Journal of Intercultural Communication, 17(2), 2017

Is there really a need for assessing intercultural competence?

Some ethical issues

Claudia Borghetti

University of Bologna

Abstract

Assessing intercultural competence (IC) is one of the hot-button issues within intercultural language education research today. Even if there are "more questions than answers" (Sercu, 2010) about this matter, it is generally accepted that assessing IC is somehow possible and necessary. This conceptual paper shifts from possibility to ethics, and discusses some reasons why assessing IC may not be considered opportune from an ethical perspective. Moreover, it casts doubt on whether assessing IC is truly necessary in language education. The paper presents four issues which appear problematic: The weaknesses of the existing models of IC with respect to assessment; the relationship between IC and interculturally competent performance; the context-based and relational nature of IC; the affective dimension of IC.

Keywords: *Intercultural Competence, Assessment, Ethics, Second Language Learning, Second Language Teaching*

Introduction

This theoretical and exploratory paper focuses on the assessment of intercultural competence (IC) in language education. It starts from the affirmation that, in the last two decades, increasing attention has been paid to assessing individuals' intercultural learning in language education. In Western language education, the phenomenon can be connected to the almost simultaneous publication of two language education policies which aim to guarantee good standards of quality in all aspects of language learning and teaching: the *National Standards for Foreign Language Education*, first published in the USA in 1996; and the *Common European Framework of Reference for Languages*, appearing in 2001.

The issue of assessing intercultural competence seems to fit well with a global society where increasingly something must be "proved", that is assessed and/or certified, to be real. This holds the consequence for individuals that there is a need to make even personal and private preferences visible and explicit, such as, in the case of IC, "interest in cultures", "empathy", "positive self-image", and so on. These matters as well as other ethical implications in assessing intercultural competence are addressed, following the idea that we may want to reconsider the importance attributed to the assessment of IC in second language classes, preferring other ways to foster the diffusion of intercultural language education and assert its importance.

Most of the reasoning used to sustain this position is not new in intercultural studies and language education, where it is often used to show how assessing IC is a difficult task and to identify ways to overcome such difficulties in order to make this assessment possible (e.g., Byram and Morgan, 1994;

Sercu, 2010). Here, however, the aim is to shift from possibility to ethics, while also addressing the issue of whether IC assessment is really necessary in intercultural language education.

1. Intercultural competence and its assessment

Although this paper focuses on IC in language education contexts, it cannot ignore the vast amount of studies dedicated to this concept in other fields, and especially in intercultural communication studies. After all, transdisciplinary influences are well attested. For example, in her attempt to reach a shared definition of 'intercultural competence', Deardorff (2006) conducted a significant interdisciplinary study where recognized experts in intercultural communication, education, political sciences, anthropology and psychology were asked to define what IC is, what its components are and the best ways to assess intercultural learning (i.e. the development of intercultural competence itself). Her study aimed to find a common way to look at this complex theoretical construct, since she could envisage at least forty-nine relevant studies, each with its own definition of IC and various labels (e.g., 'intercultural communicative competence', 'intercultural communication competence', 'cross-cultural competence', 'intercultural capability', and 'cultural sensitivity'). Many others, including Fantini (2009), Sercu (2004) and the frameworks analyzed by Spitzberg and Changnon, (2009), provide further definitions.

This potentially confusing theoretical terrain has one main consequence on this paper: It requires fixing the boundaries for the study, which, as anticipated, in this case means that this paper exclusively focuses on the assessment of intercultural competence in language educational contexts, even if to do this a wider range of literature is considered.

IC is conceived here as an integral whole of cognitive, affective and behavioural factors that influence the understanding of, and interaction with, diversity in a broad sense, and which can be developed through education and/or experience. This definition considers recent developments in IC studies in the fields of both language education and intercultural communication, where a relationship is increasingly envisaged between intercultural competence and communication on the one hand, and postmodern discourses about the notion of multiple, hybrid, fluid identities on the other (e.g., Guilherme, 2002; Noels, Yashima, and Zhang, 2012; Zarate, 2003). Intercultural competence is thus linked to how individuals socially position themselves in interactions (e.g., according to their nationally, genre, age, social status, etc.), to their awareness of such positioning, and to their willingness and ability to recognise and negotiate the others' multiple identities as much as their own. A clear advantage of this definition is that it is based on the tripartite distinction between 'cognition', 'affect', and 'behaviour', as traditionally adopted in social and cognitive psychology as well as in education (Bloom, 1956). Whatever the terms adopted to indicate each dimension (e.g., the affective dimension can be referred to as 'emotions', 'attitudes' or 'motivation' by different authors and within different models), this tripartite distinction has been the most productive in the development of IC models, both within intercultural communication (e.g., Chen and Starosta, 1996; Deardorff, 2006; Spitzberg 2000) and within language education (Byram, 1997, 2008; Fantini, 2000).

1.1 Reasons for assessing intercultural competence

The importance of assessing intercultural competence in intercultural language education can be advocated for a number of reasons and from several points of view. First of all, if IC development is one of the educational aims of language learning and teaching, it appears logical enough to assess language students' intercultural learning as well. After all, in education generally, "teaching involves assessment" (Rea-Dickins, 2004: 249), and teachers also have to monitor and improve the goals, methods and practices of their instruction in light of their learners' attainments. Paraphrasing Rea-Dickins, one might add that "learning involves assessment" and that learning benefits from assessment, in that assessment has the potential to increase students' self-awareness of their learning and learning strategies. In other words, the assessment of IC can be an "assessment for autonomy" (Lamb, 2010) in so far as it encourages learners to assume responsibility for their intercultural learning. Despite being at the heart of the learning and teaching processes, students and teachers are not the only social actors involved in formal education for whom assessing intercultural competence is important. On the political level of education systems, assessment is "part of the process of passing the values which are fundamental to the society from one generation to the next" (Byram, 2009: 222). Therefore, it is not surprising that current national and

international language education policies address the issue of assessing intercultural competence, also in light of the consideration given to intercultural dialogue and social cohesion in contemporary multicultural societies.

