

## **Integrating Instruction, Curricula, and Assessments in the L2 Classroom**

August 2015 - Volume 19, Number 2

**Kent Hill**

Meiji University, Japan

<kentokun@hotmail.com>

### **Abstract**

Simultaneous with a call for standardized assessments to assess developing rather than preexisting or surface knowledge (Sawyer, 2006) has been a growth in awareness of the unavoidable interactive nature of any form of assessment and this awareness has created a social turn towards approaches to language assessment (McNamara & Roever, 2006). Following on previous research investigating the link between instruction and dynamic and curriculum-based assessment (Lidz, 1991), this study attempts to go a step further and examines integrating instruction, the curriculum, and dynamic and standardized criterion-referenced pre/post-assessments. Results show that through integration learners performed significantly better on both forms of assessment. Learners also expressed positive washback and a higher task-intrinsic motivation to participate in more integrated forms of assessment. These preliminary results suggest that dynamic and standardized forms of assessments may be commensurable, and integrating them might also be effective in unlocking some of the discretely organized knowledge found in standardized assessment, which provides support for further direction to the social turn in assessment.

**Keywords:** assessment, dynamic, standardized, curriculum-based, development, mediated assistance, washback

### **Introduction**

Many forms of standardized assessments, although claiming to assess learners' communicative competence (Bachman, 1990; Canale & Swain, 1980), typically assess only certain components of language knowledge (Widdowson, 2001), and each component is quite often discrete, stemming from a psychometric approach to assessment and a belief that human abilities exist as discrete variables (Poehner, 2008). Standardized assessments often seek to avoid mediation in the assessment process since it could potentially contaminate the implications that could be drawn from learner performances and this in turn arises from the deeper ontological assumption that the source of human psychology is the individual (i.e., cognitive) and not the sociocultural environment.

Simultaneous with a call for standardized assessments to assess developing rather than preexisting or surface knowledge (see Sawyer, 2006) has been a growth in awareness of the unavoidable interactive nature of any form of assessment, especially with oral forms of assessment (e.g., between assessor and assessee), and this awareness has created a social

turn towards approaches to language assessment (McNamara & Roever, 2006). This turn has also created a reciprocal growing interest in classroom-based forms of assessment (Davison & Leung, 2009; Leung & Lewkowicz, 2006), which has benefitted from the discussion of sociocultural theory (SCT) and assessment research (Lantolf, 2000, Lantolf & Thorne, 2006). Classroom-based assessment, as opposed to standardized, presents a clear description of actual learner interaction (i.e., as non-discrete) and skills development (i.e., processes). In particular, dynamic assessment (Poehner, 2007, 2008) offers a specific SCT application to classroom-based assessment.

Dynamic assessment does not view variables as discrete and it does not separate instruction from assessment. Dynamic assessment and SCT are both based on the Vygotskian (1978, 1986) theory of the mind (i.e., language learning and development are the result of inter- and intra-cognitive internalization processes). Contrary to standardized assessment, which is often considered to be incommensurate with dynamic assessment (Poehner, 2008), through ongoing collaborative engagement between an assessor and learner, dynamic assessment focuses on assessing the potential learners have for language development and their ability to overcome any problem-solving difficulties they may have with language usage. It is more important in this sense to assess future development and this is determined by how learners react to mediation. At a deeper level, this assumption arises from the fundamental ontological SCT claim that human psychology is cultural and not biological, and as such is external to the individual and therefore finds its origins in social interaction.

In this approach to integrating assessment and instruction, assessment procedures are not developed *a priori* and then imposed on classroom teachers but emerge from a grounded classroom-based analysis of instructional interactions and pedagogical practices. This approach is referred to as *curriculum-based assessment* (Poehner, 2008). Combining curriculum-based and process-based approaches within a dynamic assessment model has been proposed by Lidz (1991) as the optimal approach to assessment. Addressing cognitive processes that underlie specific curriculum objectives allows the diagnostician to probe more deeply into the nature of learners' learning development and areas of challenge. Lidz adds (1991, p. 408) that remaining close to the curriculum increases the relevance of the resulting recommendations for the instructional setting and development of this approach represents an attempt to maximize the link between assessment and instruction.

Curriculum-based assessment (CBA) has increased in popularity because it links assessment to the curriculum or a good representative sampling of it (Haywood & Lidz, 2007). CBA does not in itself provide information about difficulties to learners in mastery of the curriculum. According to Haywood & Lidz (2007), the addition of dynamic assessment to CBA provides the missing link. Curriculum-based dynamic assessment (CBDA) begins with the selection of relevant curriculum content and construction of an appropriate CBA to serve as both pre- and post-assessment. The dynamic component enters after this step by applying a pre-assessment, intervention, post-assessment format that includes a dynamic intervention which addresses the processing demands of the selected tasks in relation to the processing capabilities of the learner. Interventions for attention can also address the perceptual properties of tasks in ways that capture the learner's attention, for example, introducing ways of highlighting what is important to

notice, use of contrast and novelty, and manipulation of gestures and voice (Haywood & Lidz, 2007).

