

DYNAMIC ASSESSMENT IN THE FOREIGN LANGUAGE CLASSROOM

Maria TSVETKOVA

Sofia University "St. Kliment Ohridski"

E-mail: mary.tsvetanova@gmail.com

ABSTRACT: The present article outlines a sociocultural perspective on L2 classroom evaluation. It explores a way of operationalizing the abstract constructs that inform the Vygotskian theoretical framework and examines factors involved in the continuing interplay between social and individual functioning. The main argument of this paper is that combining traditional, psychometric approaches to language ability evaluation and dynamic assessment performs a twofold function. It allows for reconceptualizing the individual student's role in the process of foreign language (FL) performance and significantly enhances subsequent achievement. The cross-cutting theme of the present paper is that eliciting FL use mainly by static testing instruments and conventional procedures implies constructing a depersonalized, 'averaged' image of the L2 learner. Such quantification presupposes a unitary and self-contained language learner detached from the context of FL use. In developing these themes, I explore the relation between identity reconstruction and increase in L2 grades. The subjects of the study consist of 50 students – 25 participate in the experimental and 25 subjects are included in the control group. The results of the conducted quasi-experimental study this paper reports on provide evidence that empowering students to act as autonomous individuals in the classroom relates positively to increased FL achievement as registered in the results of formal summative measurement.

KEYWORDS: dynamic assessment, foreign language performance, sociocultural theory, activity, Zone of proximal development, mediation, internalization

1. Introduction

The present paper discusses how Vygotsky's conceptual framework for investigating mental functioning can inform classroom-based L2 evaluation research. In order to gain a thorough understanding of students' foreign language performance, one needs to account for the interdependence between inter-psychological and intra-psychological processes involved in language use. Such a broad perspective allows for a focus on L2 students' interrelationships with the learning environment that are constitutive for identity formation.

An integrative sociocultural approach to identity construction is proposed, an approach that utilizes a number of elements of Vygotsky's conceptual work adopted and adapted to the field of foreign language acquisition and evaluation. I intend to explore language performance in a local L2-activity educational setting where participants are actively engaged in collaborative discourse creation.

In exploring inter-psychological processes and their role in students' L2 achievement, this study is nested within research endeavors aimed at challenging the polarity of inner versus outer, endogenous mind versus social environment and culture, emphasizing their situated and dynamic nature. Further, it is based on the recognition that learners are not static entities but are always in states of transition and transactional relationships with the social setting. Inspired by research conducted by Frawley and Lantolf (1988), and Poelner and Lantolf (2003; 2005; 2015), I have set out to investigate the theoretical scope and the applied practice of the construct *dynamic assessment*. The classroom-based study carried out is intended to build upon research endeavors aimed at finding the practical implications of the *dynamic assessment* framework. The research design provides for applying the theoretical concepts underpinning it to the specific conditions of a Bulgarian foreign language classroom. Having reviewed advances in FL assessment instigated by the acknowledged need for L2 learners to demonstrate practical command of communicative skills, I explore the relevance of activity-oriented L2 elicitation tasks developed within the sociocultural paradigm.

The main argument advocated throughout the paper is that approaches which disregard the dialectic relation between the inter-psychological and intra-psychological processes construe the L2 learner as an averaged and fixed object of assessment rather than a full-fledged communicator and autonomous subject. The conceiving of L2 performance within the *dynamic assessment* framework is in harmony with M. Halliday's work (1984) whose functional approach to language gives emphasis to both system and behavior. Furthermore, the *dynamic assessment* model presents a way of overcoming the

input-output metaphors in SLA that construes “mind as a container in which knowledge is memorized and stored in the form of static representations” (Dufva, 2003, p.131). FL performance in an evaluative context, as well as teaching, is seen as targeted at communication, collaboration, cooperation, and problem-solving skills. The sociocultural theory (hereafter, SCT) “takes into account the complex interactions between the individuals acting with mediational means in the sociocultural context” (Swain and Deters, 2007, p. 821).

The conceptual approach proposed here is based on the general view that the L2 learner does not operate in a social vacuum but is a part of a more complex community of practice, one in which they gain access and reaffirm their role and position through language use. The framework adopted also allows for conceptualizing the L2 classroom as an eco-social system having a profound effect on L2 achievement.

The sociocultural perspective allows for designing a knowledge-based assessment model, that is, one grounded not on declarative knowledge or procedural knowledge, but rather on knowledge constructed individually or in a group. This theoretical paradigm is oriented toward the overall communicative performance elicited from L2 learners while they are using the language as part of a goal-directed, meaning-focused activity which necessitates the integration of skills and knowledge, in terms of Brindley (1994, p.74). The application of a SCT-grounded *dynamic assessment* approach consists not only in group work, but in allowing students to project their identity on the collaborative discourse construction. It is a task-based approach aligned to issues concerning their autonomy, agency, and self-image. L2 assessment is implemented in the very process of goal-oriented language use realized while participating in collective inter-subjective exchange activities in which the individual learner could project their language ego.

Within the Vygotskian paradigm, intrinsic and extrinsic factors constitutive for (re)constructing cognition, agency and identity are not dualistically separated. The student identity could be seen as partially construed in and through their classroom interactions in the TL. It is in adapting to the social environment that these identities are reiterated and reinforced, or deconstructed. The reason I chose the Vygotskian SCT framework for a basis of the presented L2 evaluation research is that it allows for an agentive role of the student. This provides for obtaining more powerful, informative insights into the interrelatedness of identity, autonomy, interdependence, collaboration and L2 achievement that are hardly possible within structuralist accounts of L2 evaluation processes and procedures.