Whatever the perspective assumed (i.e., that of the teacher, the student or the policy maker), assessment can also be said to be important due to the 'washback' effect it may have on teaching and learning (Sercu, 2004). This can take on a strategic function in the case of intercultural language education, whose conceptual principles have not yet been fully considered by curricula planners and teachers (Sercu et al., 2005). Assessment also communicates to students that intercultural learning is important (Lessard-Clouston, 1992) and possibly affects their motivation towards intercultural learning. Importantly, Byram recognised early on that IC cannot be left "beyond the pale of respectability provided by assessment" (1988: 23).

At the same time, though, this 'washback' phenomenon has been largely discussed in language testing in relation to ethics, as one form of 'impact', that is to say, the overall effect that assessment has on test-takers (Hamp-Lyons, 1997) from an educational, socio-political or personal point of view. It has been argued that, for ethical reasons, test developers and teachers are responsible for how tests are used and for all their consequences on individuals (e.g., Hamp-Lyons, 1997; Shohamy, Donitsa-Schmidt and Ferman, 1996). Though this position is not unanimously agreed upon (e.g., Bachman, 2000; Davies, 1997), the proposal to reconceptualise washback and its social (and ethical) consequences as integral parts of validity in testing (Messick, 1996) has been intensely and productively debated.

1.2 How to assess intercultural competence

A consistent number of tools and standards for assessing intercultural competence have been developed. Considering both the educational and business fields, Fantini lists 44 principle assessment tools (2009). Additional annotated reviews are available in Humphrey (2007), SIETAR-Europa (2003), and Sinicrope, Norris and Watanabe (2007). Most of these tools are used to pursue summative assessment purposes, are based on self-awareness inventories and adopt psychometric tests (e.g., Arasaratnam, 2009; Hammer, Bennett and Wiseman, 2003). Generally they each present their own competence scales and descriptors. In contrast, the European INCA framework has had a wider application across assessment methods, at least in education. It consists of a set of six dimensions or scales ('tolerance for ambiguity', 'behavioural flexibility', 'communicative awareness', 'knowledge discovery', 'respect of otherness', 'empathy') each one articulated into three levels of competence: basic, intermediate and full (INCA, 2004a).

Versatile techniques for assessing IC are found in 'critical incidents', which Wight defines as 'brief descriptions of situations in which there is a misunderstanding, problem, or conflict arising from cultural differences between interacting parties or where there is a problem of cross-cultural adaptation' (1995: 128). More traditional uses of critical incidents, e.g. in Cushner and Brislin's 'cultural assimilator' (1996), require test takers to choose among a given list of possible interpretations for the described problematic situation, one of which is that preferred. However test takers are ever more commonly asked to give their own original interpretations, with the result that critical incidents are treated like essay examinations (Spencer-Oatey, 2013). Because of their flexibility, critical incidents are employed in various ways to assess intercultural competence: especially in cross-cultural training, they act as inputs for forms of multiple choice tests (e.g., Brislin, 1993); they may be commented on by test-takers, as in the case of the CEF-cult platform, where such comments are then evaluated by other tool users (Baten, Dusar and Van Maele, 2011); critical incidents can be written by assessees themselves, sometimes as part of their diaries for reflective self-assessment (Arthur, 2001; Jackson, 2005); finally, they can be integrated with other approaches in forms of multidimensional assessment, e.g., critical incidents are used together with a psychometric test and an autobiography in the SAILSA project (Feng and Fleming, 2009).

Multimethod and multiperspective assessment approaches like the SAILSA are increasingly encouraged (Deardorff, 2009). Assessment of intercultural competence can be conducted by means of different combinations of tasks and formats such as autobiographies, portfolios, projects, ethnographic observations, interviews, performance tasks and mini-dramas, critical incidents and questionnaires, etc. Different perspectives can be assumed by adopting a mix of expert, peer- and self-assessment. Overall, especially in language education, similar forms of alternative assessment (Fox, 2008; Gipps and Stobart, 2003) are considered the most suitable for assessing intercultural learning (Scarino, 2009; Liddicoat and

Scarino, 2013), especially when portfolio and autobiography are employed, some examples being the *Autobiography of Intercultural Encounters* (Byram et al., 2009), the INCA *Portfolio of Intercultural Competence* (INCA, 2004b), the *Culture Learning Portfolio* (Schulz, 2007), and the *PEER Model* (Holmes and O'Neill, 2010).

In spite of the number of IC assessment methods researched and employed, in the language educational field it is generally recognised that assessing intercultural competence is a difficult issue, for which there are "more questions than answers" (Sercu, 2004). However, the literature seems to illustrate a more or less explicit consensus on the fact that assessing IC in language classes is somehow possible and even necessary (e.g., Byram, 1997, 2009; Byram and Morgan, 1994; Fantini, 2009; Sercu, 2004, 2010; Schulz, 2007).

