

ILDIKÓ CSÉPES

Self- and Peer-Assessment in Focus: Stakeholders' Perceptions and Issues of Implementation

Introduction

Classroom assessment practices or formative assessment as well as high-stakes testing or summative assessment can have far-reaching implications for individuals although the consequences of the latter are more widely discussed and acknowledged. It is often the case that performance on a single standardised test may result in admittance to or rejection from specific educational programmes, but the consequences of classroom assessment seem to be regarded less far-reaching, less tangible and therefore less serious. For instance, learners may lose their motivation for learning in the long run if classroom assessment is restricted or questionable in its scope and format compared to the instruction learners have received. While low-stakes and high-stakes assessments undoubtedly differ in terms of the size and nature of impact on the learner, we believe that both of them should be based on *democratic principles* as assessment cannot become only a tool for authority for power and control. Instead, as Shohamy (2001) puts it, assessment should entail pedagogical benefits that lead to more effective learning and teaching as well as *shared authority*. This goal can be realised, for example, by “conducting and administering testing in collaboration and in cooperation with those tested” (Shohamy, *ibid.* p. 378). Collaboration and cooperation imply mutual effort between testers/assessors and test-takers/assessees, thus resulting in shared authority or partnership for the outcomes of assessment (Chen, 2008; Stefani, 1998). Self- and peer-assessment is one of the means to realise these goals in assessment.

Involving students in the assessment process can also be argued for on the ground of *constructivist theories of education*, which propose that knowledge is actively constructed by learners (von Glasersfeld, 1989); it is created by constantly interacting with the environment through an active process of sense making. In a constructivist paradigm, classroom assessment addresses learning processes as well as learning outcomes, the expectations are visible to students and they are engaged actively in evaluating their own work (Shepard, 2000). Through reflection, peer discussion and feedback learners can engage in a dialectic process that leads to the evaluation of their learning performance in the form of self- and peer assessment. By identifying strengths and weaknesses learners are also informed about how they may improve, which makes assessment an integral part of teaching and learning. Furthermore, an understanding of the assessment process as well as the evaluation criteria enhances learners' capacity as self-monitoring, autonomous

learners. This approach to assessment is often referred to as *assessment for learning* to distinguish it from *assessment of learning*, which is carried out for summative purposes (Poehner & Ableeva, 2011).

The rationale for a constructivist and democratic assessment practice with a focus on self- and peer assessment appears to be convincing, nevertheless there seems to be little empirical research carried out on how widely the above views have been adopted in practice or what stakeholders' – learners' and teachers' – perceptions are concerning the use and impact of this kind of assessment. In this paper, we will review some empirical research on self- and peer- assessment in relation to stakeholders' perceptions, the effects of training on its implementation and its scoring validity in relation to more traditional forms of assessment. In this review, the research contexts range from foreign language classrooms to other disciplines and from primary schools to higher education contexts. Our discussion will focus on assessing students' oral performances primarily.

Advantages of self- and peer-assessment

Undoubtedly, there are several benefits entailed in self-assessment (SA) and peer-assessment (PA). These assumptions include the promotion of learning and the enhancement of learner independence, autonomy and life-long learning skills as students through active involvement in assessing their own learning performance are empowered by gaining ownership of their learning (Chen 2008; Oscarson, 1989; Patri, 2002). Furthermore, PA encourages reflective learning through observing others' performances and becoming aware of performance criteria (Falchikov, 1986 as cited in Saito, 2008). When SA is informed by peers, learners are likely to assess their own performances more effectively and as a result, learners' use of meta-cognitive strategies, their evaluative attitudes and goal-orientation are enhanced (cf. Boud, Cohen & Sampson, 1999). Reflection is also believed to raise learners' self awareness of learning strengths and weaknesses as well as their confidence skills (Oscarson, *ibid.*; Chen, 2008). While working with peers in the classroom, students develop a sense of shared responsibility (cf. Somervell, 1993 as cited in Saito, 2008), which may positively influence group dynamics and at the same time it lessens the assessment burden on the teacher (Oscarson, *ibid.*)

In the case of young learners, alternative forms of assessment rather than traditional ones (e.g. pen-and-paper tests) seem to better reflect curricular goals as both achievement and the developmental process can be evaluated with the help of on-going feedback and self-evaluation. Traditional testing emphasises discrimination because its main function is to rank order students along the same scale, allocating them to different categories and thus labelling them as "strong" and "weak", which necessarily leads to competition among them. Young learners, therefore, can benefit from SA and PA because of their collaborative, non-competitive nature.

