

Review

Issues in the development of cross-cultural assessments of speech and language for children

Julie A. Carter†‡, Janet A. Lees†, Gladys M. Murira‡,
Joseph Gona‡, Brian G. R. Neville† and Charles R. J. C.
Newton†‡

†Neurosciences Unit, Institute of Child Health, The Wolfson Centre, London, UK

‡Centre for Geographic Medicine Research (Coast), Kenya Medical Research Institute, Kilifi, Kenya

(Received 26 July 2004; accepted 21 December 2004)

Abstract

Background: There is an increasing demand for the assessment of speech and language in clinical and research situations in countries where there are few assessment resources. Due to the nature of cultural variation and the potential for cultural bias, new assessment tools need to be developed or existing tools require adaptation. However, there are few guidelines on how to develop ‘culturally appropriate’ assessment tools.

Aims: To review the literature on cross-cultural assessment in order to identify the major issues in the development and adaptation of speech and language assessments for children and to illustrate these issues with practical examples from our own research programme in Kenya.

Methods & Procedures: Five broad categories pertaining to cross-cultural assessment development were identified: the influence of culture on performance, familiarity with the testing situation, the effect of formal education, language issues and picture recognition. It was outlined how some of these issues were addressed in our research. The results of the review were integrated to produce a list of ten guidelines highlighting the importance of collaboration with mother tongue speakers; piloting; familiar assessment materials; assessment location; and practice items and prompts.

Conclusions: There are few clinicians and assessors, whether in the UK or abroad, who do not assess or treat children from a culture different to their own. Awareness of cultural variation and bias and cooperative efforts to develop and

Address correspondence to: J. A. Carter, Centre for International Child Health, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK. e-mail: j.carter@ich.ucl.ac.uk

administer culturally appropriate assessment tools are the foundation of effective, valid treatment programmes.

Keywords: assessment, culture, child, language, speech.

Introduction

There is an increasing demand for the assessment of speech and language in clinical and research situations in countries where there are few assessment resources. Fundamental to the provision of an equitable speech and language therapy service is the recognition of cultural variation and the potential for cultural bias, in order to improve diagnostic accuracy and treatment outcomes (Isaac 2002). In research situations, the lack of appropriate assessment tools can be equally damaging, confounding the results and producing biased, inaccurate conclusions. In consequence, there is a growing recognition of the necessity of developing or adapting assessment tools and procedures to match the needs of these populations. However, there are few guidelines on how to develop 'culturally valid' assessments.

Culture may be defined as the set of values, beliefs, perceptions, institutions, technologies, survival systems and codes of conduct held by members of a particular group of people (Payne and Taylor 2002). The degree to which assessment components reflect the cognitive goals that are set for individuals in the target culture defines the 'ecological validity' of the assessment tool (Mishra 1996). Taylor and Payne (1983: 11) state that culturally valid assessment is

a data collection process wherein testing, measurement and evaluation are conducted using instruments and procedures that discriminate only in those areas for which they were designed (i.e. normal versus pathological behaviour) and do not discriminate unfairly either for or against a client for cultural reasons

Thus, the task of cross-cultural assessment is to identify and distinguish between behaviours that are universal, those that are culturally variable and those unique to the individual.

Background to cross-cultural assessment

Assessments designed for cross-cultural situations were first produced in the early twentieth century. The hypothesis underlying many of the first assessment tools was that behaviour is independent of the surrounding culture and merely superimposed by a cultural veneer that can be penetrated by what were termed 'culture-free' assessments. The conceptual framework of empiricism, which lay behind most scientific endeavour, posited an uninterrupted link between the brain, cognitive processes and behaviour, leading to the belief that the most reliable and valid assessment results are those obtained in a context-free clinical setting using culture-free assessments (Pérez-Arce 1999).

The later adoption of 'culture-common' or 'culture-fair' assessments demonstrated the realization that culture permeates all aspects of behaviour. These tools reflected the hypothesis that assessments could target only those experiences and expressions of behaviour common to different cultures. The use of non-verbal performance items in preference to verbal items was believed to diminish the effects

of culture, although there is now ample evidence that non-verbal assessments are also susceptible to cultural bias (Rosselli and Ardila 2003).

Anastasi and Urbina (1997), reflecting more recent opinion, comment that a single assessment cannot be universally applicable to all cultures but can only aim to reduce cultural differentials in performance. They conclude that a child's behaviour is influenced by the cultural milieu in which he or she is raised. As many developmental assessments are measures of a behaviour sample, the influence of culture will and should be detectable, a framework of analysis known as 'situated cognition' (Pérez-Arce 1999). Eviatar (2000) argues that culture should be considered a variable, similar to age or sex, that can be crucial in defining the manner in which higher cognitive processes, such as language, are related to brain organization.

Boivin (1991) found that Zairian children performed significantly below the age-related norms of US children on a number of assessments that have been widely regarded to be non-verbal, culture-fair measures of various facets of cognitive abilities: the Matrix Analogies Test (Naglieri 1985), the tactual performance task of the Halstead-Reitan Neuropsychological Assessment Battery (Reitan and Wolfson 1985) and the simultaneous and sequential reasoning subsections of the Kaufman Assessment Battery for Children (Kaufman and Kaufman 1983). He concluded that despite the 'culture-fair' label, all of the assessments originated within the context of Western psychological research and theory, thus any validity in the cross-cultural situations in which they have sometimes been used is coincidental. Similarly, Stanczak *et al.* (2001) found that normal Sudanese adults attained scores on the Arabic version of the Expanded Trail Making Test that were similar to those attained by US adults with brain damage. The authors concluded that cultural variables must be considered in the design and administration of existing assessments to new populations to minimize Type I and II diagnostic errors.

Review

Assessment tools reflect the values, knowledge and communication strategies of their culture of origin (Greenfield 1997). In comparison to the number of published assessments for majority UK populations, there are few available assessments designed for other cultural groups. Thus, in many situations, the clinician or researcher faces the task of designing novel assessments or adapting existing tools. This paper aims to review the literature on cross-cultural assessment in order to highlight the issues germane to the process of assessment development and adaptation. We have used the broader literature of cognitive testing, rather than limiting ourselves to the assessment of speech and language, because the field of cross-cultural psychology has a longer history (and more publications) than that of cross-cultural speech and language assessment but many of the issues are similar.