2. Assessing intercultural competence: a critical stance

In language assessment, 'ethics' is generally assumed - either overtly or implicitly - to be synonymous with 'morality', which pertains to "principles concerning the distinction between right and wrong or good and bad behaviour" (Stevenson, 2010: 1150). It has been advocated that what is right and good behaviour in assessment is related to

> *whether testing specialists should take any responsibility for decision about non-intended use of tests following test construction; who decides what is valid; whether professionalism conflicts with individual morality; relationship with various stakeholders; washback; and the politics of the gatekeeping use of language tests (in, for example, immigration procedures). (Davies et al. 1999: 55-56)*

In general terms, thus, the prevailing idea is that, in language assessment, ethics is linked to the responsibility of testers for the consequences of their tests on test-takers (Hamp-Lyons, 1997, 2000). Thus ethics is closely related to care and respect for individuals (Haynes, 1999). However, as stated by Lynch and Shaw, "to a certain extent, the issues of ethics in assessment can be thought of as aspects or expressions of power and the potential for abusing that power" (2005: 270). In this respect, testers' ethical conduct is framed within a broader idea of social, professional and personal responsibility (Hamp-Lyons, 2000), due to the impact that their assessment choices can have on test-takers' lives and thus to an unequal distribution of power between tester and test-taker. While this is the perspective mainly assumed in this paper, a slightly different order of considerations on ethics in assessment will be also employed, i.e. those linked to Shohamy's criticism of the instrumental use of tests to introduce changes in curriculum, textbooks and teacher training *de facto*, while the purposes of tests are not openly stated and agreed on through participative democratic decision-making processes (1997, 2007).

2.1 Models of intercultural competence and assessment

A first controversial ethical issue related to assessing IC in language education is linked to the large number of IC models available, which differently describe intercultural competence, sometimes for prescriptive and educational purposes. Stating explicitly which model is being used as reference in any given IC assessment may seem a reasonable solution to warrant that the inferences made about students' performances are valid. However, one can argue that this practice is not sufficient on the ethical level, since some crucial theoretical questions remain open about IC, which may undermine the legitimacy of assessors' decisions on other people's learning. Such conceptual questions are (see also Sercu, 2010):

- *What is culture?* It is worth noting that, with exceptions (e.g. Barrett et al., 2013), most existing IC frameworks do not provide any definition of 'culture'. This is problematic in assessment in terms of the conceptualisation and understanding of the descriptors. While it is true that some underlying assumptions about how culture is defined within a construct can be inferred by linguistic features used to described IC (e.g., whether a model employs an essentialist or non-essentialist view of culture, see Holliday, 2011), this lack of explicitness leads to cases of ambiguity when wordings

such as 'curiosity about different *cultural* practices' or 'willingness to respect *diversity*' are used in assessment descriptors.

- *What are the core IC components and what are the relationships among these components?* Existing models - in language education as well as in intercultural communication - conceptualise IC variously in terms of dimensions, components, teaching objectives, and, in some cases (Fantini, 2009), attributes or personal traits. Moreover, they do not state clearly if and how development in one component affects the others. Accordingly, while it is widely recognised that intercultural competence cannot be assessed holistically (Deardorff, 2009), what inferences - if any - can be drawn from the assessment of one IC component with respect to the others is yet to be established.
- *What is the link between intercultural competence and communicative competences in a second language?* As resulted from a comparative study conducted on a sample of models of IC (Borghetti, 2012), some frameworks conceptualise intercultural and communicative competences mostly as separate (e.g., Bennett, 1993; Deardorff, 2006), while others see the intercultural as a subcomponent of communicative competence (e.g., Kramsch, 1993, 1998) or on the contrary consider effective and appropriate communication in a second language as mostly independent from IC (Byram, 1997, 2008). Thus, it is unclear whether and how language proficiency is to be considered in IC assessment.
- *What 'levels' exist in intercultural competence development?* Apart from a few standalone attempts to establish levels of competence (e.g., Bennett, 1993; INCA, 2004a), there seems to be a general lack of confidence that "learning interculture can be viewed as a quantifiable step-by-step process from one level to the next" (Sercu, 2010: 28), thus to be assessed as such. As a matter of fact, there is a need for both conceptual investigation and empirical research on this.

Given the current models of IC - each addressing some theoretical questions pertinent to assessment and ignoring others - one is led to cast doubt on the validity of IC assessment in language education. In other words, how trustworthy can inferences drawn from students' performances be considered, when some of the features necessary for assessment are lacking in the constructs themselves, that is, in the relevant reference models of IC? In other words, "perennial validity questions" such as "are we looking at the right things in the right balance?" or "has anything important been left out?" (Messick, 1996: 246) are even more problematic if answers have to be sought within models which have internal shortcomings in respect of assessment, such as the ones listed above. Arguably, this is what Jæger sustains when she wonders whether the difficulties of assessing IC indicate a more fundamental problem regarding intercultural competence as a learning objective (1999).

If one assumes Messick's definition of validity (1989, 1996), this issue becomes ethical in principle. In Messick's comprehensive idea of validity, any matter related to construct validity is inseparable from the social consequences of assessment, including the assessment's power to influence students' lives. Thus, construct validity appears inseparable from ethics. How confident in the ethics of their professional behaviours can assessors be, when their inferences about people's IC learning - however rigorous and believable - are rooted in weak or incomplete theoretical bases? Moreover and crucially, as assessors, do we specify to the students that their level X, Y or Z of intercultural competence has been identified in respect to a specific benchmarking/model and thus its validity is somehow limited to the specific test they undertook? Is this clear to future employers and university admission offices, in case of high-stake IC tests (which, admittedly, are rare in educational contexts)? In other words, what may seem a disciplinary issue (the relation between the assessment practice and its construct) raises delicate social issues. And as will be discussed in §2.4, this matter appears to be even more sensitive if one considers that the object of assessment, intercultural competence, is a highly identity-related concept.