Several key issues impinge on the present discussion. Subjectivity is produced in a social setting, which is structured by relations of power. In interactional exchange, an individual person plays out different aspects of their identity and takes up different subject positions. Some of these may be in conflict, as Norton (1995) points out. Within a SCT framework, students, in addition to teachers, could be seen as possessing symbolic resources. The emphasis is not on the acquisition and reproduction of knowledge, where learning is seen as information seeking and sedimentation of knowledge in individuals' minds. Instead, the *participation* metaphor is emphasized. Rather than an individual act of knowledge accumulation, learning represents active processes of legitimate engagement in collaborative knowledge production. This perspective entails breaking with the unresolvable polarity of inner versus outer, endogenous mind versus external culture.

It is from this theoretical stance that L2 performance is examined here with the main focus placed on *participation* and *co-construction* rather than on *memorization* and *reproduction* of knowledge. It is not possible within the limits of the present study to evaluate all the major aspects of Vygotsky's theory of development. I have emphasized his conception of the role of social exchange and collaboration, and adapted it to the field of collective discourse creation for the purposes of L2 evaluation. Within this conceptual framework, L2 use is approached and examined as a social and participatory, creative and collaborative process. In the present paper, a dialectical view is adopted which transcends the Cartesian dichotomy between the internal and the external realms, and allows scope for L2 learners to be conceptualized as autonomous agents of their own achievement.

2. Theoretical Constructs and Premises

A historical milestone in the field of L2 testing and evaluation is marked by conceptualizing and implementing the pedagogical processes in a sociocultural perspective. The main argument in this paper is that, within this conceptual paradigm, the interplay between L2 use and identity construction could be clearly seen. This approach involves L2 assessment research and practice in an intensive dialogic process that allows for providing a multi-dimensional profile of L2 learners' competence in the TL. The distinguishing feature of this conceptual paradigm consists in the way subjects are positioned in a task, or a test. In light of this, what is decisive is the potential for opening vistas and broadening the perspective to include also identity formation. The focus in the present paper is on theoretical premises and constructs, with empirical studies used as illustration.

The sociocultural theory was conceived by the Soviet psychologist Lev Vygotsky (1978; 1981) in the early 1920s. This theoretical framework brings to the fore the pivotal role of the interpersonal relationships and culturally constructed activities in conditioning mental structures and functions. It seeks to “theorize and provide methodological tools for investigating higher cognitive processes by which social, cultural, and historical factors shape human functioning” (Daniels, 2001). Sociocultural theory is thought of as a theory of subjectivity. One of its main contributions is providing a conceptual framework for (re)conceptualizing relations between an individual person and their sociocultural context.

Vygotsky's theoretical paradigm has been widely used in educational research. In this paper it is utilized for the purpose of exploring the versatile process of measuring L2 knowledge and ability for use. This pedagogical framework could be usefully adopted when research interest is focused on its more comprehensive function that encompasses also enhancing the process of second language learning. Conceptualizing L2 evaluation in the light of a toolkit that could aid students in building and restructuring linguistic knowledge is in line with one of the fundamental theses of the sociocultural view of human development. According to this thesis, development proceeds from the inter-mental to the intra-mental plane (Vygotsky, 1997). Grounding L2 evaluation in the sociocultural research tradition allows for this complex pedagogical process to be rendered in terms of negotiation and creation of meaning in collaborative activities.

Within the *dynamic assessment* framework, L2 evaluation could be considered as a creative, dialogically constituted process rather than a snapshot product derived from the passive reception on the part of the testees of the test's content and structure put over by someone else. This approach is in harmony with the way Mitchell and Myles (1998, p. 162) see learners – as “active constructors” of their own learning environment.

L2 testing and evaluation conceptualized within sociocultural thought presents learners with a wide range of tasks and demands that include the individual in an active and dynamic participation in socially- and culturally-mediated activities. The latter process fosters the reinforcement of already acquired information as well as the construction of new linguistic knowledge. L2 assessment, conceived of as an intensive interaction process, does not rely on measuring static declarative knowledge that could be blurted out in a mechanistic way, but rather challenges L2 students to set in motion their full potential and develop their personal voice.

2.1. Zone of Proximal Development

In line with the main tenets of the sociocultural paradigm, L2 evaluation is examined in this paper as a mediated, assisted process geared to students' Zone of Proximal Development (ZPD), rather than in the light of their isolated efforts and individual achievement. The ZPD is one of the fundamental concepts constituting the sociocultural framework, and connotes the interactive processes in which social and personal elements achieve a synergetic effect. This notion provides the basis for the examined testing and evaluation approach that accounts for identity (re)negotiation as L2 learners are empowered to establish themselves as more competent, self-dependent, autonomous L2 users.

