# Assessment of Interactional Competence

DAVID WEI DAI

## Interactional Competence as an Assessment Construct

We are seeing an incessant demand unfolding from both language learners and teachers to incorporate Interactional Competence (IC) into contemporary language education in order to develop and facilitate our ability to interact in the 21st century. The field of language assessment has shown a particular abiding interest in including IC in speaking assessment, as evidenced by dedicated edited books and special issues on this topic (Betz et al., 2023; Plough et al., 2018; Salaberry & Burch, 2021; Youn & Burch, 2020). The appeal of IC to the assessment community is twofold. First, established practices in the speaking sections in large-scale standardized language tests tend to privilege a psycholinguistic understanding of speaking, dissecting speaking into componential and discrete linguistic abilities such as grammatical range, pronunciation, and lexical resources. There is little emphasis on how such isolated abilities are mobilized for effective interpersonal communication, which is what speaking in the real world is about.

Second, test end-users—people who use language test results to make judgments about a person's language ability—are concerned about whether speaking test scores can predicate test-takers' competence in real-world, real-time interpersonal communication. It generates considerable consternation to language testing researchers and companies if a speaker with a high speaking test score on measures of grammar, vocabulary, and pronunciation somehow fails to demonstrate effective interpersonal communication skills. This mismatch between what current psycholinguistically oriented speaking tests measure and what test end-users expect can seriously threaten the credibility of the language assessment enterprise (Roever & Dai, 2021). IC, with its promise of assessing the ability to interact, is a suitable candidate to address these pressing issues.

When it comes to assessing IC, the first and foremost question assessment researchers ask is what IC really means, or in testing parlance, how IC as a test construct is defined. Before I offer a more precise definition of the IC construct, I would argue that in broader terms, IC is a speaker's ability to interact with others in a range of real-world social, cultural, and interpersonal contexts. What differentiates IC from related constructs such as communicative competence and pragmatic competence is that IC embodies a constructivist epistemology to social interaction. Constructivist in the context of IC implies that the discourse where interaction takes place is concurrently *constructed* and *constructive*: it is *constructed* in that speakers use grammar, vocabulary, and pronunciation to talk identities, social relations, and community memberships into existence. Discourse is contemporaneously *constructive* because the identities, social relations, and community memberships speakers have evoked are simultaneously used as resources for further elaboration of speakers' lifeworlds where interaction unfolds.

In view of IC's epistemology, a more technical definition of IC through the constructivist lens is a speaker's ability to efficiently marshal interactional resources (e.g., grammar, vocabulary, and pronunciation) to manage the temporal progression of interaction, the enactment *of* and orientation *to* social roles and community memberships, and the moment-by-moment engagement with interlocutors to achieve context-specific interactional outcomes. The epistemological underpinning of IC maps onto the methodologies IC researchers have chosen to study social interaction. Conversation Analysis (CA) and Membership Categorization Analysis (MCA)—with the former focusing

*The Encyclopedia of Applied Linguistics*. Edited by Carol A. Chapelle.

DOI: 10.1002/9781405198431.wbeal20786

on how interaction unfolds sequentially and the latter explicating how social roles and identities are talked into existence categorically—have proved to be fitting methodologies for the study of IC (Dai & Davey, 2023).

## Interactional Competence Rubric Development

The epistemological premise and methodological praxis of IC research have direct implications for how IC constructs are operationalized. This in turn reflexively determines how IC assessment rubrics are developed as rubrics are embodiments of test constructs. IC rubric development in existing studies can be categorized into three approaches: (a) the discourse approach, (b) the synthesis approach, and (c) the universal approach.