Along the lines of this CBDA model of establishing a link between classroom and curriculum-based assessment and instruction, this present study goes a step further and investigates to what degree integrating dynamic approaches using the analogous content of a small-scale standardized form of assessment may be effective and commensurate as well as useful in overcoming some of the discrete or individual nature of standardized assessment. Following on a previous study into dynamic assessment (Hill & Sabet, 2009) which reported on dynamic approaches to speaking assessments, the study reported here involves integrating intervening dynamic speaking assessments (DSAs) using analogous content from a university-wide first-year standardized criterion-referenced pre-/post-assessment (pre-/post-CRA) to determine whether learners performed significantly better on the post-CRA. CRAs assess the amount of curriculum-based knowledge the learner has acquired during the period of study. This assessment was considered standardized because all first-year students streamed into the mid-level at this university were required to take it in order to complete the required course. The pre-/post-CRA involved conversation-based writing and listening assessments.

According to Barab (2006), variability in dynamic assessment design studies is not controlled or completely known in advance. In contrast to standardized assessment, it is seen as part of the design for analysis and some variables may emerge during the study. Along with basing the DSAs on analogous content of the pre-/post-CRA, this study also added a *development-referenced assessment* (DRA). After determining the common errors that emerged in the initial DSA and having learners attempt to correct them in the second DSA, a common-error written DRA was added to the post-CRA.

As a form of triangulation, a follow-up questionnaire was also administered to learners. The use of the questionnaire intended to determine whether learners experienced any positive or negative washback effect, in this case, as a result of integrating the pre-/post-CRA with intervening DSAs. If integration leads to teaching practices that promote and broaden learning, the effect is positive. If it results in teaching to the assessment requirements, however, and the consequence is narrowing of the curriculum, then the effect is educationally undesirable. Munoz and Alvarez (2010) state that both positive and negative washback effects have been found in previous studies, although few of these studies have also reported verifiable gains in learning.

Both of the DSAs were recorded and transcribed and excerpts from the transcriptions were analyzed using conversation analysis (CA) methods to determine whether the DSA approaches used in the study (mediated assistance, the group zone of proximal development, and collaborative engagement) had any significant transfer of learning effects from the DSAs to the post-CRA. As a form of qualitative research, the principal aim of CA is to trace the development of inter-subjectivity in an action sequence of adjacency pairs, that is, as part of the microgenetic or turn-by-turn analysis of discourse (Seedhouse, 2004). CA thus provides a micro-developmental look at learner/assistor as well as peer-to-peer interaction. CA is therefore a valuable tool when analyzing language development in DSAs because analysis of the transcripts provides invaluable information to both assessors and teachers on how to become better at assessment and instruction (Rogers, 1980).

To summarize, this study combines quantitative and qualitative data to attempt to arrive at a better understanding of how integrating individual criterion-referenced and interactive dynamic forms of assessment may improve learners' performances on both forms of assessment as well as how *sociocognitive* factors (Hill, 2006) may interact during language development, usage and assessment. These data may provide preliminary indications of the degree to which dynamic and standardized forms of assessments may be integrated or commensurable, which may reveal underlying interaction processes involved in the end-product discrete nature of standardized assessments, thereby providing further direction to the social turn in assessment. Before reporting this study, however, it is necessary to review the previous report undertaken into DSAs.

### **The Previous Dynamic Speaking Assessment Study**

Four applications to DSAs that are of particular relevance to this present study were focused on in the previous DSA report (Hill & Sabet, 2009). First, *mediated assistance* (MA) involves interaction between an assistor and a learner to reveal problems in spoken performance and the mediation by the assistor/assessor (e.g., oral or verbal highlighting of errors) is designed to aid learners in problem solving (Lantolf & Thorne, 2006). In this example, the target is the past progressive aspect (S = student; MA = mediated assistor):

1. S: What did you do before this class?
2. MA: What were you doing...
3. S: What were you doing before this class?

With MA, assistor and learner interaction attempts to develop inter-subjectivity between them to reveal to the assistor misconceptions in the learner's spoken performance. Inter-subjectivity is a perceptual sense of others that constitutes a common intentionality shared by both the perceiving subject and the perceived other (Gallagher, 2008). Analysis of the transcriptions of the DSAs in the previous study revealed examples of learners successfully learning to negotiate meaning through trial-and-error with MA (i.e., identifying how to spot their error and then knowing how to correct it as well as construct upon it).

Second, the Vygotskian (1978, 1986) construct of the *zone of proximal development* (ZPD) was an important factor to the DSAs in this study: the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving in collaboration with more capable peers. The hypothesis in this study, however, was that the same developmental and problem-solving ability between individuals would also become collective and interconnected within groups of learners (Poehner, 2008), that is, a group ZPD (GZPD). Results provided data to support this hypothesis. By observing the previous higher-level pair undertaking the DSA, pairs became interconnected and individual learners performed better. What is meant by collective and interconnected in a GZPD will be further clarified in the present study.