The effectiveness of SA and PA seems to be dependent upon some contextual factors such as the transparency or clarity of the assessment criteria, whether learners are provided with training in SA and PA, and whether the guidance and feedback in the implementation phase is based on constructive information (Chen, 2006). In high-stakes test, clear assessment criteria and rater training are considered essential to ensure rater consistency, therefore similar measures are desirable to enhance the reliability of the assessments in SA and PA. However, L2 learners as raters are different from teacher as raters because their developing language proficiency influences their perceptions and ultimately their rating. Learners' developing competence suggests that their perception of the construct of language proficiency is constantly changing and as a result, SA and PA can only be effective if they are involved in self-directed assessment that is supported by constructive feedback.

The role of feedback

As has been suggested above, self- and peer-assessment should guide future learning and not only lead to judgements to prove that students have learnt. In SA and PA, the process (reflection and arriving at judgements) should be as important as the product, i.e. the ratings. As Liu and Carless put it, "Engaging learners in thinking about achieving outcomes to certain agreed standards is a learning process and giving marks or grades is only part of that process" (2006, p. 280). The former is captured by the term peer feedback, the latter by peer assessment, although these terms are used somewhat inconsistently in the literature. Peer feedback is a precursor for peer assessment and it has greater potential for learning than peer assessment as it helps learners to become actively engaged in articulating their evolving understandings of subject matter.

Kim (2009) goes even further to suggest that the assessee, who is usually just a passive receiver of peer feedback, should give back-feedback to the peer assessor. In this way, assessees can help assessors to reflect upon their own learning, thus making them more motivated to be assessors as they assess those who assessed them.

Stakeholders' perceptions of using SA and PA

Students and teachers have shown both positive and mixed attitudes toward the use of SA and PA. It is comforting to find acceptance on the part of some stakeholders as their favourable comments reinforce many of the assumed advantages discussed above. The negative perceptions, on the other hand, can highlight those contextual variables that can undermine the successful implementation of SA and PA.

Hanrahan and Isaacs (2001) examined self- and peer-assessment in a tertiary education context, where students were asked to self-assess and evaluate one of their peers' work on their own without having a group

discussion. Their findings revealed that SA and PA can develop empathy with tutors as students reported that they had gained insights into the difficulty tutors experience in assessing students. Students also reported an increased level of motivation for task completion as they wanted to impress their peers with their work.

Mok (2011) conducted a case study in Hong Kong to explore 12 to 14-year-old school students' perceptions of the implementation of peer assessment in their English speaking classes. Along with other researchers (Dickinson & Carver, 1980; Kolláth, 1996) Mok calls for the teacher's need to prepare students for PA both methodologically (by providing training and explaining the assessment criteria) and psychologically (by explaining the underlying principles of PA, raising students' comfort level in using PA through allowing the use of L1 for instance). In the classroom context described by Mok, such preparation was missing. The four students interviewed reported difficulty in evaluating or giving feedback on their peers' oral performances in the target language. Despite their acknowledged lack of ability in peer assessment, all of them perceived peer assessment positively as it enhanced critical thinking and setting learning targets.

It seems difficult to guarantee that students benefit from SA and PA as planned because there may be unforeseen psychological forces at work. Hanrahan and Isaacs (2001) report in their study that peers are sometimes too critical but at other times they find it uncomfortable to critique other people's work. Vickerman (2009), on the other hand, suggest that peer assessment does not work as a strategy for all students because students with different learning styles seem to benefit to differing degrees from peer assessment: there are students who require more direct rather than self-directed support in assessment.

Teachers' perceptions seem to be more mixed than those of learners as they can see potential gaps between theory and practice more directly. Bullock (2011) explored EFL teachers' perceptions in relation to self assessment and, although in general it was perceived positively, there were some problems identified in relation to its implementation. The respondents pointed out that self-assessment was time-consuming and teenage students did not take it seriously. This latter concern seems to be closely related to another difficulty in the implementation mentioned by the respondents: helping learners to formulate appropriate aims for future learning. Bullock argues that if self-assessment fails to lead to choice and opportunities for self-direction, then learners' intrinsic motivation may be decreased. Despite teachers' overall positive attitudes to self-assessments, Bullock also remarks that they are not necessarily indicative of teachers' actual or future practices.

Liu and Carless (2006) collected data from a large-scale questionnaire survey on peer assessment beliefs and experiences of academics and students in tertiary education contexts in Hong Kong. They found reluctance on the part

of academics and students as well. They identified four main reasons for resistance to peer assessment processes using grades: (1) peers are thought to be less reliable assessors than academics; (2) academics are considered to have expertise in assessment (some classmates are unable to provide insightful feedback); (3) power relations are disrupted: academics may resist sharing their power, students dislike having power over peers, they may feel resentful because they have to surrender some power and control over their own work to peers; (4) more time-consuming than traditional assessment.