In terms of Little, Vygotsky's ZPD (1978) offers an explanation for the connection between autonomy, interdependence and collaboration. In Vygotsky's conceptualization, according to Shayer (2002), the creation of a ZPD is a crucial feature of learning. Interacting with “experts” and cooperating

with peers in “the dynamic region of sensitivity in which the transition from interpsychological to intrapsychological functioning takes place” (Turuk, 2008, p. 249) could rouse to life a range of functions that are still in an embryonic state. It is in this complex process that the Russian psychologist identifies one of the main functions of formal education. He reasons that effective instruction should be targeted not only at the already developed functions but also at those that are yet to mature into completely developed ones. Once the functions, dialogically produced in the course of the individual–social inter-exchange, are internalized, they turn into integrated components of the learner's independent developmental achievement. My conviction is that evaluation should follow the same path.

Vygotsky's notion of the ZPD was originally conceived as a concept to refer to children's developmental processes, and subsequently it has been picked up, adapted and employed in a wide range of academic disciplines and professional areas to fit into different applied linguistics contexts. In the area of FL evaluation, teacher- and peer-supported collaborative work in the ZPD could provide a more comprehensive profile of students' communicative competence in the target language (TL) based on foregrounding psychological activities as main units of evaluation and analysis. It is in the context of dialogic relationships that learners are (re)constituted from objects of measurement into autonomous co-participants, authors and narrators, and their subject positions are (re)negotiated. Such an approach to L2 assessment, embracing environment-triggered functions that are in a process of maturation, allows for a context-embedded measurement of learners' acquired knowledge – in the dialogic field of interaction between inter-personal and intra-personal processes. Following this line, L2 assessment grounded in Vygotskian thought could be seen in opposition to atomistic approaches centered on measuring de-contextualized discrete linguistic items and skills “contained” inside an individual learner's mind.

In addition to facilitating L2 formative and summative assessment, which entails the social and institutional dimensions of the evaluation process, measuring L2 knowledge and skills located in the ZPD could also have repercussions on the affective domain and the individual student's self-image. Unarguably, both are of crucial importance for performance in assessment situations as well as for fostering the enthusiasm and commitment for further development of language proficiency. As already alluded, the implications of the ZPD construct in L2 evaluation could also be considered from the perspective of its potential to enhance the learner's ego and self-image, as well as exert influence on related constructs.

2.2. Mediation

Comprehending the mechanism of FL development is contingent on the concept of mediation, which underlies the transference from the social to the individual domain and from elementary functions to higher, more complex ones. By using this construct, the Russian psychologist arrives at an answer to the question of how human development occurs. Mediation is determined as the distinguishing characteristic of highly developed forms of human behavior. The construct has been adopted, and adapted, in a wide range of research contexts. In this study it extends to activity-based forms of socio-communicative practices for the purpose of FL evaluation.

In terms of Vygotsky, this construct, defined in relation to the meaningful aspect of speech, embodies the significant role played by important people from the learner's social surrounding. The interaction processes with key figures for the child's development determine the nature and the course of learning. Vygotsky (1978, cited in Wertsch 1985) argues cogently that the secret of effective learning rests in the dialogic exchange between two or more people having different levels of knowledge and skills. This involves providing support to the learner for him/ her “to move into and through the next layer of knowledge or understanding” (Turuk, 2008, p. 251). Mediation through a person is one of the three major categories identified by Kozulin (1990, 2003) in elaborating upon the concept of mediation, along with mediation through material tools and mediation through symbolic or psychological systems. Vygotsky conceives of tools as mediators, and one of the most important tools is posited to be language. For the purposes of this study, the abstract construct of mediation is operationalized in terms of task-based assessment that provides for proliferation of meanings and dialogic construction of participants' roles and identity in classroom discourse.

It has been argued (Kozulin and Presseisen, 1995) that Vygotsky views the learning process not as an independent exploration of the physical and social world by the child on his/ her own, but as a process of appropriation and internalization of the set of norms and methods of behavior that function in a given culture. In the process of appropriation, symbolic tools or artifacts play a crucial role, and “scripts”, or schemes, and principles are being incorporated into an individual’s intra-mental functioning to be ready for use when approaching a wide variety of situations and events.

In the field of pedagogy, the knowledge and skills that learners initially acquire as a result of interaction are subsequently internalized as individual functions. The dialectical tension between intermental and intra-mental processes seems to be the essential prerequisite for constituting the ‘final’ state of *knowing*, or *being competent in a FL*, at a particular point in time. Obviously, this consideration has a bearing on the central concerns of L2 testing and evaluation – processes that serve mediating purposes themselves.

2.3. Internalization

A core explanatory construct of the sociocultural theory that has a bearing on foreign language evaluation is the notion of internalization. It is used to denote the process by which intermental functioning in the form of social relations among individuals and interaction with socially constructed artifacts is turned inwards and transformed into intramental functioning (Vygotsky, 1978). Applying the *internalization* concept to L2 classroom contexts could enhance interaction among students.

In addition to the significance of communication between L2 learners, scholars have also pointed out another important aspect of the operationalization of this concept, namely – how vital it is for FL teachers to realize that the role of the expert is not necessarily limited to that of an instructor. It can also be assigned to learners who have already internalized specific aspects of the language. To have the experts aid the less knowledgeable learners, reciprocal teaching can be implemented. In adopting this approach, the expert learner assumes the teacher's role within a micro-community.

Reflecting on possible ways in which the notion of internalization could be applied in specific L2 teaching and assessment contexts is a worthwhile endeavor. We could assume that what underlies the process of internalization is language use while performing a collaborative activity, which both generates and depends on the creation of a shared contextual framework. One way of achieving this is aiming for metacognitive intermental functioning.