Third, the use of *collaborative engagement* (CE) in the previous study is also of note. CE involves diagnosing problem areas during a DSA, especially when there has not been indication of transfer of learning or the internalization of any cognitive tools, and then focusing on those areas with learners during classroom interaction. Finally, *transfer of*

*learning* (TOL), the degree to which a novel behavior will be repeated and improved in a new analogous situation or context, was analyzed in the context of role-plays of graduated difficulty. Results indicated that if an isomorphic or analogous concept-structure mapping innovation can be applied from the first role-play to a new one of graduated difficulty then TOL may take place.

In the context of the present study, these DSA processes will once again be analyzed to determine whether any new forms emerge or more results obtain from: (1) using MA during dynamic practice sessions prior to DSAs, (2) developing alternative ways to construct the GZPD, (3) using CE when working with lower-level pairs, and (4) attempting TOL between assessments using analogous CRA-based content as well determining common errors which emerged from the first DSA and including them in a common-error DRA in the post-CRA.

## ***Method***

### **Research Questions**

1. Do intervening DSAs analogously based on the content of a pre-/post-CRA significantly increase learners' post-CRA scores?
2. Does correction of common spoken errors found in DSA1 during DSA2 result in significant error correction on a written post-DRA?
3. Does integrating a pre-/post-CRA with intervening DSAs result in a positive washback effect with learners (i.e., increased task-intrinsic motivation)?

### **Participants**

The participants in the study group were first-year university students (N=30) at a university in Saitama, Japan. They had been streamed into mid-level classes and they were majoring in Child Studies to become elementary or preschool teachers. Their average TOEFL PBT scores were between 300 and 320. It was a 13-week semester-long speaking class and classes met twice a week. It was the only speaking class that learners were required to take. All classes were taught by the same instructor.

The data for the control group (N=22) were taken from the previous year's data for the same major and level of class. The control group participated in two intervening speaking assessments but without any dynamic assessment approaches and the speaking assessments were curriculum-based, not CRA-based. With both groups, the pre-assessment was given on the first day of classes, the first speaking assessment was given mid-term, the second speaking assessment was given on the last day of classes, and the post-assessment was given during final assessment week the week after classes had ended.

### **Procedure**

An item analysis of the pre-/post-CRA was first conducted to determine the specific areas of learner difficulty to focus on (see Appendix A for an example section of the pre-/post-CRA). Then five dialogues similar to the CRA but based more analogously on learners' own learning situations were developed: three for DSA1 and two plus the common-error activity for DSA2. Learners practiced these dialogs in a three-class dynamic practice session prior to the DSA, first with MA, then with a new arrangement of the GZPD and, finally, with

CE. The dialogs had blanks which required learners to use some of their own personal information (see Appendix B for the corresponding example). During the DSA, learners were only allowed to look at prompts based on the practice session dialogues (see Appendix C). See Table 1 for the integrated sequence of pre-CRA, intervening curriculum-based instruction, dynamic practice sessions, CRA-based DSAs and post-CRA with the common-error DRA. As was previously mentioned, both DSAs were recorded and transcribed.

**Table 1. Treatment for Study Group**

Week 1	Pre-CRA
Weeks 2 – 4	Intervening curriculum-based instruction
Weeks 5 – 6	Dynamic practice session one (MA, the GZPD and CE)
Week 7	CRA-based DSA1 – plus analysis of learners’ common errors
Weeks 8 – 10	Intervening curriculum-based instruction
Weeks 11 – 12	Dynamic practice session two (MA, the GZPD and CE) – plus practice of learners’ common errors
Week 13	CRA-based DSA2 – plus assessment of improvement of learners’ common errors found in DSA 1
Week 14	Post-CRA – revised to include the common error DRA

**Results**

One obstacle in commensuration between dynamic and standardized forms of assessment stems from typical statistical analyses not being applicable to dynamic results (Poehner, 2008). In an attempt to overcome this obstacle, Table 2 shows the results in percentage scores for the pre-/post-CRAs and the two intervening DSAs. The study group scored 16.1% higher than the control group on the post-CRA and a one-way ANOVA analysis was extremely statistically significant (see Table 3). The study group also showed more consistent and cumulative development between the two intervening DSAs. Both the CRA-based DSAs for the study group and the non-dynamic curriculum-based speaking assessments (NDSAs) for the control group were scored by the same rater according to how well pairs of learners had individually developed and completed their conversations using the prompts. These results suggest that implementing ways to integrate forms of assessment can significantly enhance performance on both and that statistical analyses can be applicable to dynamic assessment results.

**Table 2. Pre/Post-CRA and DSAs Percentage Scores for Both Groups**

	Pre-CRA	DSA1	DSA2	Post-CRA
Study Group (n = 30)	45.4	85.7	89.3	73.6
	Pre-CRA	NDSA1	NDSA2	Post-CRA
Control Group (n = 22)	46.5	83.4	81.9	57.5

**Table 3. Results of the One-Way ANOVA**

	SS	df	MS	F	P
Between Groups	3252.924	1	3252.9241	22.61	< .0001
Total	10447.442	51			

### **The Common-Error DRA**

The common errors that emerged during DSA1 were integrated and attempted to be corrected by learners in DSA2 and then evaluated as a DRA as part of the post-CRA (see Appendix D). Results of the DRA show that 86.4% of the common errors that emerged in DSA1 were corrected on the DRA, which suggests a significant quantifiable indication of development. Table 4 shows the nine most common grammatical errors found in DSA1, examples of each, and the percentage corrected on the DRA. The two lowest percentages, errors 3 and 6, both had other potentially grammatically correct responses (e.g., *I have part-time work* or *Have you any free time?*), so these learners may have scored even higher on correcting these errors had there been partial scoring for potentially grammatically correct responses.