The perceived imbalance in power relations reported above can be explained by the sociocultural milieu of education and the social norms and roles in assessment. Hamp-Lyons (2007 as cited in Butler & Lee, 2010) proposes that there are two conflicting cultures of assessment: *a learning culture* and *an exam culture*. The former emphasises the learner's progress in learning whereas the latter recognises mastery or achievement in relation to group norms. The teaching and learning context can have an important impact on self- and peer-assessment depending on which assessment culture is dominant in the given context. To support this assumption, Butler and Lee (ibid.) found that the teachers' and students' perceived effectiveness of self-assessment was quite different depending on the teaching/learning contexts and on the role of assessment individual teachers themselves subscribed to.

Issues of validity in SA and PA

It is important to understand how the validity of self- and peer-assessments may be affected by specific contextual variables. Drawing on previous research, Butler and Lee (2010) identify three main sources of variability in self- and peer-assessments: (1) the ways in which questions and items are formulated and delivered; (2) the domain or skill being assessed; (3) learners' individual characteristics such as age, proficiency level, previous language learning experience, etc. In the next section we will discuss the effects of variability of some factors related to all the three sources mentioned above, focusing on the scoring validity of SA and PA compared to more traditional forms of assessment (proficiency tests or teacher assessments).

In relation to the first source of variability, Heilenman (1990) investigated response effects in self-report measures of second language abilities. The self-report measure focused on grammar, vocabulary, accuracy and fluency and consisted of 65 items to be rated on a five-point scale (from strongly agree to strongly disagree). The response effects the study focused on were acquiescence (a tendency to respond positively) and self-presentation (to present oneself in a favourable light, consciously or unconsciously inflating reality). In order to investigate the former, matched pairs of positive and negative statements were used. The effect of overestimation, on the other hand, was explored through statements that were assumed to represent different difficulty levels in relation to the ability level of the target group. Results show

that students find it more difficult to agree with a negative statement and a tendency toward rating high was also confirmed. The author recommends the exclusion of negatively worded questions in self-report measures.

In relation to the scoring validity of SA, Blanche (1990) reported both high and low agreement between self-ratings and teacher assessments on standardized achievement and oral proficiency tests. Similarly, the comparison of teacher assessment with peer assessment revealed mixed results, too (Patri, 2002). Clearly, some of this inconsistency may be attributed to ‘friendship marking’ (Pond, Ul-Haq & Wade, 1995 as cited in Patri, 2002), i.e. over-marking by peers and ‘decibel marking’ (ibid), i.e. when loud or extraverted individuals get high marks because they dominate a group. However, the domain or skill learners have to assess also seems to influence considerably the reliability and validity of SA and PA. Learners were reported to have more difficulty in assessing inferential dimensions, such as appropriacy and fluency, than performance dimensions (Shore, Shore, & Thornton, 1992 as cited in Patri, 2002).

The variability resulting from the skill being assessed was examined by Ross (1998), who explored the relative strengths and weaknesses of SA across skill areas (reading, speaking, writing, listening), based on a statistical analysis of 60 correlations reported in relation to self-assessments. He found variation in self-assessment validity coefficients, which suggests that accurate interpretation of the ratings is not guaranteed. When the skills were looked at separately, however, he found variation across the self-assessments of individual skills. The correlations between self-assessments and reading were found to be the strongest and therefore relatively more valid than in the case of listening skills, where the correlations showed greater variations. The reason for this, as suggested by Ross, is that the given learners probably had less frequent exposure to those skills in the higher education contexts where the studies were conducted. In the case of both speaking and writing skills, relatively lower correlations were found between self-assessments and the criterion measures. Ross concluded that “... self-assessment introduces a systematic method facet even if it has some partial concurrent validity with important criterion variables” (ibid., p. 12). He suggested that in terms of accuracy of the self-ratings, self-assessment may be appropriate for use in low-stakes decision making only. Ross also explored whether there was a difference in the accuracy of achievement-based versus general proficiency-based self-assessments. In his own study (Ross, ibid.), he found evidence in support of the assumption that learners can estimate their achievement that is directly linked to practice in a specific course more accurately than achievement that is expressed in more abstract terms and can only indirectly be linked to coursework. Ross proposed that to ensure accuracy of self-assessment of language learning achievement, the instrument should be designed on the

basis of specific curricular content rather than general ability description in relation to the criterion skill.