Observational and experimental studies of cognitive or speech and language assessment conducted outside the UK, descriptive papers about the development of tools for use in cross-cultural situations and review papers were included in the review. The specific issues pertaining to bilingualism and cross-cultural assessment of ethnically diverse groups in minority world countries are outside the scope of this paper and are dealt with elsewhere (Isaac 2002, Mahon *et al.* 2003, Marshall 2003,

Stow and Dodd 2003). However, many of the points discussed in this review are relevant to any situation in which the clinician or researcher is working with a person from a culture different to their own.

We have encountered many of the issues raised here as part of a research programme in rural Kenya and include some of our experiences during the review to provide practical illustrations of how these issues can be addressed. The research programme aimed to determine the long-term developmental outcome of severe malaria in children. We followed up 308 6–9-year-old children previously exposed to neurological manifestations of malaria (either cerebral malaria or malaria and complicated seizures) and measured their performance using assessments of cognition, motor skills, speech and language, behaviour, hearing and vision. We compared their performance with that of 179 children unexposed to either condition. The results are presented elsewhere (Carter *et al.* 2005a, b, 2006). Before starting data collection, we developed a battery of speech and language assessments (measuring receptive grammar, receptive vocabulary, syntax, lexical semantics, higher level language, pragmatics, phonology and word finding), as there were no detailed assessments available in the local language, Kigiryama. Although our experience is of assessment in a research setting, we believe the conclusions are also relevant to clinical settings.

The review is divided into five broad categories, each addressing one of the key issues highlighted in the literature: the influence of culture on performance, familiarity with the testing situation, the effect of formal education, language issues and picture recognition.

Influence of culture on performance

The experiences available in the child's immediate physical and social environment have a direct effect on his/her development and consequently, on the range of appropriate materials with which it can be measured. An example of specificity of development is the apparent precocity of early gross motor development in rural African children (Super 1976, Freedman and DeBoer 1979, Cintas 1988). Studies examining other skills have found similar effects. For example, Australian Aboriginal children and Amazonian Indians have been found to have significantly better visual-spatial memory abilities than their Western counterparts, perhaps because their environment constrains them to develop an aptitude for direction finding (Kearins 1981, Rosselli and Ardila 2003).

Cross-cultural researchers disagree as to whether such differences represent maturational lag in certain cultures in what is otherwise a universal sequence in the development of cognitive processes or whether there are qualitatively different end points in development (Dasen 1988, Bates 1997). However, what such findings suggest is that the familiarity of materials and the content and structure of a task will influence whether the child's performance is a true representation of his/her ability. Miller *et al.* (1984) describe a scenario in which the use of materials inappropriate to a cultural situation disrupted the assessment procedure. In their study, the Denver Developmental Screening Test was administered to Hmong (Laotian) children, whose families had been living in the USA for between 1 and 18 months. Most children refused to pick up a raisin (a task designed to measure fine motor skills) because it looked like a common medicine they had traditionally been told to avoid.

If we accept that language processing is, to some extent, culturally bound but interpret results acontextually, we are unlikely to reliably predict how the child will act in his/her own environment or what compensating mechanisms may be facilitated or weakened within the family or community context.

In addition to the materials used in language and cognitive assessments, the assumptions underlying the assessments may be incongruous in cross-cultural situations. For example, Greenfield (1997) tested the Piagetian stage of concrete operations with unschooled Wolof children in Senegal. The children were able to give quantity judgements about water transferred from a short, fat cup to a long, thin one but did not understand the question 'Why do you think it is the same/more/less?' or a slightly reworded question: 'Why do you say it is the same/more/less?' Only when the children were asked 'Why is the water the same/more/less?' were they able to give justifications for their original quantity judgement. Greenfield concludes that, in contrast to the assumptions underlying the structure of the questions, Wolof children had an 'epistemology of mental realism', not distinguishing between their own thoughts or statements about something and the object of knowledge itself.

Greenfield's experiences underline the need to be sensitive to the fact that the normative strategies for processing information in British culture — rational and dualistic thinking — are not necessarily the key ingredients for successful work or education in other cultural groups. Helms (1992) lists what she calls the 'Eurocentric' values and beliefs that have influenced many cognitive and language assessments, including individual achievement, competitiveness, the value of speed, adhering to time schedules, the emphasis on logical thought and objectivity and the belief that performance can be quantified. For example, there is evidence that the concepts of fast or accurate performance vary between cultures. Mulenga *et al.* (2001) administered the Developmental Neuropsychological Assessment (NEPSY) (Korkman *et al.* 1998) to Zambian children and found that despite clear instructions to perform the task as quickly as possible, most children worked slowly. Similarly, Zairian children have been found to be slower than US or Canadian children in the Trail Making Test (Boivin *et al.* 1995). Differences in normative strategies also apply to communication. For example, many observational assessments of pragmatics measure eye contact, which is seen as a component of successful face-to-face communication in the UK. However, in rural Kenya, children are taught that it is disrespectful to make eye contact with adults who are in an authoritative position. Therefore, the normative strategy for a rural Kenyan child would be to lower his or her eyes when conversing. Pragmatic bias is recognized as a pervasive issue in cross-cultural communication research and differences in conversational and discourse rules may occur in turn-taking during conversations; interruptions; silence as a communicative device; appropriate topics of conversation; the use of humour; appropriate amount of speech; how to open and close a conversation and non-verbal strategies such as eye contact (Payne and Taylor 2002).