Of interest to research question 2, because learners significantly improved on the common errors on the DRA, it might be suggested that learners not only improved their procedural knowledge through improved speaking ability on DSA2 but also their declarative written knowledge with improved DRA scores. Results also suggest that finding the common errors in DSA1 and having learners attempt to correct them in DSA2 contributes to raising their development to a level appropriate to significantly improve on the post-CRA.

**Table 4. Common Errors and Percentage Corrected on the post-DRA**

Common Errors	% Corrected on the post-DRA
1. Use of the article “a” with a plural	
<i>How often do you have a classes?</i>	85.2
2. Lack of “do” support	
<i>What you do after classes?</i>	85.2
3. Lack of use of “to”	
<i>I have work part-time.</i>	66.6
4. Overgeneralization of “o’clock”	
<i>At 5:30 o’clock.</i>	100
5. Lack of use of “of”	
<i>Are you a member a club?</i>	96.3
6. Incorrect word order	
<i>Have you do any free time?</i>	74.1
7. Combining progressive and infinitive verb forms	
<i>I like to go to shopping.</i>	88.9
8. Using “in” in the question “Where do you live?”	
<i>Where do you live in now?</i>	96.3
9. Use of the singular rather than plural noun.	
<i>Oh, do you like animal?</i>	85.2
Total Corrected	86.4

### **Comparing DSA1 and DSA2**

Excerpt 1 from the transcript for the corrected common-error section in DSA2 shows the common errors found in DSA1. The words in brackets are errors which were not repeated and the words or letters in bold are corrections added by learners (learners’ names have been replaced with capital letters):

Excerpt 1

1. X: How often do you have (*a*) classes?
2. Y: I have classes every day in the morning.
3. X: What **do** you do after classes?
4. Y: I have **to** work part-time.



5. X: What time do you work?
6. Y: At 5:30 (*o'clock*).
7. X: Are you a member **of** a club?
8. Y: Yes, I belong to the tennis club.
9. X: Wow, you are busy. **Do** you **have** any free time?
10. Y: Yeah, on the weekends.
11. X: What do you like to do on weekends?
12. Y: I like to go **shopping**.
13. X: I like shopping, too. Where do you live (*in*) now?
14. Y: Actually, I live near a zoo.
15. X: Oh, do you like animals?
16. Y: Yeah, I really like them. They are so cute!

It must be kept in mind that because of the low level of learners a less structured mediation was not possible. That said, the use of prompts raised awareness in learners of the errors they had made in DSA1. Learners performed better than expected on both DSAs, suggesting a faster rate of development than anticipated, possibly as a result of the analogous CRA-based DSAs. Of note, in DSA2, with the common-error assessment, the use of contractions emerged as the next form to focus on or stabilize in order for learners to become more self-regulating or “native-speaker-like” (e.g., *They’re* so cute!).

### **Comparing DSA2 with post-CRA Results**

Another useful indication of L2 development was to compare the transcripts of DSA2 with the post-CRA results. Excerpt 2 provides an example of overgeneralization of the indefinite article with a plural:

#### Excerpt 2

1. S: How often do you have a classes?
2. E: I have a classes every day in the morning.

In this case, S got the form correct on the DRA, but E did not. Therefore, CE would most likely have been effective with E given the opportunity for more curriculum-based classroom interaction after the DRA. Comparison of the DSAs’ transcripts with the post-CRA/DRA scores reveals which learners are continuing to have problem-solving difficulties and are in need of more CE. A missed opportunity for MA places the responsibility for lack of development not only on the learner, but, and perhaps more importantly, also upon the assistor.

### **The Group Zone of Proximal Development**

In typical standardized assessments and oral proficiency interviews, assessors and assesseees have minimal interaction and assesseees are normally assessed individually. With the GZPD, on the other hand, an entire classroom of learners is being assessed individually

but also collectively in the context of being interconnected pairs and a group, highlighting the potential for more social approaches to assessment.

In this study, learners were originally paired according to their pre-CRA scores (the top N = 1 – 15 paired respectively with the lower N = 16 – 30). Each higher–lower pair observed the previous higher pair participate in the dynamic practice sessions. After the first practice in the three-class dynamic practice session, the sequence of pairs was initially rearranged according to the problem-solving needs of each pair’s performance. After the second practice, partners within each pair were rearranged according to the same needs of each individual performance. In the final practice, as a form of CE, the order of pairs was reversed (i.e., the lower pairs were first) so the pairs in most need of MA then began the practice. In total, perhaps 50% of learners either changed their position in the GZPD or their partner. With the other half, no alteration was found to be necessary.