2.2 Intercultural competence vs interculturally competent performance

A second potentially controversial point for the assessment of IC is the relationship between intercultural competence and interculturally competent performance. Most existing models of IC ignore this issue and implicitly describe competence, while they also introduce features which clearly refer to competent performance. Thus, when adopted as reference frameworks for assessment, these models may, for example, lead to an evaluation of an individual's "interest" towards the others and "ability to establish and maintain relationships" (Fantini, 2009: 459), as if the process by which the two were elicited from the

assessees' performance were straightforward. This is the problem that Spitzberg and Changnon particularly identify in so-called 'compositional models', namely that:

> *[They] often mistake what constitutes an internal affective or cognitive factor, as opposed to a behavioral factor (i.e., skill). Thus, engaging in 'self-reflection' and 'taking multiple perspectives' are arguably internal information-processing activities and do not have obvious referents in the behavioral realm" (2009: 11).*

These problems related to the relationship between competence and performance for assessment purposes "are not confined, of course, to the evaluation of intercultural communicative competence, nor to language and culture teaching *per se*. They are common to all forms of assessment" (Byram and Morgan, 1994: 173). McNamara defines the relation between competence and performance in assessment as "obscure" referring to language testing (2001). At least partially the responsibility of the confusion between competence and competent performance can be ascribed to the fact that, while for conceptual thinking it is less problematic to maintain a certain degree of blurriness between the two, lack of theoretical clarity on this point is untenable when assessment - or pedagogy - are at the stake. Assessors and teachers deal with individuals' performances, and not with their competences, which are underlying and unattainable phenomena (Byram and Morgan, 1994). Jæger (1999) interestingly argues that the ambiguity between competence and performance is linked to the fact the construct of intercultural competence has been influenced more by Chomsky than by Hymes' idea of competence. For the former, competence is idealised and not fully reflected in performance, which is constrained by subjective and environmental factors; on the contrary, the latter denies this detachment, affirming that "appropriateness is *included* in the *definition* of communicative competence" itself, thus "meeting the standards of communicative acceptability in one's performance is demonstrating communicative competence" (1999: 77).

It would seem that the only model of IC which in principle tries to distinguish performance from competence is Deardorff's Pyramid Model of Intercultural Competence (2006) where "external outcomes" such as 'being able to carry out appropriate intervention in case of misunderstanding' are conceptually distinct from "internal outcomes" (for example 'willingness to seek out or take up opportunities to engage with otherness'). Deardorff's model clearly states what is blurry, though often present, in other constructs. For this reason, the labels 'internal' and 'external outcomes' are extremely useful to illustrate the present argument.

When assessors evaluate what can be referred to as the assessees' 'internal outcomes', they in fact (try to) infer competence from performance: They interpret the individuals' acts and personal accounts on the basis of assumptions on how 'recognition of the other's expectations', 'interest in other cultures', 'readiness to suspend judgement', etc. manifest themselves in students' performance. However, it can be argued that if eliciting competent behaviours is a serious difficulty, the attempt to assess internal outcomes poses even greater challenges and delicate ethical questions since it implies inferring what a student is thinking, willing to do, worrying about, etc. while performing any given intercultural assessment task. On the other hand, one cannot simply avoid assessing internal outcomes, if the intention presumably is to assess intercultural competence, i.e. in all its components. Moreover, exclusively assessing external outcomes can be a controversial solution as well, since in the case of intercultural competence, it is relatively easy for an individual to achieve external outcomes (such as 'apply effective and appropriate behaviours') without having fully achieved the internal outcome, thus without truly learning. And this would lead to the theoretical mistake of giving more credit to shallow than to deep learning, i.e. to changes which occur in surface behaviour, without affecting the individual's values and beliefs at a deeper level (Byram, 1997, 2008).

2.3 Intercultural competence as context-based and relational

A further challenging issue in the conceptualisation of intercultural competence pertains to how to establish what is an interculturally competent performance in a given assessment task or process, particularly when forms of alternative assessment are employed. Indeed, even when assessors decide to limit their evaluations to external communicative IC outcomes, the concepts of 'effectiveness' and 'appropriateness' are far from being unambiguous. To start with, Spitzberg points out that

"communication in an intercultural context […] is competent when it accomplishes the objectives of an actor in a manner that is appropriate to the context. Context here implies several levels including culture, relationship, place, and function" (2000: 375). Accordingly, if interculturally competent performance depends on several contextual factors, it can only be benchmarked against such a plurality of factors. Among the contextual variables, interlocutors in particular play a major role in eliciting what can be interpreted as interculturally competent or incompetent behaviours. For example, one cannot underestimate that competent performance greatly depends on the interlocutor's attitudes and (verbal) acts. What if, for example, the interlocutor shows a lack of interest in the interaction or even has bad intentions (Dervin, 2010)? Is students' lack of "analytic empathy" (Ting-Toomey and Kurogi, 1998) or interest in the other's perspectives (Byram 1997, 2008) necessarily negative in these cases? Moreover, "the success of any communicative activity is heavily determined by the way the participants perceive the context of situation and shape it accordingly through their verbal and non-verbal behavior" (Kramsch, 1993: 49-50). In other words, roles and expectations are co-constructed in interaction, and the interlocutors are the ones who can best establish what is appropriate and effective in the specific situation they are negotiating.