The discourse approach displays a strong orientation to the conversation-analytic tradition of CA and MCA by developing the IC rubrics bottom-up (Ikeda, 2021; Youn, 2015). Researchers first build a collection of test-taker sample performances, followed by discourse analysis on these performances, mostly via CA so far but with growing incorporation of MCA (Dai & Davey, 2023; Dai, 2024), and finally identify IC features based on which features stand out to the researchers as distinctive and differentiable of test-taker ability levels. The advantage of the discourse approach is that by staying true to the constructivist methodological tradition of CA and MCA, the resulting IC rubrics demonstrate a close fit to the IC tasks and test-taker profiles on hand. Because rubrics are derived directly from test-takers' actual speaking performances, the descriptors in the rubrics offer a close resemblance to what is taking place in test-taker discourse. This can facilitate the training of IC raters and the rating of IC performances as raters will be able to identify quickly in test-taker performances construct-relevant features, as captured by the rubric descriptors. The limitation of the discourse approach is that it is a laborious process since test developers need to conduct fine-grained discourse analysis on a sizable collection of test-taker performances to develop rubrics that are comprehensive enough to cover crucial IC features in their test tasks and target test-taker demographics. The selection of IC features to incorporate in the rubrics is also driven by the researcher's sole judgment, which can leave the rubrics vulnerable to concerns about whether other test stakeholder groups (e.g., language teachers and language users in the real world) agree with the researcher in terms of construct definition.

The synthesis approach draws on existing IC research and selects IC features that the researchers consider relevant to their test tasks. A typical example using this approach is Ockey et al. (2023) where the researchers canvassed previous IC studies and selected 15 IC features (e.g., "checking comprehension" and "connecting topics") for their rubric. The advantage of the synthesis approach is that test developers can utilize IC features previous research has already shown to be capable of separating test-takers into different IC levels. It improves the practicality of IC assessment by circumventing the resource-intensive nature of conducting discourse analysis on test-taker performances to inductively arrive at an IC rubric. However, the limitation of the synthesis approach lies precisely in the absence of discourse analysis. The epistemology of IC hinges on a constructivist understanding of social interaction: an IC construct without an intimate connection to discourse leaves open the question of to what extent the IC features the researchers selected based on previous studies actually reflect the features observable in their particular test tasks. As IC is *shaped by* and *shapes* its immediate interactional context, it is incumbent on researchers adopting the synthesis approach to demonstrate how the IC features they have selected can offer a good fit for the interactional context pursuant to their test tasks.

The third approach, termed the universal approach, commences from a holistic, theoretical IC construct that lends itself to adaptation to local interactional contexts. The thinking behind this approach is motivated by the resource-intensive nature of the discourse approach and the reliance on test developer judgment in both the discourse and the synthesis approach. Dai (2024) represents an initial attempt at the universal approach. Starting from everyday-life linguistic laypersons'

criteria of effective interaction and locating these criteria using CA and MCA in test-taker discourse, Dai proposes a universal IC construct/rubric that measures IC at the sequential, emotional, logical, moral, and categorial dimensions. The descriptors of interactional conduct in Dai's rubric are deliberately theoretical—drawing on concepts in CA and MCA—to depict the universal, context-free mechanism of interaction. As this theoretical IC construct/rubric represents interactional processes that cut across any language, assessment task, and test-taker cohort, it allows test developers to localize it to their specific target languages, speaking tasks, and test-taker demographics. The limitation of the universal approach is that it still requires a stage of adaptation—in the form of discourse analysis—to translate its theoretical rubric descriptors to local test-taker behavior for rater training and rating. However, the amount of discourse analysis in the universal approach is smaller than the one in the discourse approach as its focus is on adaptation and localization of an existing rubric, compared to building one from scratch in the discourse approach. In this sense, the universal approach amounts to a middle-ground one between the discourse and synthesis approaches.

## Interactional Competence Task Development

Apart from rubric development, another central consideration in IC assessment is to develop language tasks where IC can be observed and assessed. There are two pivotal questions to IC task development: (a) what types of speaking tasks are suitable for the assessment of test-takers' interactional abilities, and (b) what are the principles behind IC task development.