These three new methods of arrangement within the GZPD during practice sessions seemed to be the most effective at interconnecting pairs to collectively enhance individual performance. After this practice sequence was completed, each DSA was then conducted. The main finding from having an initial practice sequence in this study was that pairs’ interaction showed much less repetition of phrases during the DSA, especially with the common errors found in DSA1. In other words, the interaction between pairs was much more native-speaker-like (NS-like), although repetition is commonly found in typical NS interaction.

### **Mediated Assistance**

In their previous study, Hill & Sabet (2009) found recasts, comprehension checks, and negotiation of meaning to be the main forms of MA during DSAs. Most likely as a result of the practice sequence prior to each DSA, this follow-up study found three new forms of MA to emerge from the interaction between the assistor and pairs: (1) peer-to-peer MA, (2) shadowing by the learner, and (3) intonation, facial expressions, and fillers.

### **Peer-to-peer MA**

Peer-to-peer MA involves the partner providing the MA rather than the assistor. If the partner (K) can notice the place of her partner’s error (M), she may provide assistance before the mediated assistor (MA):

Excerpt 3

1. M: Yeah, yeah...
2. K: On
3. MA: On
4. M: On the weekends

Peer-to-peer MA is not simply giving the answer to one’s partner but providing cognitive support similar perhaps to a private speech utterance one might make to oneself in a cognitively demanding situation. The significance of peer-to-peer mediation is that it further strengthens the central claim of dynamic assessment—that assessment and instruction should be a unified activity (Poehner, 2008). The use of peer mediators is a

result of the establishment of triadic inter-subjectivity. It in no way replaces MA but shifts the focus away from what isolated assistors can do and emphasizes instead what each individual is capable of when functioning inter-subjectively to co-construct discourse because if peers develop the ability to provide peer-to-peer MA they are then the ones who are aware of their more competent role in the ZPD.

### **Shadowing**

Shadowing or simultaneous repetition of the assistor's MA by the learner seems to be the result of the practice sequence. Shadowing occurs when the learner repeats the MA in unison with the assistor:

Excerpt 4

1. MA: One hour AND A half (*K in unison "one hour half"*)
2. K: About one...
3. MA: hour
4. K: half
5. MA: And a half
6. K: And a half (*K in unison*)

Simply reading the transcript does not adequately convey what is meant by shadowing. Listening to the recording would better exemplify that the learner's utterances occur in unison with the assistor and exhibits the anticipatory and predictive inter-subjective nature that develops in the learner as a result of previous MA practice sessions.

### **Intonation, Facial Expressions and Fillers**

If learners have already practiced their dialogs and have had more explicit forms of MA at that time, then during the DSA implicit forms of MA can be used to better assess development. These include (1) a rising intonation as a signal of the need for self-repair, (2) a facial expression such as a raised eyebrow to indicate the need for self-repair, and (3) a filler (e.g., *hmm?*) to indicate the need for self-repair. In the previous DSA study, Hill & Sabet (2009) found the opposite sequence to be the case during the DSA (i.e., implicit to explicit). That sequence, however, was the result of learners having to spot the intent of the MA as well as the development in MA from recasts to comprehension checks to negotiation of meaning and in the present study these had already occurred previously during the practice sessions prior to the DSA. Therefore, as a result of the practice sessions in this study, during the actual DSA, the MA occurred from a more top-down or inter-cognitively internalizing perspective, which was more implicit as a result of the shared inter-subjective perspective towards MA that had developed between the assistor and learners.

The important finding from the development of these newer forms of MA as a result of practice is that in each kind the learner becomes responsible for his or her own self-repair. These forms of MA, along with revealing a continuum from explicit to implicit MA, also reveal its dynamic nature (i.e., from variability to stabilization). After a practice session, MA becomes much less explicit during the DSA and more stable in an implicit sense because, through inter-subjective co-construction with the assistor, learners are self-organizing

their language usage, which perhaps might assist learners in unlocking or understanding some of the discretely organized componential knowledge assessed in standardized assessments.