Similar challenges are well-known in alternative language assessment, especially when employing forms of performance assessments involving interactions between individuals. McNamara highlights them as follows:

> *If we are to take the point that everything is co-constructed in interaction, then it seems that we may only have performances, in Hymes's sense of 'instances of use', not performance in the sense of underlying potential for performance, ability for use. How are we to generalize from these actual instances? And how are we to speak of communicative competence as residing in the individual, if we are to 'include the hearer in the speaker's processes'? (1997: 457).*

As happens in the field of language assessment, in the case of interculturally competent performance when this is intended as an inherently social (contextual, relational, co-constructed) activity, a conceptualisation of the competence as belonging to the individual seems at the least debatable. With respect to assessment, this concept of IC creates a problem with generalizability, since a person cannot but manifest different levels of IC in different situations, depending on the contextual factors involved. At a broader theoretical level, this issue is linked to the very relationship between competence and performance discussed earlier (§2.2), as McNamara's words make clear. This is a fundamental theoretical challenge, which transcends methodological assessment choices (e.g., what test format is most suitable for assessing IC). As summarised by Lynch (who, once more, alludes to language assessment), the issue is in fact "epistemological and ontological", entailing a view of the construct "as something that is created and exists in the act of our using, inquiring and interpreting, not as an independent, objective entity waiting to be discovered and measured" (2001: 361). From this perspective, even if we had maps of possible alternative instances of interculturally-competent and interculturally-incompetent performances in order to benchmark and judge those of students (Scarino, 2009), this way to proceed in assessment would face the theoretical challenge that the (in)appropriateness of a performance from an IC point of view can only be established in the specific situation in which the particularly alternative was adopted. Thus, from an ethical perspective, who should take (or be given) the responsibility and *power* of establishing what is desirable and appropriate in an individual's specific performance? This question can partially be answered if assessment is made in interaction with others, where these others offer standardised behaviours specifically designed to maximise the learner/assessee's performance. However, apart from a single study (Harsch and Poehner, 2016), to my knowledge there are no empirical investigations on this delicate terrain.

2.4 The assessment of the IC affective dimension

A fourth issue related to IC which may be problematic from an ethical point of view concerns the assessment of its affective dimension, i.e. the individual's attitudes. Compared to the other three factors discussed above, this matter is the one that has been addressed in the literature most often for its questionable fairness with respect to teaching and, above all, assessment (Byram and Morgan, 1994; Jæger, 1999; Sercu, 2004). For example, Sercu states:

> *With respect to the assessment of the affective dimension of intercultural competence, the question arises as to whether it is desirable that learners be assessed with respect to particular attitudes or personality traits. Does education want to be prescriptive about the intercultural attitudes learners should develop and can learners be punished for not having particular desired personality traits, such as 'interest in cultures' or positive self-image', which are traits which have been identified as characteristic of the effective intercultural person? (2004: 78).*

In spite of their diverse combinations of components or dimensions, every model of IC usually encompasses an affective dimension, which pertains to the attitudes that the individual ought to develop towards diversity (e.g., openness, curiosity, interest in novelties, tolerance for ambiguity) and, in some frameworks, also a number of desirable personal traits or 'attributes', such as 'flexibility', 'humour', 'patience', 'empathy', and 'suspending judgements' (Fantini, 2009: 459). In other words, to be considered interculturally competent, an individual should possess such attitudes (and traits) to a certain degree, presumably in addition to, among others, being self-confident, centred in her/his opinions, and as free as possible from social anxiety.

A main ethical challenge in assessing the affective dimension of IC is of course linked to the fact that, to a certain degree at least, IC assessment implies evaluating private feelings and identity-related personal characteristics. This is especially true for personal traits or 'attributes', and it is probably not by chance that only some models of IC explicitly include them (Fantini, 2009; Spitzberg, 2000; Ting-Toomey, 1999; see also Kim's discussion about 'identity security', 2009). Others simply mention personal attributes as being part of the IC construct but not expand on them (Deardorff, 2006) and still other deliberately leave them out (Byram, 1997). However, the difference between personal traits and attitudes can be considered marginal in the present discussion, since crucial IC attitudes (including the most frequently-mentioned ones, 'curiosity' and 'openness') relate to personality at least to a certain extent. While it is unquestionable that some attitudes are catalysts for IC development, one wonders to what extent assessors should feel licensed to express evaluations or even simply their opinions when someone else's identity (and diversity) is at stake. For example, a test-taker may be shy or may not have a great level of self-esteem when interacting with people; would it be fair or even appropriate to assess that shy and anxious student negatively compared to other students or to a set of given standards?

The adoption of forms of alternative assessment like portfolios may well help reduce this ethical challenge in IC assessment, as they generally manage students' privacy more carefully. Similarly, a combination of observations of behaviour and of subjective judgment tasks, especially when discussed in class, can limit the impact of tests on students. More generally, formative assessment, where the student-teacher (institutional and personal) relationship helps frame the teacher's judgements about the student's personality, may be a valid alternative. The question appears more critical - especially with high-stake tests - with respect to the expected final attainment of intercultural development: Every construct sets standards, and guiding individuals towards standards which prescribe them *to be something* and *not to be something else* can be problematic. Not in the least because this implies that individuals are not encouraged to foster their own distinct characters, but are led in the direction of one pre-established, preferred personality. Once more, this matter is not new to language education: the *Common European Framework of Reference for Languages* warns that the development of an "intercultural personality" raises ethical and pedagogic issues, such as "the extent to which personality development can be an explicit educational objective" (2001: 106). However, comparatively, the question appears even more delicate in the case considered here, because of the impact that being assessed on their attitudes and personal traits can have on test-takers' self-image and self-esteem. Thus, in light of the concept of "consequential validity" (Messick, 1989, 1996), the assessor's responsibility assumes marked relevance in the case of IC.

3. A Shift from possibility to ethics: towards the idea of not assessing IC

The argument that assessing IC is important may appear indisputable, since intercultural education is (or should be) an integral part of language curricula. However, at least for what concerns language learning and teaching, this stance ignores the fact that in most schooling and higher education contexts, intercultural competence is not promoted at all. Therefore, either teachers are asked to assess what they did not teach - which is clearly not ethical - or assessment is used to introduce IC into the curricula, which raises a different order of ethical concerns.