2.4. Implications for L2 Classroom Evaluation

A number of practical implications could be derived from the key theoretical concepts constituting the sociocultural framework. However, the possible ways in which they could be operationalized in research grounded in the dialectic between the social and the personal need a much deeper discussion than space limitations would allow for. The present paper is focused specifically on one way of putting them in operation which allows for a holistic approach to teaching, learning and assessment. It has been argued that teachers' efforts put into increasing the L2 competence of their students should not be anchored around mastering separate skills (Zimmerman, 1997). When applied to the field of L2 testing and evaluation, this tenet is highly pertinent to the notion of the *washback effect*. Arguably, prioritizing knowledge of isolated facts and skills development might result in learners failing to make progress from one proficiency level to the next. It could be a stumbling block to both acquiring skills necessary to cope with challenging tasks and mastering various aspects of literacy, such as meaning construction, flexibility and fluency, indispensable for working with pieces of discourse as readers or writers. Additionally, it could prevent L2 learners from developing the ability to express their personal and social identity.

In contrast, engaging students in joint, goal-directed activities, in which they team up to construct together their *intermental zone* in mutual scaffolding, provides opportunities for integrating teaching and assessment. It opens the way to reconceptualization of the L2 measurement process where evaluating testees is imbued with new meanings and consideration is given to both the social and the personal dimensions constituting an inseparable dialectic unity. Within this approach, which has its theoretical

basis in Vygotsky's thought, testing and evaluation are not seen as a standardized solitary activity, unifying and lumping students together, but as a meaningful language-related experience that allows for self-presentation and ego enhancement, and encourages the L2 learner to play out their identity.

The implications of these abstract claims for the practices of L2 evaluation come into focus when we consider how L2 use is shaped by various mediational means. Within the SCT framework the focus is placed on the agentive role of the learner as an individual who imposes their personality and intentionality upon the test or task in hand. In view of this, Donato argues that it is important for teachers to concern themselves to a greater extent with students' orientation to tasks than solely with task outcomes. Consequently, continuing Donato's theme of scaffolding, tasks need to be seen in terms of 'emergent interactions', collaborative sharing and construction of knowledge, rather than as recipes for ensuring specific kinds of language performance. These assumptions could be seen as reinforcing the position that L2 evaluation might be employed as a means of L2 teaching. One of the fundamental tenets of Vygotsky's theoretical approach to human cognitive functioning is related to the role of the interactive behavior in the formation of intra-mental processes constitutive for identity (re)construction.

The tasks developed within a SCT-based module are conceptually grounded on a view that brings to the fore issues related to reconceptualizing students' identity and agency. By stimulating interaction, reaction, and collaboration in FL evaluation, students could be transformed from objects of evaluation working in isolation to a network of writers who co-construct, represent and negotiate meaning in a particular context of communication, in terms of Miller (1984) and Bakhtin (1986).

Designed within a sociocultural constructivist paradigm, the *dynamic assessment* module is intended to result in communication processes and formation of collaborative relationships that could potentially encourage students to share in the challenging endeavor of seeking and planning solutions to emerging problems. It presents itself as a way to elicit authentic language output produced in scaffolding-enhanced student-centered activities, while the micro-social community of L2 learners is operating in their *zone of proximal development*. Using as a starting point the constructivist argument that "all cognitive development is first and foremost inter-psychological" (Donato, 1994), I make the tentative proposal that the evaluation of the acquired L2 knowledge and skills should be premised on the same principle and follow the same general access route to learners' intra-psychological processes.

3. Research Perspectives

3.1. Setting the Scene

After providing an account of what the overall idea of SCT-based evaluation is, it would be of value to tie these abstract considerations and concepts back to some of the specific processes in the FL classroom. In order to ground the theoretical and abstract into the practical and concrete, and put a more "real-world" spin on what may seem far-fetched, I have conducted a specially-designed quasi-experimental study.

For the purpose of exploring the effect of systematically administered socioculturally-grounded performance tasks and procedures, a *dynamic assessment* model has been applied and integrated in FL evaluation in a Bulgarian high-school setting.

3.2. Research Parameters

The **research design** includes applying a *dynamic assessment* model to the experimental group and making a between-group comparison of learners' L2 success at the end of the research period to determine if there was a statistically significant variation in grades. The treatment was administered for 14 sessions held twice a week with 90 minutes allocated for each session.

The **general aim** of the study is to explore a performance assessment model developed within the Vygotskyian tradition. Testing empirically the proposed theoretical model of L2 assessment entails its design, implementation and analysis of the results. The model is intended to provide a multidimensional, cumulative profile of students' L2 knowledge and skills as well as to support interlanguage development.

The **major task** is to explore the pedagogical impact of the performance model by making a cross-group comparison of L2 achievement. The variation in the levels of language achievement has been rendered in quantitative terms by calculating the inter-group difference in the results of the formal evaluation as registered at two different points of the research period.

The examined construct of L2 achievement is quantified and computed as the average of the scores obtained on standard evaluation forms. I have used as a pretest measure the first-school term grades – as obtained before the administration of the treatment, and as a posttest measure – the second-school term grades, as received after subjecting the experimental group to the treatment. It should be noted that the grades obtained by the participants on the experimental tasks and procedures included in the newly integrated task-based assessment module are not taken into account in calculating the mean values of the formal assessment grades for the purposes of the quasi-experimental study.