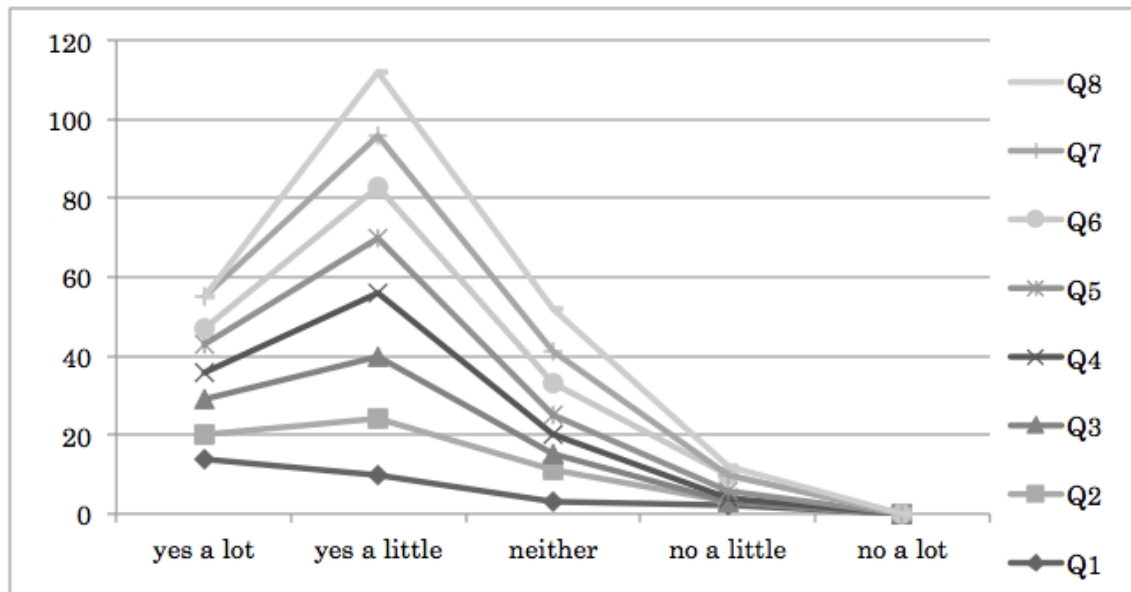
### **The Follow-up Learner Questionnaire**

Research question 3 intended to investigate whether the analogous CRA-based DSAs could make both forms of assessment more task-intrinsically motivating to learners. This is exactly what the responses to the learner questionnaire showed (see Table 5 and Figure 1).

**Table 5. Results of the Follow-up Learner Questionnaire (N = 29)**

1. Did practicing for the speaking test help when doing the written test?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
14	10	3	2	
2. Do you think speaking tests are useful?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
6	14	8	1	
3. Do you think practicing for the speaking test is useful?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
9	16	4		
4. Would you try harder on the speaking test if you knew it could help your written test score, too?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
7	16	5	1	
5. Did you improve your errors when practicing for the speaking test?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
7	14	5	2	
6. Did watching the pair before you help with the speaking test?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
4	13	8	4	
7. Did the teacher correcting your mistakes help with the speaking test?				
Yes, a lot	Yes, a little	Neither yes or no	No, a little	Not at all
8	13	8		
8. Do you have more or less motivation to take English tests after this class?				
A lot more	A little more	The same	A little less	A lot less
	16	11	2	

Responses to questions 1, 6 and 7 add qualitative support from learners to research question 1 in the context of the quantifiable significant gains found from the pre-/post-CRA. Additionally, question 5 is another supportive form of triangulation for research question 2 and improvement on the common-error DRA. Finally, questions 4 and 8 especially reveal that learners found the DSAs to be intrinsically motivating.



**Figure 1: Learners' Responses to the Follow-up Questionnaire**

**Positive Washback**

The follow-up questionnaire responses suggest that learners found integration of the DSAs with the pre-/post-CRA broadened the curriculum and did not narrow learning. They also seemed more willing to interact meaningfully with the dynamic practice sessions in this learning environment. Therefore, it can be suggested that learners' experienced a positive washback effect from the integration of assessments which in turn resulted in more task-intrinsic motivation with them towards both forms of assessment. In this sense, the integration of assessments led to teaching practices that promoted learning, and the effect was educationally desirable. Additionally, this study found a positive washback effect coinciding with verifiable gains in learning.

***Discussion***

Although these are initial findings and should be approached with caution, this study has shown significant quantifiable longitudinal gains and qualitative micro-development from the initial pre-CRA, through the two CRA-based DSA sessions, to the final post-CRA. These results suggest that not only integrating dynamic assessment and the curriculum is effective as Lidz (1991) found but that further integration of forms of assessment (i.e., classroom-based, curriculum-based, dynamic-referenced, development-referenced and criterion-referenced—along with standardized) can lead to even more significant and

accurate assessment results. The linking of forms of assessment also provides a clearer indication of learner development from one assessment to the next.

These results suggest an indication of cumulative language development through dynamic self-organizing processes which stabilize underlying grammatical forms, interactive processes which might also unlock, or allow learners greater understanding of, some of the discretely organized end-product knowledge found in standardized assessments. Dynamic assessment methods do seem to be an effective method to allow learners to access their procedural knowledge, perhaps suggesting inter-/intra-cognitive directional clines from procedural, spoken or social to declarative, written or individual knowledge development. As Gergen (2009, p. 40) points out, interaction does not occur according to the rules in the sense of following rules "inside the head"; rather, patterns of coordination are generated from which what seem to be the rules are later extracted. Standardized assessment seems to be more closely related to this rule-based end result of discrete forms of knowledge, and it is through dynamic social interaction that this knowledge can be extracted. Results of this preliminary study into integrating instruction, the curriculum and forms of assessment suggest that further research should be conducted into attempting to establish commensuration between them as well as between dynamic and other forms of assessment.

Attempts to integrate procedural or dynamic with declarative or standardized assessments appears to establish a link between their opposing clines, resulting in more effective performance in both forms of assessment. Rather than thinking of assessment methods or learning processes as dynamic or discrete, thinking of them as interrelated clines along a dynamic/discrete, social/individual, spoken/written, implicit/explicit, top-down/bottom-up, inter-/intra-cognitive or procedural/declarative dialectical continuum may help to tune more into language development processes. Perhaps then the psychometric control of variables within standardized pre-/post-assessment research could also allow for the inclusion and analysis of variables which emerge during the study.