From the macro-political perspective of education systems, assessment in general is an efficient (i.e., not time-consuming) way to monitor, improve and renew teaching, by inducing changes. However, this practice means that educational reforms are often introduced through changes in assessment, instead of being addressed directly through modifications in curricula and teacher training. In other words, assessment is used as a powerful device to make changes in the educational systems without consulting teachers, students, and experts in the field. This instrumental practice of using assessment for different purposes from which it is meant is non-democratic and thus could be categorised as "unethical" (Shohamy, 2007). This is even more so the case for a relatively recent educational area such as intercultural language education, whose theoretical framework and methodological principles will unlikely improve in their soundness, unless IC professionals together with students and teachers set the agenda for their promotion. In addition, it is worth remembering that, since it is harder to assess some IC components (e.g., attitudes and certain skills) than others (e.g., cultural knowledge), the impact of washback on curricula may be misleading and prevent intercultural pedagogy from focusing precisely on attitudes, etc., which are the most challenging - and presumably the most educational - of IC dimensions.

Finally, assessment is usually considered important because all the social actors involved at various levels in formal education (students, teachers and also parents) are interested in seeing evidence of learning. Given the specific, recognised challenges imposed by IC assessment, a way to meet this need in intercultural language education has been explored through focussing on low-stakes tests and adopting a general, substantial shift from assessment-of-learning to assessment-as-learning. As mentioned, formative or pedagogical assessment is receiving special attention, and portfolios, autobiographies and other forms of alternative assessment which variously integrate expert with self- and peer- assessment are being employed; this can indeed circumscribe the problem, making assessment a more participatory practice in which teachers and students' mutual trust also plays a role. However, in light of the ethical doubts raised above, we may also consider the idea of asking ourselves if assessing intercultural competence is truly necessary at all.

4. Conclusion: What do we do, then, with intercultural competence?

The paper has focused on four issues which appear problematic in the assessment of intercultural competence from conceptual and ethical perspectives:

- The weaknesses of the existing models of IC with respect to assessment;
- The relationship between IC and interculturally competent performance;
- The context-based and relational nature of IC;
- The affective dimension of IC.

In the light of these points, one may question the opportunity of assessing intercultural competence in intercultural language education. While it is advisable to specify the IC model which any given assessment process is based upon in order to better interpret (and circumscribe) test results, this practice may not suffice on an ethical level. The same can be said for the delicate issues of internal IC outcomes and contextual nature: Assessors should be asked to firmly rein in their assumptions about assessees' internal skills and behaviours, in order to avoid simplistic explanations and respect unpredictable human social processes. But perhaps this is too much to ask of assessors in terms of personal and professional responsibility, considering that IC conceptual thinking offers such weak support. On this respect, one also

needs to consider that, while a relatively strong emphasis is on the importance of assessing IC, teacher training in this area it is not as widespread as one would expect; this poses ethical challenges to the language teachers themselves, as they are asked to perform assessments that they are not qualified - or haven't been trained - to perform. Finally, if we assume (as I do) that in order to be able to say we are truly assessing IC, all its components should be addressed, then some doubts arise about the possibility to detach personality traits from the rest, and thus not assess test-takers' personalities.

However, if we avoid assessing IC altogether, a question arises: What do we do with intercultural competence in language education? I venture to say that several things can be done, some of which assume educational value precisely because assessment does not take place. First of all, efforts can be put on researching IC pedagogy (or IC with respect to pedagogy) in order to identify ways to promote intercultural learning. After all, as pointed out by Kramsch, "if tests can suggest to teachers what to teach, they don't tell the language teacher how to teach it. The procedure through which a competence is tested is not the process through which this competence is taught and acquired" (1998: 15). Moreover, while it remains true that leaving assessment out can entail both the risk that IC be considered less significant than other educational goals, and that students not be ready to be assessed in their intercultural competence, one may also argue for the *importance of not assessing IC*: leaving IC assessment out can convey a crucial, indirect message to students, namely that not everything must be proved and assessed to be valuable; and this may well represent a new, challenging educational aim in the present 'assessment era' (Broadfoot and Black, 2004). Finally, it is worth considering that, after all, interculturality is a value. Like other values, such as honesty, generosity, politeness, etc., it should be promoted in language classes (and elsewhere), without necessarily implying that, since it has been taught, it should automatically be assessed. Teachers can make the difference in the development of intercultural competence through their own behaviours, critical attitudes, motivations, expectations and so on, just as they do for other values, which are not marginal simply because they are not assessed.

References

Arasaratnam, L. A. (2009). The development of a new instrument of intercultural communication competence. *Journal of Intercultural Communication*, 20, 1-11.

Arthur, N. (2001). Using critical incidents to investigate cross-cultural transitions. *International Journal of Intercultural Relations*, 25, 41-53. doi:10.1016/S0147-1767(00)00041-9.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing,* 17(1), 1-42. doi:10.1177/026553220001700101.

Barrett, M., Byram, M., Lázár, I., Mompoint-Gaillard, P. and Philippou, S. (2013). *Developing intercultural competence through education*. Strasbourg: Council of Europe Publishing.

Baten, L., Dusar, L. and Van Maele, J. (2011). CEFcult toolkit. Intercultural communicative competence training materials. Retrieved from http://www.lscoop.com/files/cms/1cefcult_toolkit_students_2011-11fin.pdf

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (pp. 21-71). Yarmouth, ME: Intercultural Press.

Bloom, B. S. (1956). *Taxonomy of educational objectives*. New York, NY: David McKay.

Borghetti, C. (2012) Pursuing Intercultural and Communicative Goals in the Foreign Language Classroom: Clues from Selected Models of Intercultural Competence. In P. Alderete-Díez, L. Incalcaterra McLoughlin, L. Ní Dhonnchadha & D. Ní Uigín, (Eds.), *Translation, Technology and Autonomy in Language Teaching and Learning* (pp. 333-359). Frankfurt am Main: Peter Lang.