Familiarity with the testing situation

The social rules of language interaction — who may speak to whom, in which situations — are culturally bound. In many cultures, children have no freedom to

engage in prolonged dyadic play with adults (Mbise and Kysela 1990) and spend much of their time in polyadic situations with siblings or other children (Lieven 1994). For example, our experience suggests that it is unusual for Kenyan children to sit and converse with an adult, especially a strange adult, so the expectation of most language and cognitive assessments — for a child and adult to interact — is an unfamiliar activity. In many cultures, knowledge flows from the top down and adults do not normally solicit the views of children. Beliefs about the relative value of individual versus group perspectives also differ. In Mexico, Greenfield (1997) conducted individual interviews with children but found that her worldview — that each girl would have an independent viewpoint or perspective — conflicted with that of the community, who expected family groups to answer cooperatively, the group effort producing the most valid information.

Lack of familiarity with aspects of the testing situation in children from different cultural groups has long been cited as a possible reason for low performance levels (Brislin *et al.* 1973, Miller-Jones 1989). Saxe (1988) found that Brazilian children who sell sweets on the street displayed highly sophisticated reasoning and problem solving abilities when confronted with everyday mathematical tasks, although they performed poorly in classroom mathematics tests. More recently, Peña *et al.* found that Puerto Rican children in the US performed significantly better on tasks requiring descriptions than tasks requiring object labelling, which was an unfamiliar activity, not being taught in their home or community setting (Peña *et al.* 1992, Peña and Quinn 1997). This hypothesis is supported by the fact that practice and prompts enhance performance levels in African children previously unfamiliar with standardized testing situations (Brislin *et al.* 1973, Mwamwenda 1992). However, unintended disadvantages may accompany the benefits of adaptations. For example, we used a teach–test format for a non-verbal functioning assessment, in which the child received prompts and verbal feedback for each item. This extended the length of the task compared to the original format, as a result of which many children complained of boredom.

Functional equivalence concerns the extent to which cognitive/language ability, as defined by the clinician, operates in the same manner between cultures. For example, a fundamental conversational convention about the question form is that a questioner is asking for information that he/she lacks. However, most assessment questions (particularly those with a right/wrong answer) presuppose that a questioner who already knows a piece of information can sensibly ask a listener for the same information, essentially requesting redundant information. Assessment questions are also a key component of teaching, therefore children with some experience of formal education are usually more familiar with them (Greenfield 1997). However, for many cultural groups, answering questions to satisfy an adult's request, rather than to accomplish a functional goal is counterintuitive (Helms 1992).

Formal education

Many children in majority world countries do not attend school or their attendance is spasmodic, depending on the family's financial circumstances. Formal education is hypothesized to affect assessment performance, possibly mediated by the effects of literacy or linguistically mediated problem-solving capacity and strategies for

completing such tasks (Pérez-Arce 1999). The schoolchild may be more familiar with assessment strategies, thus less likely to make errors due to misunderstanding task demands. In addition, the epistemological assumptions of school culture — that materials exist whose only purpose is to test — gives schoolchildren an advantage in understanding the framework of assessments (Greenfield 1997).

The response requirement of most language and cognitive assessments is for the child to exhibit their best behaviour to a stranger (the assessor) in a structured, often unfamiliar environment (Baine 1990). In investigations among the Kpelle people of West Africa, unschooled children performed poorly compared to their schooling counterparts on memory tasks such as the free recall of word lists (Cole and Scribner 1974). Their performance level improved when cues were used in the unstructured tasks or when naturalistic story tasks were used. Similarly, in our work in rural Kenya, unschooled children performed significantly poorer than those attending school on assessments of receptive skills, syntax, vocabulary, word finding, higher level language skills, memory and non-verbal skills.

Schooling of poor quality may not diminish all of the positive effects on cognitive performance. Das (1992) found that schooling of comparatively poor quality in remote areas of India enhanced children's performance on tasks of simultaneous and successive processing and planning. Conversely, differences in measures of recall have been found between Indian children attending 'good' schools (characterized by sufficient space for staff and children, transportation, trained teachers, library and recreational facilities and the use of new teaching technology) and those attending 'ordinary' schools (characterized by a relative lack of these facilities) (Mishra 1996). Mwamwenda and Mwamwenda (1991) assessed performance on Piagetian tasks in Batswana children attending English-medium and Setswana-medium schools. They hypothesized that the fee-paying English-medium schools would deliver superior quality education compared to the government-funded schools working in Setswana, which generally employ teachers with less or no training and have fewer educational resources. The results of the study showed significantly better performance in children from English-speaking schools relative to those from Setswana-speaking schools on two of the three tasks (class inclusion and conservation). However, socio-economic status was not taken into account in the study and may have influenced the performance of the respective groups.

There is disagreement as to whether it is the effect of formal schooling itself or the impact of literacy that has the most profound effect on cognitive processes (Mishra 1996). Literacy in itself is hypothesized to promote the development of different conceptual frameworks and there is empirical evidence that literacy promotes phonological awareness, explicit visual analytic skills and working memory, as evaluated by digit span (Morais and Kolinsky 2000). Mishra (1996) describes two effects of literacy at the individual level proposed by Scribner and Cole (1981). First, the assimilation of knowledge and information transmitted by written texts promotes the growth of the mind and second, influences the content of thought and the processes of thinking. According to Olson (1976), the oral relay of information via proverbs, stories and other mnemonic tools is most commonly encountered in predominantly illiterate societies. The content of information in such societies focuses on the functional description of concrete objects. By contrast, in literate societies, communication is relayed by written prose and content is more often related to relationships between more abstract concepts.

Language issues

Language plays an essential role in cognitive development, first as a means of social communication and later in shaping cognitive processes and organizing thought (Vygotsky 1962). The relationships between culture, language and cognition have been discussed in the fields of philosophy, anthropology and linguistics, as well as psychology. The Sapir–Whorf linguistic relativity hypothesis is one of the most famous of these debates, positing that culture, through the medium of language, influences people's thinking (Whorf 1956). Cultural influences on cognitive processes — what Mishra (1996) describes as the 'subjective orientation' of languages — have been described in ethnobiology, in which the economic importance of plant and animal species has been found to determine classification systems (Mishra 1996); categorization, in which some cultural groups categorize objects functionally rather than taxonomically (Sternberg and Grigorenko 2001); picture interpretation, in which children speaking different languages were found to interpret pictures in different ways (Slobin 1991) and time, in which the way English speakers and Mandarin speakers talk about time (as if it were horizontal for English speakers; as if it were vertical for Mandarin speakers) was found to correspond to their relative conception of time (Boroditsky 2001).