When analogously basing a DSA on the content of the pre-/post-CRA, it is important to keep TOL in mind (i.e., learners' ability to transfer what they learned in the DSA to the more discrete post-assessment), because significant TOL and development can result when both assessments correspond analogously. Some degree of TOL emerges from dynamically practicing each part separately and then assessing all three parts together in the DSA (i.e., the DSA interaction is more NS-like with much less repetition). During a dynamic practice session, any amount or form of MA is acceptable. During the DSA, however, the more explicit form of MA that is necessary reveals the lower potential development of the learner, especially if the learner received explicit MA for the same form in the practice session. Additionally, it is not a matter of assessing learners negatively if problem-solving difficulties persist during the DSA; rather, a different MA response is required, i.e., other more NS-like MA forms (e.g., peer-to-peer mediation, shadowing, or intonation, facial expressions and fillers).

During the practice session prior to the DSA, MA is bottom-up. However, in the DSA, to determine any cumulative developmental effects from practice, it stems from a more top-down inter-cognitive perspective of the learning continuum and if intra-subjective development has occurred learners are then encouraged to make their own self-repair.

Receptivity to MA develops inter-subjectivity and frees learners' cognitive tools to focus on their performance in order for them to stabilize form and become self-regulating. MA, GZPD, CE and TOL can all be considered as *co-constructionism* approaches (i.e., learning by constructing one's knowledge through focusing on how social context provides opportunities for making connections to what is being learned), as opposed to *instructionism* approaches (i.e., learning by transmission of knowledge) (Sawyer, 2006).

### **Limitations of the Study and Future Research**

One limitation of this study was the lower level of the learners. It required the form of dynamic assessment to be similar to other less dynamic and more conventional L2 acquisition methods. Also the Control Group's speaking assessments were non-dynamic and curriculum-based. The previous study (Hill & Sabet, 2009) showed that DSAs outperform NDSAs so a follow-up study would need to determine whether CRA-based DSAs outperform curriculum-based DSAs. Additionally, in order to develop a common-error pre-/post-DRA, it is necessary to determine exactly which learners made which errors in DSA1 and their ability to correct them in DSA2 and on the DRA. Finally, future research should look at the commensurability of integrating dynamic assessment with large-scale or high-stakes norm-referenced forms of assessment.

### ***5. Conclusion***

This study attempted to go a step further than investigating the link between instruction and dynamic and curriculum-based assessment and investigated integrating instruction, the curriculum, and dynamic and criterion-referenced assessments. Results showed that through integration learners performed significantly better on both forms of assessment. Learners also indicated positive washback and a higher task-intrinsic motivation to participate in more integrated forms of assessment. These preliminary results suggest that dynamic and standardized forms of assessments may be commensurable and integrating them might also be effective in unlocking some of the discretely organized knowledge found in standardized assessment, which would provide support for further direction to the social turn in assessment. These results have shown how untangling and coordinating the complex interaction of cognitive and social factors occurring in instruction, curricula and assessment may improve learners' performances on all forms of assessment, thereby revealing how sociocognitive factors interact during language development, usage and assessment.

### **About the Author**

**Kent Hill** is a lecturer at Meiji University. He received his Ph.D. from the University of Nottingham in 2006. His main area of interest is sociocognitive and dynamic approaches to language development. His current interests include integrating forms of assessment (i.e., dynamic, cognitive diagnostic and standardized).

## References

- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Barab, S. (2006). Design-based research: A methodological toolkit for the learning scientist. In R. K. Sawyer (Ed.), *The Cambridge handbook of learning sciences* (pp. 153-169). Cambridge: Cambridge University Press.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to language testing. *Applied Linguistics*, 1, 1-47.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43, 393-415.
- Gallagher, S. (2008). Understanding others: Embodied social cognition. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 439—452). San Diego: Elsevier.
- Gergen, K. (2009). *Relational being*. Oxford: Oxford University Press.
- Haywood, H., & Lidz, C. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge: Cambridge University Press.
- Hill, K. (2006). A sociocognitive perspective: The best of both worlds. *TESOL Quarterly*, 40, 819-825.
- Hill, K., & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly*, 43, 537-545.
- Lantolf, J. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Lantolf, J., & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Leung, C., & Lewkowicz, J. (2006). Expanding horizons and unresolved conundrums: Language testing and assessment. *TESOL Quarterly*, 40, 211-234.
- Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York: Guilford.
- McNamara, T., & Roever, C. (2006). Language testing: The social dimension. *Language Learning*, 56: Supplement 2.
- Munoz, A., & Alvarez, M. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27, 33-49.
- Poehner, M. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, 91, 323-340.
- Poehner, M. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Berlin: Springer.
- Rogers, C. (1980). *A way of being*. Boston: Houghton Mifflin.
- Sawyer, R.K. (2006). *The Cambridge handbook of learning sciences*. Cambridge: Cambridge University Press.



Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Oxford: Blackwell.

Vygotsky, L. (1978). *Mind in society*. Cambridge, Mass.: Harvard University Press.

Vygotsky, L. (1986). *Thought and language*. Cambridge, Mass.: The MIT Press.

Widdowson, H. (2001). Communicative language testing: The art of the possible. *Studies in language testing* (pp. 12-21). Cambridge: University of Cambridge.

## Appendix A

This is the example section from the pre/post-CRA:

Part 3

Questions 16 – 20. (5 x 1 point)

Complete the conversation between Paul and Tomoko.

For questions 16 – 20, mark A, B, C, or D on the answer sheet. Choose from the box below.

EXAMPLE: 0.

A. class B. period C. schedule D. study

ANSWER: C

Paul: What's your school (EXAMPLE: 0) like?

Tomoko: It's pretty busy. I have classes every day (16) the morning.

Paul: Sounds tough. How are your classes?

Tomoko: They're all right. But I don't get much time to study, because right after school I have to work (17).

Paul: Wow, you're really busy. Do you have to work on the weekends?

Tomoko: Yes, I have to work on Saturdays, too. Hey, could you (18) me the time?

Paul: Uh, it's almost four.

Tomoko: Oh no! Sorry, but I've got to hurry.

Paul: (19) How come?

Tomoko: I've got an appointment with my math professor to talk about some homework. I can't (20) math!

Paul: Well, let's talk later when you're not so busy.

Tomoko: Okay, maybe at lunch tomorrow. See you.

Paul: Bye.

16. A. on B. at C. in D. for

17. A. restaurant B. at part-time C. in part-time D. part-time

18. A. have B. tell C. say D. teach

19. A. Do you see? B. Really? C. Perhaps? D. Got it?

20. A. stood B. hate C. stand D. like

Write ONE word for each space (21 – 25) on the answer sheet.

Sayuri: Where was this picture (EXAMPLE: 0. **taken**)?

Takumi: In Niigata. That's where I lived for two years.

Sayuri: Where do you (21) now?

Takumi: In Saitama City.  
Sayuri: Do you live with your family?  
Takumi: No, I live by ( 22 ) in an apartment.  
Sayuri: What's your neighborhood ( 23 )?  
Takumi: Oh, it's pretty noisy. There are a few game centers and karaoke bars.  
Sayuri: Do you have ( 24 ) brothers or sisters?  
Takumi: Yeah, I have an older sister. She works part-time at a hospital.  
Sayuri: How about you? Do you work part-time anywhere?  
Takumi: Yeah, I'm a cashier at a convenience store.  
Sayuri: ( 25 ) often do you work?  
Takumi: Usually four days a week.  
Sayuri: What do you like to do in your free time?  
Takumi: I like to play basketball and watch TV.

## Appendix B

This is an example of what pairs practiced during the dynamic practice session. Learners provided their own answers for the underlined parts:

A: How often do you have classes?

B: I have classes Mon / everyday in the morning / afternoon.

A: What do you do after classes?

B: I have to work part-time / have club activities / go home.

A: How often do you have to work part-time / have club activities / go home?

B: Always / usually / sometimes.

A: Where do you live now?

B: I live in Akabane.

A: Do you live with your family?

B: Yes, I do. / No, I live by myself in an apartment.

A: What's your neighborhood like?

B: It's quiet. / noisy.

A: Do you have any brothers or sisters?

B: Yeah, I have one brother. / No, I don't.

A: Could you tell me the time, please?

B: It's 10:30.

A: Oh, sorry. I have to go.

B: Really? Where?

A: I have a dentist's appointment.

B: Oh, I'm crazy about / I can't stand the dentist.

A: Me, too. / Me neither. See you.

B: Bye.

## Appendix C

These are examples of the prompts that learners were allowed to look at during the DSA:

A:

B: Mon / everyday ... morning / afternoon.

A:

B: have to work part-time / have club activities / go home.

A: have to work part-time / have club activities / go home?

B: Always / usually/ sometimes.

A:

B: Akabane.

A:

B: Yes, I do. / No, I live by myself in an apartment.

A:

B: quiet. / noisy.

A:

B: Yeah, I have ... / No, I don't.

A:

B: 10:30.

A: Oh, sorry.

B: Really?

A: dentist's appointment.

B: Oh, ... the dentist.

A: ... See you.

B: Bye.

## Appendix D

This is the common-error post-DRA which was added to the post-CRA:

Look at the sentences. There is a mistake in the sentence. Find the mistake and then write the sentence correctly. For example:

1. How often do you have a classes? \_\_\_\_\_
2. What you do after classes? \_\_\_\_\_
3. I have work part-time. \_\_\_\_\_
4. At 5:30 o'clock. \_\_\_\_\_
5. Are you a member a club? \_\_\_\_\_
6. Have you do any free time? \_\_\_\_\_
7. I like to go to shopping. \_\_\_\_\_
8. Where do you live in now? \_\_\_\_\_
9. Oh, do you like animal? \_\_\_\_\_

© Copyright rests with authors. Please cite TESL-EJ appropriately.