Brislin, R. (1993). A culture-general assimilator: preparation for various types of sojourns. In M. Paige (Ed.), *Education for the intercultural experience* (pp. 281-297). Yarmouth, ME: Intercultural Press.

Broadfoot, P. and Black, P. (2004). Redefining assessment? The first ten years of assessment in education, *Assessment in Education*, 11(1), 7-27. doi: 10.1080/0969594042000208976

Byram, M. (1988). Foreign language education and cultural studies. *Language, Culture, and Curriculum*, 1(1), 15-31. doi:10.1080/07908318809525025

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters.

Byram, M. (2008). *From foreign language education to education for intercultural citizenship. Essays and reflections*. Clevedon: Multilingual Matters.

Byram, M. (2009). Evaluation and/or assessment of intercultural competence. In A. Hu & M. Byram (Eds.), *Intercultural competence and foreign language learning: Models, empiricism, assessment* (pp. 215-34). Berlin: Gunter Narr Verlag.

Byram, M., Barrett, M. D., Ipgrave, J., Jackson, R. and Mendez-Garcia, M. C. (2009). *Autobiography of intercultural encounters: Context, concepts and theories*. Strasbourg: Language Policy Division, Council of Europe Publishing.

Byram, M. and Morgan, C. (1994). *Teaching-and-learning language-and-culture*, Clevedon: Multilingual Matters.

Chen, G-M. and Starosta, W. J. (1996). Intercultural communicative competence: A synthesis. *Communication Yearbook*, 19, 353-383.

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

Cushner, K. and Brislin, R. W., (1996). *Intercultural interactions: A practical guide. Second edition*. Thousand Oaks, CA: Sage.

Davies, A. (1997). Demands of being professional in language testing. *Language Testing*, 14(3), 328-339. doi:10.1177/026553229701400309.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. & McNamara, T. (1999). *Dictionary of language testing. Studies in language testing, Volume 7*. Cambridge: Cambridge University Press.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internalization. *Journal of Studies in International Education*, 10(3), 241-266. doi:10.1177/1028315306287002.

Deardorff, D. K. (2009). Implementing intercultural competence assessment. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 477-491). Thousand Oaks, CA: Sage.

Dervin, F. (2010). Assessing intercultural competence in language learning and teaching: A critical review of current efforts. In F. Dervin & E. Suomela-Salmi (Eds.), *New approaches to assessing language and (inter-)cultural competences in higher education* (pp. 157-174), Frankfurt: Peter Lang.

Fantini, A. E. (2000). A central concern: developing intercultural competence. *SIT Occasional Paper Series*. Brattleboro, VT: The School of International Training, 25-42.

Fantini, A. E. (2009). Assessing intercultural competence: Issues and tools. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 456-476). Thousand Oaks, CA: Sage.

Feng, A. and Fleming, M. (2009). Assessing intercultural competence for purpose – the SAILSA project. In A. Hu & M. Byram (Eds.), *Intercultural competence and foreign language learning: Models, empiricism, assessment* (pp. 235-252). Berlin: Gunter Narr Verlag.

Fox, J. (2008). Alternative assessment. In E. Shohamy & N. Hornberger (Eds.) *Encyclopaedia of language and education. Volume 7. Language testing and assessment* (pp. 97-108). New York, NY: Springer.

Gipps, C. V. and Stobart, G. (2003). Alternative assessment. In T. Kellaghan & D. L. Stufflebeam (Eds.), International handbook of educational evaluation (pp. 549-575). Dordrecht: Kluwer.

Guilherme, M. (2002). *Critical citizens for an intercultural world: Foreign language education as cultural politics*. Clevedon: Multilingual Matters.

Hammer, M. R., Bennett, M. J. and Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations*, 27, 421-443. doi:10.1016/S0147-1767(03)00032-4.

Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language Testing*, 14(3), 295-303. doi:10.1177/026553229701400306.

Hamp-Lyons, L. (2000). Social, professional and individual responsibility in language testing. *System*, 28, 579-591. doi:10.1016/S0346-251X(00)00039-7.

Haynes, F. (1999). Ethics and education. In M. Peters, P. Ghiraldelli, B. Źamič, & A. Gibbons (Eds.), *Encyclopedia of philosophy of education*. Retrieved from http://eepat.net/doku.php.

Harsch, C. and Poehner, M. E. (2016). Enhancing student experiences abroad: The potential of Dynamic Assessment to develop student interculturality, *Language and Intercultural Communication*, 16(3), 470-490. doi:10.1080/14708477.2016.116804

Holliday, A. (2011). *Intercultural communication and ideology*. London: Sage.

Holmes, P. and O'Neill, G. (2010). Autoethnography and self-reflection: Tools for self-assessing intercultural competence. In Y. Tsai & S. Houghton (Eds.), *Becoming intercultural: Inside and outside the classroom* (pp. 167-193). Newcastle upon Tyne: Cambridge Scholars Publishing.

Humphrey, D. (2007). *Intercultural communication competence: The state of knowledge*. London: The National Centre for Languages.

INCA (2004a). INCA: The theory. Retrieved from http://ec.europa.eu/ewsi/en/resources/detail.cfm?ID_ITEMS=9372.

INCA (2004b). INCA: Portfolio of Intercultural competence. Retrieved from http://ec.europa.eu/ewsi/en/resources/detail.cfm?ID_ITEMS=9372

Jackson, J. (2005). Assessing intercultural learning through introspective accounts. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 11, 165-186.

Jæger, K. (1999). Intercultural competence and the problem of assessment. In T. Vestergaard (Ed.), *Language, culture and identity* (pp. 71-81). Aalborg: Aalborg University Press.

Kim, Y. Y. (2009) The identity factor in intercultural competence. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 53-65). Thousand Oaks, CA: Sage.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.