The **study population** is formed by eighth-grade students of English as a foreign language (EFL). Due to practical constraints related to the implementation of random sampling, the subjects are selected on the basis of previously formed groups in a concrete educational institution. The school involved in the quasi-experiment has been chosen through *a random selection*. The **subjects** of the study consist of 50 EFL learners. The students in the experimental class are assigned to a “treatment” condition (n = 25) and the students in the control class – to a “standard” (n = 25) condition where L2 performance is elicited by standard forms and procedures.

3.3. Research Question and Hypothesis

In line with the primary purpose of the study, the main **research question** is as follows:

Will there be a statistically significant increase in L2 achievement in any of the research groups, as reflected in the results of the formal evaluation?

The **research hypothesis (RH)** states that when comparing students' L2 grades in the control and the experimental group, there will be a considerable inter-group difference in the variation in L2 success.

3.4. Types of Tasks Used for Eliciting Linguistic Data

What follows is a description of the first component that has been included in the L2 *dynamic assessment* system under investigation. The proposed model grounded on a sociocultural approach to educational assessment is hypothesized to offer a form of eliciting written discourse in a way approximating the dialogic co-construction of meaning in oral communication. Assigning students the role of situated interlocutors, it is intended to stimulate them to perceive dialogue as discovery of meaning and knowledge in opposition to mere performing a teacher-developed exercise required for obtaining grades. This component is based on the rationale that language is socially constructed, that meaning resides not so much in the grammatical forms and words *per se* but in the understanding that emerges between participants in an interactional exchange.

Students have been required to put together collective stories in several steps and perform the collaboratively constructed piece of work in the public discursive field represented by the L2 classroom. Performance in the target language is scored for sociolinguistic elements as well as linguistic knowledge. Assessment has been based on the participation in this peer-mediated activity including both co-creating written pieces of discourse and providing feedback to other participants. Thus, when performing this task, L2 learners have been assigned the role of narrators and assessors at the same time, in accordance with their locally determined goals.

Subjects selected for the study from the experimental class have been divided into groups. Each group of students has been presented with a sequence of pictures. The task of creating an episode of a collective narrative has been assigned to every participant. In addition, each one is required to edit the piece of discourse compiled by one of his/ her peers.

The second component that has found place in the L2 evaluation system entails **dictogloss tasks** (De La Colina and Garcia Mayo, 2007; Swain and Lapkin, 2001). An audio recording has been played through and students have been allowed to take notes while listening to it. They have been given 20

minutes to think together and share what they have heard with each other. Then, participants are required to reconstruct the text following the oral exchange of information and ideas.

The **dictogloss technique** provides for operationalization of the abstract idea of the class as a social network of interdependence, as well as representation of the classroom discussion as constituting a dynamic, discursively created space that allows for self-regulation, exercising autonomy, and self-organization. The dictogloss tasks are selected for the purposes of this research study because they are directed at the metaskills that are indispensable for L2 learners to become autonomous agents capable of controlling and self-directing their language performance experiences. Semiosis, and the concomitant L2 identity re-constitution, emerges in the process of meaning-making – during individuals' engagement with others.

The third component nested within a cooperative L2 evaluation approach consists of **jigsaw tasks**. Again, L2 learners have been put in groups. The outcome of the task is a story collaboratively compiled by each student group. Participants are required to jointly manage components of the task accomplishment. In the process of working on the task, they are expected to construct collectively a scaffold for each other's L2 performance. As with the previous types of tasks described above, jigsaw tasks are posited to promote both the negotiation of meaning, the negotiation of form, and negotiation of subject position.

The jigsaw task is mainly aimed at encouraging learners to ask for information, to seek clarification, to use circumlocution as well as a whole range of linguistic and nonlinguistic resources they have mustered to negotiate meaning, to stick to the communicative task (Savignon, 2006).

This assessment approach is premised on the understanding of narrating as a process in which learners represent and organize their experience in a way that allows for eliciting their knowledge of morpho-syntactic structures within social contexts. Task-based L2 assessment allows for alternating the social conditions under which expression of ideas in the target language and interaction take place. It provides a straightforward way for evaluating L2 learners' ability to encode 'ideational meaning', to create 'textual meaning' as well as 'interpersonal meaning' (Halliday, 1973) – if we refer to Halliday who foregrounds inter-individual rather than intra-individual mechanisms.

Designed within a sociocultural constructivist paradigm, the *dynamic assessment* module is intended to result in language use processes and formation of collaborative relationships that could potentially encourage students to share in the challenging endeavor of seeking and planning solutions to emerging communicative problems. It presents itself as a way to elicit authentic language output produced in scaffolding-enhanced student-centered activities, while the micro-social community of L2 learners is operating in their zone of proximal development.

Using as a starting point the constructivist argument that “all cognitive development is first and foremost interpsychological” (Donato, 1994, p. 35), I make the tentative proposal that the evaluation of the acquired L2 knowledge and skills should be premised on the same principle and follow the same general access route to learners' intrapsychological development. Cognitive processes unfold in meaning-focused participation in cultural, linguistic, and historically formed contexts. In a similar vein, IL development takes place, and is most efficiently observed and assessed, in the process of participation in L2 communicative exchange. Following these lines, taking part in inter-psychological processes, in terms of Vygotsky, is also posited to provide for considering the perspective of students' identity (re)construction in L2 classroom performance.