Language is a medium for conveying and internalizing culture and is thus embedded in culture (and vice versa). Many studies conducted in countries in which assessments have not been standardized on the population have used translated versions of UK or US assessments. However, it is important to consider the language environment of the host country, as a translated question may not elicit the same response as originally intended. Sperber *et al.* (1994) argue that translation of assessments from the source language to a second, 'target' language is complex and often does not overcome the problem of biased or unfair assessment. Practical problems include test directions, which are often technical and difficult to translate; ensuring equivalence of meaning, which necessitates several translations back and forth between languages; the fact that the psychological constructs underlying the assessment are not always universal across cultures and the variation in test-taking behaviour and orientation towards assessment procedures between cultures (Sperber *et al.* 1994).

We adapted the Test for the Reception of Grammar (TROG) (Bishop 1989), altering the content of items to reflect the structure of Kigiryama (the local language), while maintaining the principles of using a restricted vocabulary, easily recognizable pictures and lexical and grammatical distractors. For example, questions testing the third person pronoun were irrelevant ('he/she' are both represented by 'a' in Kigiryama), thus were removed and the singular/plural question was expanded to take account of the complex structure of noun classes in the language. Syntactic differences may also be found between cultures using the same language, as reported by Baddeley *et al.* (1995), who attempted to use the TROG with children speaking Jamaican patois. Adaptation and revision are usually required in addition to translation because item content and assessment format may be more familiar in one culture than another.

These issues are further illustrated in a study in which a translated version of the Denver Developmental Screening Test was administered to Vietnamese children (Miller *et al.* 1984). One task required the children to select coloured blocks in response to a particular colour being named by the examiner. The Vietnamese

children seemed to consistently confuse the colours blue and green: it later became apparent that the names for these two colours were from the same overall colour group in their language, thus indistinguishable without a qualifying description. Similarly, Howard and De Salazar (1984) adapted the Denver Developmental Screening Test for Costa Rican children. Although they translated the assessment into Spanish, piloting highlighted further changes to the content needed for valid administration. For example, the instruction 'draw a man' did not elicit the desired response, nor did the replacement of 'man' with 'boy', 'girl', 'mother' or 'father'. However, there was an almost unanimously positive response to the instruction 'draw a doll'. These examples highlight the importance of having someone familiar with the language of the area involved at every stage of assessment construction and administration.

Pahl and Kara (1992) assessed the applicability of the Renfrew Word Finding Scale (RWFS) to the South African context, administering the assessment to 60 white or Indian children from English-speaking homes of a 'middle socio-economic group'. They found that only six children obtained scores equal to or above their chronological age levels using British norms and that 39 performed at a level suggestive of expressive vocabulary problems, despite the fact that their teachers considered their language abilities to be normal. Specific items on the assessment were consistently named incorrectly: for example, none of the Indian children were able to name 'spire/steeple', although nearly half of the white children named it correctly, suggesting that lack of familiarity with the items was responsible for the errors. Pahl and Kara (1992) concluded that although the unmodified RWFS was not suitable for any of the South African children in their study, modifications could make the tool appropriate for the white children in their study. Interestingly, they considered that the lexical differences between the South African English speakers and the South African Indian English speakers to be sufficiently great that even with modifications, the assessment would be unsuitable for Indian children. This study highlights the fact that even in children from the same overall language group and socio-economic background, cultural differentials may still influence performance.

Other language issues to consider when developing assessments are those common in multi-lingual societies. Many African and Asian countries, although having one national or official language, have numerous indigenous languages. For example, English and Kiswahili are the official languages of Kenya but there are 61 languages spoken in the country representing each of its major ethnic groups and other emigrant groups (Grimes 2000). The children we assessed were part of the Mijikenda group, which represent nine distinguishable ethnic and language groups of the coastal Kenyan strip. All of the Mijikenda languages are closely related and are, to a large extent, mutually understandable. Exposure to Mijikenda languages other than the child's mother tongue usually occurs early in life, thus most older children attain a level of understanding and fluency in other languages spoken in the geographical area and the number of cognate words can be as high as 89% (Van Otterloo and Van Otterloo 1980). As Kigiryama (the largest of the Mijikenda languages) was the mother tongue of most of the children in our area, we selected it as the language of assessment in our study. Occasionally, children expressed a preference for assessment in one of the other Mijikenda languages, therefore we employed some assessors from Mijikenda groups other than the Kigiryama. In our case, the level of mutual comprehension between ethnic groups facilitated the transition of assessment content into the different Mijikenda languages.

Similar issues may affect assessment in predominantly monolingual societies in which there are different dialects, therefore the assessor must also be careful that their assessment does not unfairly penalize speakers of non-standard dialects, even when the language of assessment is the same (Payne and Taylor 2002). Speakers of non-standard dialects have been reported to be more sensitive to the context of assessment, sometimes reducing their verbal output when relating to a standard dialect speaker or to 'hypercorrect' their speech (i.e. inappropriately apply standard dialect features) (Seymour and Miller-Jones 1981).

Picture recognition

The majority of assessments of language and cognition use picture stimuli. As with language, picture recognition involves symbolic representation. Children begin to recognize pictures of objects as early as 5 months of age (De Loache *et al.* 1979). Reynell (1980) described the development of symbolic understanding as beginning with the recognition of real objects, progressing to that of miniature objects then the understanding of more arbitrary symbols such as the recognition of photographs of an object. Finally, the child learns to recognize stylized drawings of objects. A picture is not an exact replica of an object, merely a two-dimensional representation that presents certain cues, which as a result of previous experiences, trigger the perception of the object (Anastasi and Urbina 1997).