Kramsch, C. (1998). Teaching along the cultural faultline'. In D. L. Lange, C. A. Klee, M. R. Paige, & Y. A. Yershova (Eds.), *Culture as the core: Interdisciplinary perspectives on culture teaching and learning in the second language curriculum. Selected conference proceedings* (pp. 15-31). Minneapolis, MN: University of Minnesota.

Lamb, T. (2010). Assessment of autonomy or assessment for autonomy? Evaluating learner autonomy for formative purposes. In A. Paran A. & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 98-119). Clevedon: Multilingual Matters.

Lessard-Clouston, M. (1992). Assessing culture learning: Issues and suggestions. *The Canadian Modern Language Review*, 48(2), 326-341.

Liddicoat, A. J. and Scarino A. (2013). *Intercultural language teaching and learning*. New York: Wiley and Sons.

Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351-372. doi:10.1177/026553220101800403.

Lynch, B. and Shaw, P. (2005). Portfolios, power, and ethics. *TESOL Quarterly*, 39(2), 263-297.

McNamara, T. (1997). 'Interaction' in second language performance assessment: Whose performance? *Applied Linguistics*, 18(4), 446-466. doi:10.1093/applin/18.4.446.

McNamara, T. (2001). Language assessment as social practice: challenges for research. *Language Testing*, 18(4), 333-349. doi:10.1177/026553220101800402.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement. Third Edition* (pp. 13-103). New York, NY: Macmillan.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. doi:10.1177/026553229601300302.

National Standards in Foreign Language Education Project (1996). *Standards for foreign language learning: Preparing for the 21st century (SFFLL)*. Lawrence, KS: Allen Press

Noels, K.A., Yashima, T. and Zhang, R. (2012). Language, identity and intercultural communication. In J. Jackson (Ed.). *The Routledge handbook of language and intercultural communication* (pp. 52-66). New York, NY: Routledge.

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Teaching*, 21(3), 249-258. doi:10.1191/0265532204lt283ed.

Scarino, A. (2009) Assessing intercultural capability in learning languages: Some issues and considerations, *Language Teaching*, 42(1), 67–80. doi:10.1017/S0261444808005417.

Schulz, R. A. (2007). The challenge of assessing cultural understanding in the context of foreign language instruction. *Foreign Language Annals*, 40(1), 9-26.

Sercu, L. (2004). Assessing intercultural competence: a framework for systematic test development in foreign language education and beyond. *Intercultural Education*, 15(1), 73-89. doi:10.1111/j.1944-9720.2007.tb02851.x.

Sercu, L. (2010). Assessing intercultural competence: More questions than answers. In A. Paran A. & L. Sercu (Eds.), *Testing the untestable in language education* (pp. 17-34). Clevedon: Multilingual Matters.

Sercu, L., Bandura E., Castro, P., Davcheva, L., Laskaridou, C., Lundgren, U., Méndez García, M. C. and Ryan, P. (2005). *Foreign language teachers and intercultural competence. An international investigation*. Clevedon: Multilingual Matters.

Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*, 14(3), 340-349. doi:10.1177/026553229701400310.

Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14(1), 117-130. doi:10.1080/09695940701272948.

Shohamy, E., Donitsa-Schmidt, S. and Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), pp. 298-317. doi:10.1177/026553229601300305.

SIETAR-Europa (2003). Intercultural assessment and instruments. Retrieved from http://www.sietar.de/SIETARproject/Assessments&instruments.html.

Sinicrope, C., Norris, J. and Watanabe, Y. (2007). Understanding and assessing intercultural competence: A summary of theory, research, and practice (technical report for the Foreign Language Program Evaluation Project). *Second Language Studies*, 26(1), 1-58.

Spencer-Oatey, H. (2013). Critical incidents. A compilation of quotations for the intercultural field. *GlobalPAD core concepts*. Retrieved from http://go.warwick.ac.uk/globalpadintercultural.

Spitzberg, B. H. (2000). A model of intercultural communication competence. In L. A. Samovar & Porter, R. E. (Eds.), *Intercultural communication: A reader. Ninth Edition* (pp. 375-387). Belmont, CA: Wadsworth.

Spitzberg, B. H. and Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 2-52). Thousand Oaks, CA: Sage.

Stevenson, A. (2010). *Oxford dictionary of English. Third Edition*. Oxford: Oxford University Press.

Ting-Toomey, S. (1999). *Communicating across cultures*. New York, NY: Guilford.

Ting-Toomey, S. and Kurogi, A. (1998). Facework competence in intercultural conflict: An updated face-negotiation theory. *International Journal of Intercultural Relations*, 22, 187–225.

Wight, A. R. (1995). The critical incident as a training tool. In S. M. Fowler (Ed.), *Intercultural sourcebook: Cross-cultural training methods. Volume 1* (pp. 127-140). Yarmouth, ME: Intercultural Press.

Zarate, G. (2003). Identities and plurilingualism: preconditions for the recognition of intercultural competences. In Byram, M. (Ed.), *Intercultural competence* (pp. 84-117). Strasbourg: Council of Europe.

About the Author

Claudia Borghetti is Research Fellow in Language Learning and Teaching at the Department of Modern Languages, Literature, and Cultures, University of Bologna, Italy. She was the project manager of the IEREST Project. She researches on intercultural language learning and teaching, language education, and teaching Italian as a foreign/second language.

Author's Address

Dr. Claudia Borghetti
Department of Modern Languages, Literature, and Cultures
University of Bologna
via Cartoleria 5, 40124
Bologna, Italy
Email: claudia.borghetti@unibo.it
tel. +39 0512097182

Journal of Intercultural Communication, ISSN 1404-1634, issue 44, July 2017.
URL: http://immi.se/intercultural