3.5. Reporting Descriptive and Inferential Statistics

A pretest-posttest design is employed in order to investigate the impact of a socioculturally-grounded L2 evaluation approach on students' subsequent L2 achievement. To obtain the present dataset, standardized school grades are used as the quantitative estimate to determine the size of the population effect. I have used as a pretest measure the participants' *first school term grades* obtained before administration of the treatment. The average of their *second school term grades* and *end-year exam scores* serve as a posttest measure.

The results of the descriptive analysis presented in *Table 1* show research participants' scores for the first and the second school term – the means (M) and the standard deviations (SD). The grades of students in the treatment condition increase from the pretest M = 5.09 (N = 25, SD = .45) to the posttest – M = 5.45 (SD = .34). As to the control group, the pretest and posttest mean scores are as follows: M = 5.04 (N = 25, SD = .41) and M = 5.11 (SD = .41), respectively.

Table 1. Descriptive statistics

		Report	
	TreatmEffect	FirstSchoolTerm	SecSchoolTerm
Cntrl	Mean	5,0351	5,1095
	N	25	25
	Std. Deviation	,41474	,40937
Exp	Mean	5,0851	5,4531
	N	25	25
	Std. Deviation	,44933	,34212
Total	Mean	5,0601	5,2813
	N	50	50
	Std. Deviation	,42869	,41174

Having inspected the descriptive statistics based on students' L2 achievement, as measured in terms of formal classroom assessment, I move on to determine whether the difference in the means is statistically significant at the .05 level.

In order to test the significance of the variation in the pre-test and post-test measures, I use *Univariate Analysis of Variance* as a test method. The main research hypothesis is that the average value of the dependent variable is not the same for the experimental and the control group. The null hypothesis is set as “the mean is the same for the two groups”. $H_0: \rho \geq 0.05$ $H_a: \rho < 0.05$. Before interpreting the ANOVA results, I will first check an important assumption underlying this statistical test – by using Levene’s test for equality of variance. The null hypothesis is that there is no difference between the variance of the experimental group and the variance of the control group: $H_0: \mu_1 = \mu_2$; $H_1: \mu_1 \neq \mu_2$; OR: $H_0: \mu_1 - \mu_2 = 0$; $H_1: \mu_1 - \mu_2 \neq 0$ (μ_1 and μ_2 are the population means for the two groups).

Table 2. Test of homogeneity of variances

Levene's Test of Equality of Error Variances ^a			
Dependent Variable: GainSchTERM			
F	df1	df2	Sig.
,015	1	48	,904
Tests the null hypothesis that the error variance of the dependent variable is equal across groups.			
a. Design: Intercept + TreatmEffect			

As shown in *Table 2*, the obtained p-value ($p = .904$) is higher than the significance level set ($p < 0.05$), which means that differences in sample variances are unlikely to have occurred based on random

sampling from a population with equal variances. Thus, in this case, there is not enough evidence to reject the null hypothesis of homogeneity of variances. It could be inferred that there is no statistically significant difference between the variances in the population. The assumption of homogeneity of variance has been met, i.e., a conclusion could be drawn that the variances are homogenous (or, approximately equal) across the two groups. Therefore, the results indicating a causal relationship may not have occurred because of random chance, perhaps a sampling error. In short, the result obtained is not a chance finding.

The interpretation of the results of the gain score analysis (*Table 3*) shows that the change in scores from the pretest to the posttest is greater for the experimental group than for the control group. Therefore, a conclusion could be drawn that it is the treatment that has produced the observed effect, and not other factors unaccounted for.

Table 3. Results of Univariate ANOVA

Tests of Between-Subjects Effects						
Dependent Variable: GainScores						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1,078 ^a	1	1,078	19,812	,000	,292
Intercept	2,446	1	2,446	44,983	,000	,484
Treatment Effect	1,078	1	1,078	19,812	,000	,292
Error	2,611	48	,054			
Total	6,135	50				
Corrected Total	3,688	49				
a. R Squared = ,292 (Adjusted R Squared = ,277)						

The results of the *F*-statistic reveal a significant influence $F(1, 48) = 19.812$, with a probability of chance occurrence $p < 0.001$ (*Table 3*). A conclusion could be drawn that the obtained *F*-ratio is not likely to occur by chance. The statistical significance attained by the outcome results is high. It furnishes evidence of the probability that the observed effect, i.e., the statistically significant difference between the two research groups, is the result of the treatment, and is not due to random error.

The effect size ($\eta^2 = .292$, $p < .05$) is large, which indicates that a large proportion of variance in the post-treatment data could be explained by the experimental conditions rather than by experimental error. The value of eta squared indicates a very strong association between the treatment and effect. Put otherwise, the change in the continuous dependent variable could be attributed to a considerable degree to the role of the examined predictor variable.

The results of the analysis of gain scores indicate that for the control group students, the gain in L2 achievement is .074 ($SD = .23$), while the average change between the first and second term for the participants from the experimental class is .368 ($SD = .24$). The changes from pretest to posttest are not identical across the research groups. Since the increase in school grades reflecting enhanced L2 achievement is higher in one group as compared to the other group, a case could be argued for a cause-and-effect relationship ($p < 0.001$). That is, the main hypothesis is confirmed by statistical evidence.