Some cross-cultural studies have suggested that the ability to recognize and interpret pictures is culture-specific. Serpell and Deregowski (1980) consider picture recognition to be a 'culturally restricted perceptual skill': Western children learn to perceive pictures in the functional contexts of learning to speak and listening to stories with the help of picture books, activities unavailable to most rural African children. Hudson (1960, 1962) described picture recognition as related to literacy, rather than solely to culture, finding that non-literate Bantu and European participants were unable to interpret a picture in three dimensions. Other studies have found that lack of previous exposure to pictures does not negatively affect recognition (Deregowski 1968) and that advanced age, schooling or an especially high level of intelligence are not essential for successful picture recognition (Kennedy 1974).

We found that rural Kenyan children, especially those not attending school, had difficulties in interpreting picture stimuli. Some children confused the target item with objects of a similar shape (e.g. 'pencil' instead of 'broom') or knew the definition of an object but had never seen either the actual object or a picture of it (e.g. most children knew what a crocodile was but did not know what one looked like). Therefore, we piloted all picture stimuli with children from the target population for level of recognition. If more than 80% of children recognized the picture, it was accepted; otherwise it was redrawn or discarded. This level of recognition was selected to represent the majority of children but not inordinately alter the assessment materials for the small numbers of children who gave incorrect answers during piloting. A second issue, relating to picture perception, concerned the assessment of prepositions, which we originally tested by asking children to name pictorial representations of prepositions. During piloting, most children found this task problematic, especially the items 'in front' and 'behind': more than 75% of children responded 'on top' or 'under', possibly due to

difficulties in understanding the manner in which depth is represented in two dimensions (Serpell 1972). Consequently, we decided to use real objects in further assessments.

Approaches to the development of speech and language assessments in cross-cultural situations

The development of novel assessments in cross-cultural situations necessitates ethnographically based research on the language and culture (Ochs 1983, Schieffelin and Ochs 1986) with the aim of generating a developmental model for the intended study population. This approach is both resource- and time-consuming, making it impractical in many research and clinical situations. A number of alternative approaches are available: renorming an existing assessment for the target population, using dynamic assessment techniques to assess a child's 'modifiability' and modifying existing, standardized assessments.

Renorming a standardized assessment requires a large sample of typically developing children from the target population, preferably comprising individuals of the same ethnic, cultural and linguistic background, age, sex and educational level and type. The reliability and validity of the renormed assessment should also be established for the intended group. There are obvious advantages for the clinician or researcher to having an assessment for which there is culturally applicable normative data. However, in addition to the time and resources required for such an undertaking, renorming can be problematic in that there may not be suitable equivalents for some of the original words and concepts and pictorial stimuli may not be appropriate, meaning that the absolute and relative difficulty of assessment items may change (Saenz and Huer 2003).

Dynamic assessment typically involves a sequence of testing, including pre-testing, a teaching element and post-testing and focuses on how the child learns rather than what the child knows (Isaac 2002). The same assessment is administered at pre- and post-testing. The aim of the intervening teaching element is to teach the child concepts that will enable him/her to perform to maximum ability on the post-test. The child's scores on the two attempts are compared and analysed according to the theory that a child with poor pre-test results but normal language ability may achieve a substantially higher score on the post-test because they have a typical ability to learn the rules and procedures of the assessment when given adequate exposure and teaching (Peña *et al.* 1992). The aim of dynamic assessment is to overcome the bias in results from lack of familiarity with assessment requirements or formal education. The disadvantage is that repeating the assessment and teaching the child is time-consuming; our experiences of using dynamic assessment techniques for non-verbal functioning assessments in Kenya is that children become bored with the long assessment session and lose concentration. Another disadvantage of this approach is if the teaching element is not standardized, the reliability and validity of the scores may be compromised, which is problematic in diagnostic or research situations (Saenz and Huer 2003).

We adopted the approach of assessment modification, whereby assessments were derived from validated instruments in common use in the UK, when a Kenyan alternative was not available; or used assessment methods routinely in use in the UK, with revisions and modifications made to make them appropriate for the study

population. The adaptation or reformulation of standardized assessments, as undertaken by Mbise and Kysela (1990) in Tanzania, offers the chance to create culturally and linguistically valid assessment materials, yet follow an established framework. In this way, 'imported' assessments can be adapted or new assessments developed from the principles of standardized tools, using methods and materials derived from the local culture. This enables the clinician/researcher to determine the assessment's cultural equivalence (whether the construct has similar meanings within and across cultural groups) and functional equivalence (described earlier) (Pérez-Arce 1999).

Pakendorf and Alant (1997) comment that when assessment items are translated, omitted or altered in content, pictorial stimuli are adapted for suitability to the culture and the sequence of presentation of assessment items is reordered, it is debatable how many of the original characteristics are retained. Thus, in a research setting, it is important to have a group unexposed to the condition of interest, drawn from the same population as those exposed. Sbordone (1996) discusses the ecological validity of the normative group, stating that if the exposed group does not 'fit' the unexposed group in attributes such as age, sex, cultural and linguistic background, any comparisons between the groups or judgements about the presence of impairments will be spurious. This is particularly pertinent in the situation of transferring assessments across cultures. For example, Anderson (2001) found that the use of imported normative data resulted in an unacceptably high diagnostic rate of neuropsychological impairment in healthy South African adults and emphasized the importance of establishing norms and assessment criteria in the local population. In addition, it is essential that assessors are blind to the group status of the participants to prevent bias in scoring decisions.

Considerations in cross-cultural assessment development

Drawing upon the literature and our experiences in Kenya, the processes outlined below are suggested in studies attempting to create 'culturally valid' tools and assessment procedures:

- Tools should be developed in conjunction with mother tongue speakers of the assessment language, who grew up in the local area (thus familiar with the culture). We specifically employed local fieldworkers and assessment specialists but local teachers are often a good source of advice.
- Adapted assessments should be piloted on a representative sample of the target population, which should take account of age, sex, socio-economic status and schooling status. Children in majority world countries often face multiple risk factors to their health and development, therefore samples should not be drawn exclusively from populations that have been exposed to infectious diseases, severe malnutrition or other conditions that may seriously affect their performance.
- All pictures should be piloted with children from the target population for level of recognition before inclusion in assessments. This is a simple procedure whereby the assessor presents each picture to the child and asks what it is. The sample must include unschooled children (if they will be part of the population the tool will eventually be used with), as they may be less familiar with interpreting picture material. Any pictures consistently named

incorrectly should be redrawn or discarded. If the target population predominantly consists of children who are unfamiliar with pictures, consider using real objects or photographs instead of drawings.