In terms of task suitability, since IC is concerned with the ability to interact in real-world sociocultural contexts, IC *tasks* by this definition are incompatible with speaking assessment *activities* such as "reading aloud" or "repeating sentences," which, modeled on psycholinguistic constructs of speaking, assess explicitly and solely abilities such as one's control over speech fluency and pronunciation. Here, I purposefully distinguish between activities and tasks as the latter are undertakings that correspond to real-world, authentic communicative events where IC can be observed. Speaking assessment activities, for example, repeating sentences, are on the very individualist end of an interactiveness continuum, assuming that test-takers are performing speaking with no specific interlocutor or sociocultural context present. At the other end of this continuum, we have highly interactive speaking tasks such as pair/group discussions and roleplays that lend themselves well to the observation of IC. Discussion and roleplay tasks are usually modeled on real-world, two-way/multiparty interaction with rich social, cultural, and contextual cues, allowing test-takers to display their ability to interact with others on a moment-by-moment basis. These interactive tasks can generate test-taker performance for assessors to observe IC on various dimensions:

- **Sequential**: how test-takers structure interaction temporarily (e.g., turn-taking management and topic development).
- **Categorial**: how test-takers orient to the social, cultural, and relational aspects of interaction (e.g., making oneself sound like a *student* and designing their interaction to suit the *lecturer* identity their interlocutor takes on in academic discourse).
- **Emotional**: how test-takers promote rapport with their interlocutor (e.g., using language to build a therapeutic relationship with patients in clinical communication assessment as in the Occupational English Test, see OET, 2018).
- **Logical**: how test-takers display reasoning skills in interaction (e.g., the reasons and evidence they provide for their arguments in discussion tasks or debates).
- **Moral**: how test-takers manage the self-display of their personae in interaction (e.g., do they sound like a *diligent* student, a *responsible* doctor, or an *approachable* accountant).

An interesting question about IC task development concerns the tasks that are somewhat in the middle of the interactiveness continuum: they are the monologic tasks such as "producing a verbal

summary after listening to a short university lecture" or "sending a voice message after receiving one from one's neighbour." These monologic tasks offer more stable estimates than interactive, real-time tasks that involve live interlocutors, who can introduce variability in the measurement of test-takers' IC. Monologic tasks, provided that they are embedded in realistic sociocultural contexts, can also elicit a range of IC features such as topic development, social role orientation, and emotional engagement. The shortcoming of monologic tasks is that they offer only partial coverage of the IC construct, unable to account for more interactive, co-constructed IC features such as turn-taking and back-channeling. It is therefore appropriate to combine monologic with dialogic tasks in IC test batteries (e.g., see Dai, 2024; Ikeda, 2021; Youn, 2015).

As to the second question of IC task development, the principles behind IC task development, so far two approaches have emerged: the IC-additive approach and the IC-centered approach. The IC-additive approach sees IC as an additional dimension of the speaking construct in juxtaposition to established speaking ability indicators such as pronunciation, grammatical range, and lexical resources (see UCLES, 2008 for an example of this approach used in Cambridge English). Extending previous psycholinguistic models of speaking to accommodate an IC dimension, this approach requires lesser adaptation of existing speaking assessment instruments. Test developers only need to explore what IC features their validated speaking tasks, whether they are interview-based or discussion-based, can elicit, and then add a new rating category on IC to incorporate IC in their overall psycholinguistically oriented speaking constructs.

A more radical and potentially paradigm-shifting approach to IC task development is the IC-centered approach. This approach puts the ability to interact at the heart of effective speaking, requiring a fundamental rethink of what speaking assessment is trying to assess. Adopting this approach aligns IC assessment, and speaking assessment in general, more with assessing communicative competence and pragmatic competence. Assessing IC essentially becomes assessing one's ability to manage *functional language use*, which Kramsch (1986) describes as ranging from "survival as a tourist or a student to negotiating treaties" (p. 366). Dai (2023) proposes that social action—such as requesting, refusing, and complaining—can serve as the unit of assessment for IC. The IC-centered approach explicitly tests speakers' ability to manage specific social actions, embedded in specific speech events and macro speech situations (Galaczi & Taylor, 2018). When asking test-takers to demonstrate the competence to handle a task centered on, for example, negotiating a meeting time with classmates for a group project or complaining to their manager about unfair work practices, test developers can assess test-takers' ability in deploying a range of IC resources for sequential organization, social role orientation, emotional engagement, logical reasoning, and self-display management. As social actions are language-, context- and culture-specific, the development of IC-centered assessment tasks requires test developers to conduct needs analysis to understand the specific social actions that are of pertinence to their respective test-taker cohorts (see Dai, 2023 and Youn, 2018 for examples). Although needs analysis calls for time and resources, it provides an empirically grounded approach to developing language tasks that are more meaningful, relevant, and useful for test-takers as the resultant tasks target the genuine interactional needs test-takers experience in their respective real-world language use domains.

