know enough about current assessment practices to articulate their problems and possible solutions that practitioners and researchers can respond to.

The development of the Business Processing Language Assessment Scales

BUPLAS was developed in Manila in 2003 because of the lack of communicative testing practices in the BPO industry. The test scales, descriptors and tasks were specifically developed by the author to address the English language assessment needs of the BPO industry in:

• recruiting staff with the right language skills
• identifying the language needs of new hires for training
• providing information about the shortfall of those CSRs who are almost at the right level but require more training
• providing training needs analysis information by generating diagnostic profiles
• conducting pre-program and post-program assessments
• auditing job functions and different accounts/ departments within the company
• evaluating quality on the job in terms of English communication
• providing English language assessment for internal promotion.

BUPLAS was used for the consultancy for Call Centre X and was subsequently mapped into all their business processes, including the QA account client...
scorecards. In developing BUPLAS, the consultant aimed to bring into the BPO industry current research and best practice in language assessment. BUPLAS is based within a sociolinguistic framework and assesses the sociolinguistic criteria of:

- discourse capability – the ability to speak logically when, for example, explaining, instructing, describing a product/process etc in an extensive turn
- strategic competency – the ability to repair communication when it breaks down (eg asking for repetition and clarification)
- interactive capability – the ability to build relationships with people (eg customers on the telephone)
- sociolinguistic competency – the ability to understand the culture embedded in the language, such as nuances in what is said, idiomatic expressions and paralinguistic features (eg how to treat silences, sighs, jokes etc).

These sociolinguistic criteria were felt to be very important when making assessments of candidates working in the BPO context, particularly in call centres and for email/chat customer support, because cultural understanding and relationship building are key competencies in this industry.

BUPLAS has now been embedded into the business processes and transferred into six call centres in Manila and forms the basis of consulting assessment projects in call centres in India and the Philippines. It is sold at a modest licence fee cost.

**Conclusions and further research**

Call centres require embedded end-to-end language assessment solutions serving a number of business requirements, such as the selection and recruitment of new CSRs, account auditing, pre-course and post-course training, ongoing coaching and mentoring using language measurement, and communications QA reporting. These homegrown language assessment solutions developed by local managers had been computerised and were used to closely monitor the communication levels of each CSR from the time of recruitment to regular QA reporting once on-site. One problem was that language communication scores yielded from such homegrown solutions bore little relation to the CSAT scores, and the call centre received regular complaints from client groups about serious communication breakdown on the telephones, despite favourable feedback from the QA specialists and local communications scoring approaches. The other problem was that the language assessment processes were invalid and unreliable.

Call centre communication research needs have been brought into sharp focus because of the experience English-speaking multinational companies are having in non-English speaking offshore and outsourced destinations such as India, the Philippines, South America, China and Eastern Europe.

An agenda for English language assessment research in the future may focus around research and development issues that pose the following questions:

- How many potential CSRs are being lost at recruitment due to flawed assessment tools and practices?
- How fast does the language proficiency level of CSRs develop once they are on the telephones?
- What is the nature of the communication breakdown on the telephones?
- Can communication breakdown be identified and diagnosed through conventional language assessment scales and descriptors?
- What are the LSP needs of language assessments in call centres?
- Can non-English language specialists be trained as language assessors?
- Is there a threshold language proficiency level that the language assessors must have in order to carry out reliable language scoring?
- How can scorecards be informed by communicative language assessment frameworks?
- How does research into world Englishes inform call centres about levels of English to train for and expect?
- Would an interdisciplinary approach assist applied linguists when working on language assessment solutions for the workplace?

The findings from this consultancy research
in Call Centre X perhaps raise more questions than they answer. However, they did provide pragmatic success in meeting the business needs and a blueprint to move forward. There is no doubt that a prolonged and systematic involvement of language assessment researchers and practitioners in the field, including sustained dialogue with stakeholders, will positively impact the current workplace practices in call centres. Call Centre X reported at the conclusion of the consultancy a commitment to the frameworks and the recommendations suggested, which they methodically adhered to over the next 18 months with improved results. No follow-up evaluation study has yet been undertaken.

Reflecting stakeholder needs into the assessment and evaluations processes of call centres necessarily opens up the importance of incorporating these new sets of perspectives into the whole curriculum and quality process. Perhaps applied linguistic researchers and English Language Teaching (ELT) workplace practitioners will require new knowledge and new tools to better understand the workplace communities they are training and assessing. Wyndham (1998: 12) states that:

Corporate managers need the right tools to measure the English proficiency of their employees. Human Resource development managers need the right tools to measure which training programs yield the ‘biggest bang for the buck’ (or yen as the case may be). Trainers and assessors need the right tools to help them gauge the quality of programs and provide information on how to improve these programs. What are needed are instruments that measure both the learning that takes place as a result of training and, more importantly, the changes in workplace behaviour resulting from the training and support.

Notes
1 BUPLAS is owned by Lexicon Consulting Group Ltd

References


Cleveland, B., & Mayben, J. (2004). Call center management on fast forward: Succeeding in today’s dynamic inbound environment. Maryland: Call Center Press, A Division of ICMI, Inc.


What does the Business Processing Outsourcing industry want from English language assessment?


Lockwood, J. (2004). Consultancy report for Call Centre X