- All aspects of the adapted tool should be piloted: not just the assessment itself but also the instructions, response formats, practice items, prompts and setting.
- Mother tongue speakers of the assessment language should be trained to carry out or assist in assessment procedures. The focus of cross-cultural testing has recently shifted from the assessment itself to the assessor's behaviour in the testing situation, which needs to be adapted to the needs of the test-taker (Anastasi and Urbina 1997). Assessors from the test-takers' ethnic group and language background would be assumed to make such adaptations most effectively.
- Assessment materials familiar to children in the local area should be used. This ranges from the toys used in the 'settling-in' period to the picture material used in the assessment itself. Simple but often overlooked items such as skin colour of people in pictures, style of houses, types of trees and flowers make a difference to children's comprehension of the assessment.
- When possible, children should be assessed in their own homes or in rooms away from a hospital or clinic setting to minimize the unfamiliar aspects of the assessment situation.
- When working with children who are unfamiliar with the testing situation, consider using practice items and prompts, where appropriate, to reduce the chance of errors due to misunderstanding of task requirements. However, also consider the impact on assessment length.
- When designing timed tasks, consider the cultural view of speed and performance and whether the task will measure the desired construct. Advice from local fieldworkers or teachers and piloting should provide this information.
- Having followed the above measures, the normative data should show the expected distribution. If not, consider possible reasons for any outliers and further modifications to the assessment that would result in a normal distribution. Piloting should be seen as an iterative process.

Conclusions

There is an increasing demand for speech and language therapists and psychologists in clinical and research situations in countries where there are few tools for measurement or therapy. This coincides with the growing evidence of the negative impact of infectious diseases (Carter *et al.* 2003a, b, 2005a, b, 2006), malnutrition (Chavez *et al.* 1995, Sigman 1995) and low socio-economic status (Bradley and Corwyn 2002) on speech and language functioning and the recognition that, due to the prolific nature of these conditions in many countries, there are large numbers of children in need of rehabilitation services. Even for therapists practising exclusively in the UK, the impact of asylum seekers and immigrant workers will increase the proportion of schoolchildren from minority groups from one in eight in 2001 to one in five by 2010 (DfES 2003), meaning that there will be few therapists who do not face the issues of cross-cultural assessment.

Of course, speech and language assessments are only a tool to understanding communication. Although emphasizing the need for culturally appropriate tools, the clinician must also be careful not to miss different styles of communication in which, for example by using gesture, simple signs or eye pointing, children function efficiently with relatively unsophisticated speech and language. This is especially important in communities in which high-level verbal and written communication skills are not the key components of a successful or fulfilling life. Ideally, therefore, formal assessments should be complemented by a measure of the social effect of the child's communication abilities to ascertain whether those identified as having speech and language problems would be recognized as functioning poorly in their homes and communities.

For the clinician assessing in a cross-cultural situation, there is inevitably the ensuing question of treatment. This brings its own issues, such as the clinician–patient relationship (Flores 2000), parental beliefs in the origin of the disorder and their view of other treatment options and support systems, such as traditional and faith healers (Semela 2001, Marshall 2003). However, awareness of cultural variation and bias and cooperative efforts to develop and administer culturally appropriate assessment tools are the foundation of effective, valid treatment programmes.

Acknowledgements

J. A. C. (059336) and C. R. J. C. N. (050533) were supported by The Wellcome Trust, UK. This paper was published with the permission of the Director of KEMRI.

References

- ANASTASI, A. and URBINA, S., 1997, *Psychological Testing*, 7th edn (Upper Saddle River, NJ: Prentice-Hall).
- ANDERSON, S. J., 2001, On the importance of collecting local neuropsychological normative data. *South African Journal of Psychology*, **31**, 29–34.
- BADDELEY, A., MEEKS GARDNER, J. and GRANTHAM-MCGREGOR, S., 1995, Cross-cultural cognition: developing tests for developing countries. *Applied Cognitive Psychology*, **9**, S173–S195.
- BAINE, D., 1990, Guide to the development, evaluation and/or adoption and modification of tests for early childhood education in developing countries. In M. J. Thorburn and K. Marfo (eds), *Practical Approaches to Childhood Disability in Developing Countries: Insights from Experience and Research* (Tampa: Global Age), pp. 235–263.
- BATES, E., 1997, Origins of language disorders: a comparative approach. *Developmental Neuropsychology*, **13**, 447–476.
- BISHOP, D. V. M., 1989, *Test for Reception of Grammar*, 2nd edn (Manchester: University of Manchester).
- BOIVIN, M. J., 1991, The effect of culture on a visual–spatial memory task. *Journal of General Psychology*, **118**, 327–334.
- BOIVIN, M. J., GIORDANI, B. and BORNEFELD, B., 1995, Use of the tactual performance test for cognitive ability testing with African children. *Neuropsychology*, **9**, 409–417.
- BORODITSKY, L., 2001, Does language shape thought?: Mandarin and English speakers' conception of time. *Cognitive Psychology*, **43**, 1–22.
- BRADLEY, R. H. and CORWYN, R. F., 2002, Socio-economic status and child development. *Annual Review of Psychology*, **53**, 371–399.
- BRISLIN, R. W., LONNER, W. J. and THORNDIKE, R. M., 1973, *Cross-cultural Research Methods* (New York: Wiley).
- CARTER, J. A., LEES, J. A., GONA, J., MURIRA, G., RIMBA, K., NEVILLE, B. G. R. and NEWTON, C. R. J. C., 2006, Severe *Falciparum* malaria: a common cause of acquired childhood language disorder. *Developmental Medicine and Child Neurology* (in press).