## Future Research Directions

Having discussed what IC as an assessment construct is, how we can develop rubrics to assess IC, and how we can create speaking tasks to observe IC, this entry concludes with three future research directions for IC assessment.

First, there is continuing work to build and validate holistic construct definitions of IC that go beyond our current focus on sequence in IC assessment (Galaczi & Taylor, 2018; Ikeda, 2021; Youn, 2015). Dai (2024) represents one such attempt by assessing IC using a holistic IC construct

covering the sequential, emotional, logical, moral, and categorial dimensions of social interaction. Using Many-facet Rasch measurement (MFRM), Dai observed high MFRM test reliability, high MFRM item reliability, and high intra- and inter-rater reliabilities. These findings are a promising development in IC assessment as they demonstrate the achievability of assessing volitive and attitudinal factors (cf. Hymes, 1972) of interpersonal communication, which assessment researchers have tended to consider too messy and unreliable for large-scale, standardized assessment (Elder et al., 2017). Further research along this line can shed more light on the feasibility of assessing holistic IC constructs, bringing us closer to a more comprehensive assessment construct of effective interpersonal communication (see how the original definition of communicative competence in Hymes, 1972 includes a range of non-cognitive, volitive, and attitudinal factors).

Second, there is growing evidence that IC is distinct from the psycholinguistic definition of speaking proficiency, raising questions about whether L1 speakers should also be tested for IC. Comparing L2 test-taskers' performance on an IC test and a TOEFL speaking test, Roever and Ikeda (2023) noticed that there are certain IC features that are more susceptible to test-takers' proficiency levels as measured by TOEFL, whereas other IC features are less impacted by, and in some cases independent from proficiency. By recruiting both L2 and L1 speakers for his L2-Chinese IC test, Dai (2024) noted a weak correlation between IC and proficiency, with many L1 speakers performing worse on IC measures compared to L2 speakers of lower proficiency. This disconnect between IC and proficiency suggests that IC is a unique ability that requires dedicated instruction and attention in language learning, teaching, and assessment. It also poses an interesting question as to whether there is a need to teach and assess L1 speakers' IC since they can no longer be safely assumed to be the gold standard of effective interpersonal communication, insofar as IC is taken into account. More research is therefore required to compare L2 and L1 speakers' IC performances to ascertain if IC is an ability that everyone needs to develop for effective interpersonal communication.

Finally, advances in technology, most notably in generative artificial intelligence (AI), promise significant potential for the assessment of IC. One of the main challenges to rolling out IC assessment on a large-scale basis is its resource-intensive nature, which, at the far end of the interactiveness continuum described above, requires live human interlocutors. The incorporation of monologic IC tasks can mitigate this burden but only to a limited extent. Generative AI, on the other hand, can drastically reduce human labor costs in testing IC, provided that AI is developed to the extent that it can elicit meaningful IC features comparable to human interlocutors. We are already witnessing technology making strides in this direction: LANGX Speaking (Equmenopolis, 2022) is an AI-generative speaking test that allows test-takers to have conversations in an interactive manner with AI interlocutors. This has significant implications for teaching and testing IC, especially for test-takers from lower socioeconomic backgrounds who otherwise cannot afford IC instruction and assessment if IC relies solely on human interlocutors. Additionally, AI can offer more adaptability and authenticity for assessment as it requires less effort to configure an AI-interlocutor of a particular language and demographic background than expecting a human interlocutor to consistently and credibly demonstrate these features. More research is required in these areas to evaluate how AI can contribute to the assessment of IC.