The mean, standard error, and 95% confidence interval for the mean are shown in *Table 4*. The calculation of the 95% gain score confidence intervals for the control and the experimental group indicates that the confidence interval (ranging from .155 to .288) does not overlap zero. This means that there is a significant improvement for the students who are subjected to the experimental treatment.

Table 4. Treatment Effect

Grand Mean			
Dependent Variable: GainScores			
Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
.221	.033	.155	.288

3.5. Summary

The results of formal classroom assessment have been analyzed in a univariate analysis of variance adopting a pretest – posttest design with *treatment effect* as the independent variable. I use ANOVA to partially adjust for preexisting differences among the groups. The results obtained indicate that the increase in L2 achievement is greater for participants in the treatment condition (M = .37, SD = .24) than for those in the control condition (M = .07, SD = .23). Inspection of the 95% confidence intervals around each mean indicates that there is a significant increase in L2 achievement for students in the treatment group, and no statistically significant increase for participants in the control group. This means that the experimental condition has a positive causal effect on L2 achievement as measured by standardized formative and summative school assessment.

The results of the conducted univariate ANOVA analysis (*Table 3*) indicate that **the research hypothesis** of causal impact **could not be rejected** ($p = 0.00$). The conclusion that can be drawn on the basis of statistical evidence is that applying a *dynamic assessment* approach **results in significant posttest gains**. The main implication made on the basis of the results yielded by this research study reiterates the crucial role of *meaning-focused interpersonal interaction* for L2 achievement.

4. Results and Discussion

In summary, this paper, focused on the evaluation of L2 development using a sociocultural theoretic lens, demonstrates the beneficial effect of integrating dynamic assessment in L2 classroom pedagogical practices. The results furnish evidence that the experimental group students significantly outperform the participants assigned to the control group ($p < 0.05$) on formal L2 classroom assessment which is the form of post-test measurement employed for the purposes of the quasi-experiment.

The statistical results substantiate the speculation that the expected increase in ESL achievement in the second school term is caused by the treatment variable. The findings supporting the hypothesized causal effect point to the advantage of incorporating dynamic assessment in L2 classroom evaluation – in terms of both measuring students’ communicative competence and promoting subsequent L2 development.

The approach adopted in this study is premised on the view that a *dynamic assessment* model allows for examining L2 testing and evaluation processes in the light of (re)constructing L2 students’ identity. The L2 learner is assigned the role of an autonomous individual, a personality construed as actively shaping the evaluation processes, and transformed from a “devoiced”, instrumentalized object of measuring to an agentive participant who has a “unique psychobiography”, to borrow the term from Norton (2013).

From this vertex, examinees are not conceived of as mere recipients of a language test’s input or producers of the required output being constrained and preconditioned by the content and structure of the elicitation instrument. Rather, within the sociocultural approach, students are construed as autonomous

communicators involved in a situated interaction, which encourages them to enter different discursive roles and play out various aspects of their personal and social identity. While participating in joint oral discussions preceding the writing activity itself, they are acting as authors, evaluators, critics, narrators and agents co-constructing discourse-creation processes by way of assisted, scaffolded performance.

The complex process, in which a learner's institutionally-imposed role and position are converted, has been explored in this study as unfolding in a local L2-related activity setting where students are engaged in negotiation of meaning, and in renegotiating and reconstituting their identities. Subject positions are construed as being mediated both on the individual level and on the inter-psychological level.

In this paper, I have attempted to provide a rationale for incorporating collaborative discourse creation into FL evaluation. 'Co-storying' applied as part of *dynamic assessment* is seen here as having the potential to establish a vibrant L2 performance environment. Engaging students in narrative processes constitutes an important dimension of tapping into their comprehensive ability to use the target language. It could be an effective form of gathering evidence to integrate collective meaning-and-identity-construction work into L2 classroom evaluation.

In suggesting a sociocultural approach to classroom-based evaluation research, I do not wish to advocate a novel model studied as a replacement for standard classroom processes but rather as an approach to both evaluating FL knowledge and instigating learning, which complements psychometric methods. I offer the *dynamic assessment* framework as a contribution to what has been hitherto achieved in studying L2 performance evaluation in terms of collaboration, dialogic behavior and identity expression.

The presented results are only tentative. The implications, of course, need to be interpreted with consideration given to the specific L2 assessment context, the students' particular needs and their reasons for studying the foreign language. There could be no universal panaceaic approach to testing and evaluation that is equally effective for all settings. In view of this, test designers and classroom practitioners should be aware of, and address, students' idiosyncratic context-contingent communicative needs. The pivotal role of the social context for evaluating L2 performance and exploring issues of identity is reaffirmed by the findings of this study. The results show that an individual-focused, contextualized approach to eliciting samples of TL use is significantly related to L2 achievement.