- CARTER, J. A., MUNG'ALA-ODERA, V., NEVILLE, B. G. R., MURIRA, G., MTURI, N., MUSUMBA, C. and NEWTON, C. R. J. C., 2005a, Persistent neuro-cognitive impairments associated with severe *Falciparum* malaria in Kenyan children. *Journal of Neurology Neurosurgery and Psychiatry*, **76**, 476–481.
- CARTER, J. A., MURIRA, G. M., ROSS, A., MUNG'ALA-ODERA, V. and NEWTON, C. R. J. C., 2003a, Speech and language sequelae of severe malaria in Kenyan children. *Brain Injury*, **17**, 217–224.
- CARTER, J. A., NEVILLE, B. G. R. and NEWTON, C. R. J. C., 2003b, Neuro-cognitive impairment following acquired central nervous system infections in childhood: a systematic review. *Brain Research Reviews*, **43**, 57–69.
- CARTER, J. A., ROSS, A. J., NEVILLE, B. G. R., OBIERO, E., KATANA, K., MUNG'ALA-ODERA, V., LEES, J. A. and NEWTON, C. R. J. C., 2005b, Developmental impairments following severe *Falciparum* malaria in children. *Tropical Medicine and International Health*, **10**, 3–10.
- CHAVEZ, A., MARTINEZ, C. and SOBERANES, B., 1995, Effects of early malnutrition on late mental and behavioural performance. *Developmental Brain Dysfunction*, **8**, 90–102.
- CINTAS, H. M., 1988, Cross-cultural variation in infant motor development. *Physical and Occupational Therapy in Pediatrics*, **8**, 1–20.
- COLE, M. and SCRIBNER, S., 1974, *Culture and Thought: A Psychological Introduction* (New York: Wiley).
- DAS, J. P., 1992, The relevance of psychology in a developing country: examples from cognitive assessment. *Psychology and Developing Societies*, **4**, 5–19.
- DASEN, P. R., 1988, Between the universal and the specific: the contribution of the cross-cultural approach. *Archives de Psychologie*, **56**, 265–269.
- DE LOACHE, J. S., STRAUSS, M. S. and MAYNARD, J., 1979, Picture perception in infancy. *Infant Behavior and Development*, **2**, 77–89.
- DEPARTMENT, FOR, EDUCATION, AND, SKILLS, 2003, *Aiming High: Raising the Achievement of Minority Ethnic Pupils*. Consultation Paper (London: DfES).
- DEREGOWSKI, J. B., 1968, Picture recognition in subjects from a relatively pictureless environment. *African Social Research*, **5**, 356–364.
- EVIATAR, Z., 2000, Culture and brain organisation. *Brain and Cognition*, **42**, 50–52.
- FREEDMAN, D. G. and DEBOER, M. M., 1979, Biological and cultural differences in early child development. *Annual Review of Anthropology*, **8**, 579–600.
- FLORES, G., 2000, Culture and the patient–physician relationship: achieving cultural competency in health care. *Journal of Pediatrics*, **136**, 14–23.
- GREENFIELD, P. M., 1997, You can't take it with you: why ability assessments don't cross cultures. *American Psychologist*, **52**, 1115–1124.
- GRIMES, B. F. (ed.), 2000, *Ethnologue: Languages of the World*, 14th edn, Vol.1 (Dallas, TX: SIL International).
- HELMS, J. E., 1992, Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, **47**, 1083–1101.
- HOWARD, D. P. and DE SALAZAR, M. N., 1984, Language and cultural differences in the administration of the Denver Developmental Screening Test. *Child Study Journal*. 1984 **14**, 1–9.
- HUDSON, W., 1960, Pictorial depth perception in subcultural groups in Africa. *Journal of Social Psychology*, **52**, 193–208.
- HUDSON, W., 1962, Cultural problems in pictorial perception. *South African Journal of Science*, **58**, 189–195.
- ISAAC, K. M., 2002, *Speech Pathology in Cultural and Linguistic Diversity* (London: Whurr).
- KAUFMAN, A. S. and KAUFMAN, N. L., 1983, *Kaufman Assessment Battery for Children* (Circle Pines, MN: American Guidance Services).
- KENNEDY, J. M., 1974, *A Psychology of Picture Perception* (New York: Josey-Bass).
- KEARINS, J. M., 1981, Visual spatial memory in Australian Aboriginal children of desert regions. *Cognitive Psychology*, **13**, 434–460.
- KORKMAN, M., KIRK, U. and KEMPT, S. L., 1998, *NEPSY — A Developmental Neuropsychological Assessment* (San Antonio, TX: Psychological Corporation).
- LIEVEN, E. V. M., 1994, Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway and B. J. Richards (eds), *Input and Interaction in Language Acquisition* (Cambridge: Cambridge University Press), pp. 56–73.
- MAHON, M., CRUTCHLEY, A. and QUINN, T., 2003, New directions in the assessment of bilingual children [Editorial]. *Child Language Teaching and Therapy*, **19**, 237–243.
- MARSHALL, J., 2003, International and cross-cultural issues: six key challenges for our professions. *Folia Phoniatrica et Logopaedica*, **55**, 329–336.