**SEE ALSO:** Rating Scales and Rubrics in Language Assessment, Assessment of Speaking, Needs Analysis, Discourse Analysis in Language Assessment, Conversation Analysis and Membership Categories, Pragmatics in the Analysis of Discourse and Interaction

## References

Betz, E., Malabarba, T., & Barth-Weingarten, D. (Eds.). (2023). Describing and assessing interactional competence in a second language. [Special issue]. *Applied Pragmatics*, *5*(2). https://doi.org/10.1075/ap.00020.mal

Dai, D. W. (2023). What do second language speakers really need for real-world interaction? A needs analysis of L2 Chinese interactional competence. *Language Teaching Research*, *1–38*. https://doi.org/10.1177/13621688221144836

Dai, D. W. (2024). *Assessing Interactional Competence: Principles, test development and validation through an L2 Chinese IC test*. Peter Lang. http://doi.org/10.3726/b21295

Dai, D. W., & Davey, M. (2023). On the promise of using Membership Categorization Analysis to investigate Interactional Competence. *Applied Linguistics*, *1–26*. https://doi.org/10.1093/applin/amad049

Elder, C., McNamara, T., Kim, H., Pill, J., & Sato, T. (2017). Interrogating the construct of communicative competence in language assessment contexts: What the non-language specialist can tell us. *Language & Communication*, *57*, 14–21. https://doi.org/10.1016/j.langcom.2016.12.005

Equmenopolis (2022). *LANGX speaking*. https://www.equ.ai/en/langx/

Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualizations, operationalizations, and outstanding questions. *Language Assessment Quarterly*, *15*(3), 219–236. https://doi.org/10.1080/15434303.2018.1453816

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.

Ikeda, N. (2021). Assessing L2 learners' pragmatic ability in problem-solving situations at English-medium university. *Applied Pragmatics*, *3*, 51–83. https://doi.org/10.1075/ap.19039.ike

Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal*, *70*, 366–372. https://doi.org/10.1111/j.1540-4781.1986.tb05291.x

OET (2018). *Speaking assessment criteria and level descriptors*. https://prod-wp-content.occupationalenglishtest.org/resources/uploads/2018/08/22102547/speaking-assessment-criteria-updated-2018.pdf

Ockey, G. J., Chukharev-Hudilainen, E., & Hirch, R. R. (2023). Assessing Interactional Competence: ICE versus a Human Partner. *Language Assessment Quarterly*, *1–22*. https://doi.org/10.1080/15434303.2023.2237486

Plough, I., Banerjee, J., & Iwashita, N. (2018). Special issue on interactional competence [Special issue]. *Language testing*, *35*(3).

Roever, C., & Dai, D. W. (2021). Reconceptualising interactional competence for language testing. In M. R. Salaberry & A. R. Burch (Eds.), *Assessing speaking in context: Expanding the construct and its applications* (pp. 23–49). Multilingual Matters. https://doi.org/10.21832/9781788923828-003

Roever, C., & Ikeda, N. (2023). The relationship between L2 interactional competence and proficiency. *Applied Linguistics*, *1–23*. https://doi.org/10.1093/applin/amad053

Salaberry, M. R., & Burch, A. R. (Eds.). (2021). *Assessing speaking in context: Expanding the construct and its applications*. Multilingual Matters.

UCLES (2008). *Assessing speaking performance – Level B2*. https://www.cambridgeenglish.org/images/168619-assessing-speaking-performance-at-level-b2.pdf

Youn, S. J., & Burch, A. R. (Eds.). (2020). Where Conversation Analysis meets language assessment. [Special issue]. *Papers in Language Testing and Assessment*, *9*(1).

Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, *32*(2), 199–225. 10.1177%2F0265532214557113

Youn, S. J. (2018). Task-based needs analysis of L2 pragmatics in an EAP context. *Journal of English for Academic Purposes*, *36*, 86–98. https://doi.org/10.1016/j.jeap.2018.10.005

## Suggested Readings

Hall, J. K., Hellermann, J., & Doehler, S. P. (Eds.). (2011). *L2 interactional competence and development*. Multilingual Matters.

Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, J. C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics & Language Learning* (vol. *11*, pp. 281–314). University of Hawai'i at Manoa: National Foreign Language Resource Center.

Roever, C. (2021). *Teaching and testing second language pragmatics and interaction: A practical guide*. Routledge.

Stokoe, E. (2012). Moving forward with membership categorization analysis: Methods for systematic analysis. *Discourse Studies*, *14*(3), 277–303. https://doi.org/10.1177/1461445612441534

ten Have, P. (2007). *Doing conversation analysis*. Sage.