REFERENCES:

- Bakhtin, M. (1986)** The Problem of Speech Genres. – In: Speech Genres and Other Late Essays. Austin: University of Texas Press, p. 60 – 102.
- Brindley, G. (1994)** Competency-Based Assessment in Second Language Programs: Some Issues and Questions. // *Prospect*, 9 (2), p. 43 – 55.
- Daniels, H. (1996)** An Introduction to Vygotsky. New York. NY: Taylor and Francis.
- Dixon-Krauss, L. (1996)** Vygotsky in the Classroom. Mediated Literacy Instruction and Assessment. White Plains, NY: Longman Publishers.
- Donato, R. (1994)** Collective Scaffolding in Second Language Learning. – In: Lantolf, J.P. (Ed.). *Vygotskian Approaches to Second Language Research*. London: Ablex Publishing, p. 33 – 56.
- Dufva, H. (2003)** Beliefs in Dialogue: Bakhtinian View. – In: A. Maria Barcelos & P. Kalaja (Eds.). *Beliefs About SLA: New Research Approaches*, p. 13 – 152.
- Halliday, M. (1973)** Foreword. – In: D. Mackay, B. Thompson, & P. Schaub (Eds.). *Break-through to Literacy* (American ed.). Glendale, CA: Bowmar.
- Halliday, M. (1984)** Language as Code and Language as Behaviour: A Systemic Functional Interpretation of the Nature and Ontogenesis of Language. – In: R. Fawcett, M.A.K. Halliday, S.M. Lamb, & A. Malckai (Eds.), *The Semiotics of Culture and Language*, Vol. 1. London: Frances Pinter.
- Kozulin, A. & B. Z. Presseisen (1995)** Mediated Learning Experience and Psychological Tools: Vygotsky's and Feuerstein's Perspectives in a Study of Student Learning. // *Educational Psychologist*, 30(2), p. 67 – 75.

- Kozulin, A. (2002)** Sociocultural Theory and the Mediated Learning Experience. School Psychology International: Sage Publications.
- Kozulin, A. (Ed.). (2003)** Vygotsky's Educational Theory in Cultural Context. NY: Cambridge University Press.
- Lantolf, J. P. & W. Frawley (1988)** Proficiency: Understanding the Construct. // *Studies in Second Language Acquisition*, 10, pp. 181 – 195.
- Lantolf, J.P. & G. Appel (1994)** Theoretical Framework: an Introduction to Vygotskian Approaches to Second Language Research. – In: Lantolf, J. P. (Ed.). *Vygotskian Approaches to Second Language Research*. London: Ablex Publishing.
- Lantolf, J. P. (2000)** Sociocultural Theory and Second Language Learning. Oxford: Oxford University Press.
- Little, D. (2000)** Learner Autonomy and Human Interdependence: Some Theoretical and Practical Consequences of a Social-Interactive View of Cognition, Learning and Language. – In: B. Sinclair, I. McGrath & T. Lamb (Eds.). *Learner Autonomy, Teacher Autonomy: Future Directions*. Harlow: Longman/Pearson Education.
- Miller, R. (2011)** Vygotsky in Perspective. New York, NY: Cambridge University Press.
- Mitchell, R. & F. Myles (1998)** Second Language Learning Theories. London: Arnold.
- Moll, L. C. (1990)** Vygotsky and Education: Instructional Implications and Applications of Sociocultural Psychology. Cambridge, UK: Cambridge University.
- Norton, B. (1995)** Social Identity, Investment, and Language Learning. // *TESOL Quarterly*, 29, p. 9 – 32.
- Norton, B. (2013)** Identity and Language Learning: Extending the Conversation (2nd ed.). Bristol, UK: Multilingual Matters.
- Nunan, D. (2004)** Task-Based Language Teaching: A Comprehensively Revised Edition of Designing Tasks for the Communicative Classroom. Cambridge, UK: Cambridge University Press.
- Ohta, A. (2000)** Rethinking Interaction in SLA: Developmentally Appropriate Assistance in the Zone of Proximal Development and the Acquisition of L2 Grammar. – In: J. P. Lantolf (Ed.). *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.
- Poehner, M. E. & J. P. Lantolf (2003)** Dynamic Assessment of L2 Development: Bringing the Past into the Future. – In: CALPER Working Papers Series, No. 1. The Pennsylvania State University, Center for Advanced Language Proficiency, Education and Research.
- Poehner, M. E. & J. P. Lantolf (2005)** Dynamic Assessment in the Language Classroom. // *Language Teaching Research*, 9 (3), pp. 1 – 33.
- Poehner, M. E. (2015)** A Casebook of Dynamic Assessment in Foreign Language Education. University Park, PA: Center for Advanced Language Proficiency Education and Research, The Pennsylvania State University.
- Schunk, D. H. & B. J. Zimmerman (1997)** Social Origins of Self-Regulatory Competence. // *Educational Psychologist*, 32, pp. 195 – 208.
- Swain, M. & P. Deters (2007)** “New” Mainstream SLA Theory: Expanded and Enriched. // *The Modern Language Journal*, 91, pp. 820 – 836.
- Turuk, M. (2008)** The Relevance and Implications of Vygotsky's Sociocultural Theory in the Second Language Classroom. // *ARECLS*, Vol. 5, pp. 244 – 262.
- Vygotsky, L. S. (1978)** Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1981)** The Instrumental Method in Psychology. – In: J. Wertsch (Ed.). *The Concept of Activity in Soviet Psychology*. New York: Sharpe.
- Vygotsky, L. S. (1986)** Thought and Language. Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1997)** Educational Psychology. (V. V. Davydov, Introd.; R. Silverman, Trans.). Boca Raton: St. Lucie Press.
- Wertsch, J. V. (1985)** Vygotsky and the Social Formation of Mind. Cambridge, MA: Harvard University Press.