- MBISE, A. S. and KYSELA, G. M., Developing appropriate screening and assessment instruments: the case of Tanzania. In M. J. Thorburn and K. Marfo (eds), *Practical Approaches to Childhood Disability in Developing Countries: Insights from Experience and Research* (Tampa, FL: Global Age), .
- MILLER, N., 1984, *Bilingualism and Language Disability — Assessment and Remediation* (London: Chapman & Hall).
- MILLER, V., ONOTERA, R. T. and DEINARD, A. S., 1984, Denver Developmental Screening Test: cultural variations in Southeast Asian children. *Journal of Pediatrics*, **104**, 481–482.
- MILLER-JONES, D., 1989, Culture and testing. *American Psychologist*, **44**, 360–366.
- MISHRA, R. C., 1996, Cognition and cognitive development. In J. W. Berry, P. R. Dasen and T. S. Saraswathi (eds), *Basic Processes and Human Development*, 2nd edn, Vol. 2 (Boston: Allyn & Bacon).
- MORAIS, J. and KOLINSKY, R., 2000, Biology and culture in the literate mind. *Brain and Cognition*, **42**, 47–49.
- MULENGA, K., AHONEN, T. and ARO, M., 2001, Performance of Zambian children on the NEPSY: a pilot study. *Developmental Neuropsychology*, **20**, 375–384.
- MWAMWENDA, T. S. and MWAMWENDA, B. B., 1991, Africans' cognitive development and schooling. *International Journal of Educational Development*, **11**, 129–134.
- MWAMWENDA, T. S., 1992, Improving African children's cognitive development through training. *Perceptual and Motor Skills*, **74**, 1040–1042.
- NAGLIERI, J. A., 1985, *Matrix Analogies Test — Expanded Form* (Columbus, OH: Merrill).
- OCHS, E., 1983, Cultural dimensions of language acquisition. In E. Ochs and B. B. Schieffelin (eds), *Acquiring Conversational Competence* (London: Routledge & Kegan Paul).
- OLSON, D. R., 1976, Culture, technology and intellect. In L. B. Resnick (ed.), *The Nature of Intelligence* (Lawrence Erlbaum).
- PAHL, J. and KARA, M. B., 1992, The Renfrew Word Finding Scale: application to the South African context. *South African Journal of Communication Disorders*, **39**, 69–73.
- PAKENDORF, C. and ALANT, E., 1997, Culturally valid assessment tools: Northern Sotho translation of the Peabody Picture Vocabulary Test — Revised. *South African Journal of Communication Disorders*, **44**, 3–12.
- PAYNE, K. T. and TAYLOR, O. L., 2002, Multicultural influences on human communication. In G. H. Shames and N. B. Anderson (eds), *Human Communication Disorders. An Introduction*, 6th edn (Boston: Allyn & Bacon).
- PEÑA, E. and QUINN, R., 1997, Task familiarity: effects on the test performance of Puerto Rican and African American children. *Language Speech and Hearing Services in Schools*, **28**, 323–332.
- PEÑA, E., QUINN, R. and IGLESIAS, A., 1992, The application of dynamic methods to language assessment: a nonbiased procedure. *Journal of Special Education*, **26**, 269–280.
- PÉREZ-ARCE, P., 1999, The influence of culture on cognition. *Archives of Clinical Neuropsychology*, **14**, 581–592.
- REITAN, R. M. and WOLFSON, D., 1985, *The Halstead–Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation* (Tucson, AZ: Neuropsychology Press).
- REYNELL, J., 1980, *Language Development and Assessment*, Vol. 1 (Lancaster: MTP Press).
- ROSSELLI, M. and ARDILA, A., 2003, The impact of culture and education on non-verbal neuropsychological measurements: a critical review. *Brain and Cognition*, **52**, 326–333.
- SAENZ, T. I. and HUER, M. B., 2003, Testing strategies involving least biased language assessment of bilingual children. *Communication Disorders Quarterly*, **24**, 184–193.
- SAXE, G. B., 1988, The mathematics of street child vendors. *Child Development*, **59**, 1415.
- SBORDONE, R. J., 1996, Ecological validity: some critical issues for the neuropsychologist. In R. J. Sbordone and C. J. Long (eds), *Ecological Validity of Neuropsychological Testing* (Florida: G. R. Press/St Lucie Press).
- SCHIEFFELIN, B. B. and OCHS, E., 1986, *Language Socialisation Across Cultures* (Cambridge: Cambridge University Press).
- SCRIBNER, S. and COLE, M., 1981, *The psychology of literacy* (Cambridge, MA: Harvard University Press).
- SEMELA, J. J. M., 2001, Significance of cultural variables in assessment and therapy. *Folia Phoniatrica et Logopaedica*, **53**, 128–134.
- SERPPELL, R. and DEREKOWSKI, J. B., 1980, The skill of pictorial perception: an interpretation of cross-cultural evidence. *International Journal of Psychology*, **15**, 145–180.
- SERPPELL, R., 1972, Intelligence tests in the Third World. *Race Today*, **4**, 46–49.
- SEYMOUR, H. and MILLER-JONES, D., 1981, Language and cognitive assessment of black children. In N. Lass (ed.), *Speech and Language: Advances in Basic Research and Practice*, Vol. 4 (New York: Academic Press).

- SIGMAN, M., 1995, Nutrition and child development: more food for thought. *Current Directions in Psychological Science*, **4**, 52–55.
- SLOBIN, D., 1991, Learning to think for speaking: native language cognition and rhetorical style. *Pragmatics*, **1**, 7–26.
- SPERBER, A. D., DEVELLIS, R. F. and BOEHLECKE, B., 1994, Cross-cultural translation: methodology and validation. *Journal of Cross-Cultural Psychology*, **25**, 505–524.
- STANCZAK, D. E., STANCZAK, E. M. and AWADALLA, A. W., 2001, Development and initial validation of an Arabic version of the Expanded Trail Making Test: implications for cross-cultural assessment. *Archives of Clinical Neuropsychology*, **16**, 141–149.
- STERNBERG, R. J. and GRIGORENKO, E. L., Ability testing across cultures. In L. A. Suzuki, J. G. Ponterotto and P. J. Meller (eds), *Handbook of Multicultural Assessment* (San Francisco: Jossey-Bass).
- STOW, C. and DODD, B., 2003, Providing an equitable service to bilingual children in the UK: a review. *International Journal of Language and Communication Disorders*, **38**, 351–377.
- SUPER, C. M., 1976, Environmental effects on motor development: the case of “African infant precocity”. *Developmental Medicine and Child Neurology*, **18**, 561–567.
- TAYLOR, O. L. and PAYNE, K. T., 1983, Culturally valid testing: a proactive approach. *Topics in Language Disorders*, 8–20.
- VAN OTTERLOO, R. and VAN OTTERLOO, K., 1980, *A Sociolinguistic Study: The Bantu Language Groups of the Kenya Coastal Area* (Nairobi: Summer Institute of Linguistics).
- VYGOTSKY, L. S., 1962, *Thought and Language* (Cambridge, MA: MIT Press).
- WHORE, B. L., 1956, *Language, Thought and Reality* (Cambridge, MA: Technology Press of MIT).