Several studies have focused on specific individual learner characteristics to examine how they influence the scoring validity of SA and PA. Among the variables investigated, the age and proficiency level of the learner as well as the previous learning experience, more specifically the presence or absence of training in SA and PA seem to emerge as the most important ones. Butler and Lee (2010), for example, correlated young learners' self-ratings and external proficiency ratings. They found that over time learners' accuracy improved, which was taken as evidence in support of the contention that even young learners can be trained to self-assess. Most of the studies, however, focused on more mature and proficient learners, pursuing studies in tertiary education. In Hong Kong, Cheng and Warren (2005) compared teacher assessments to how undergraduates assessed their peers' spoken language proficiency. Although the student and teacher assessments were not widely different and by the end of the experiment they even got reasonably close, when the peers awarded marks, they actually reported that they had failed to include all the elements of language proficiency that the teachers assessed. For example, students associated oral proficiency more with oral fluency. This is, in fact, not so surprising as some of the students admitted that they were unsure to make peer assessments as they did not regard themselves competent for the task either linguistically or professionally. This finding reinforces the need for learner training, which may not only increase the reliability of peer assessments but, as Cheng and Warren also point out, it can enhance learners' confidence in their own ability to carry out such tasks and thus increase their positive attitude towards the peer assessment exercise itself.

Patri (2002) investigated the agreement amongst teacher-, self- and peer-assessments of students in the presence of peer feedback while testing oral presentation skills. 56 Chinese university students, divided into control and experimental groups, participated in her study, in which participants in the former group practised the tasks of self- and peer-assessment for four weeks, complemented with peer feedback while the latter did so without any peer feedback. Performances were assessed for organization of content, language use, manner and interaction with the audience. These criteria were all included in a questionnaire, which was used for teacher-, peer- and self assessments with minimal modifications. The behaviour of peer assessment was found to be somewhat different from that of self assessment. When rating their own performances individually, participants were not able to judge them in the same way as the teacher had done either in the presence or absence of peer feedback, but in the experimental group peer-assessment was in high agreement ($r=0.85$, $p\leq 0.005$) with the teacher-assessment when there was peer feedback. As a result of peer discussion, students were able to make

judgements of their peers' oral presentations comparable to those made by the teacher. Patri's conclusion suggests that peer feedback can enhance learners' ability to make judgements on their peers' oral presentation skills therefore peer feedback is essential in order to achieve scoring validity in relation to teacher judgments.

Chen (2008) also compared student assessment with teacher assessment in order to explore the validity of self-ratings of oral performance in English. In this case study, 28 Chinese university students went through a 12 week instruction period that included both training and a 10-week practice phase in peer and self-assessment. The assessment components were collaboratively developed by the teacher and the students including content (30%), language (30%), delivery (30%), and manner (10%). Students were asked to assess their own performance using the same criteria as those used for peer and teacher assessment and all three forms were conducted jointly. There was a significant match between self- and teacher scorings towards the end of assessment practice, i.e. there was improvement in the assessment accuracy. Chen noted that the students and the teacher emphasized different aspects of the established criteria as the former commented on delivery and manner most of the time while the latter commented more on aspects of content, language, and delivery. As time passed by, learners' negative comments decreased and more positive and neutral comments were offered. By the end of the experiment, the learners' comments were more similar to the teacher's, becoming more positive and constructive as they began to notice their own strengths. Giving and receiving feedback in class for ten weeks thus helped students to change the negative tone of their comments to a reflective and suggestive one. Chen concluded that this assessment procedure involved two simultaneous processes: learning to assess and assessing to learn.

The impact of training on peer assessment of L2 oral presentation was explored by Saito (2008) in a Japanese university context, using treatment and control groups. Both groups received a series of instructional inputs on 12 skill aspects of presentation but only the treatment group was offered rater training, which involved viewing and rating three videotaped presentations of former students. Ratings were discussed and feedback was given by pointing out over- and under-rating. Later on the course, students in both the treatment and control groups rated and commented on all classmates' performances in their own class. Saito found that "instruction on skill aspects without rater training may be sufficient for peer assessment to correlate with instructor scores to a certain degree; however, rater training seems to help students provide more relevant comments, citing more skill aspects on peer performance" (*ibid.*, p. 575). On its own, instruction on skill aspects that learners can rate is not likely to increase correlations between learner and teacher ratings. With the help of rater training, however, learners' awareness of skill aspects can be enhanced, which is reflected in the frequency and relevance of the comments learners

give. Saito suggests that “peer assessment training is a meta-cognitive activity in which student attention is drawn to the features of a language learning task” (ibid., p. 577).

In summary, the findings reported above have emphasised the need for learner or rater training that involves awareness-raising of skills aspects of performance and practice in applying specific assessment criteria. By creating opportunities for interaction and clarification instructor feedback and peer discussion both seem to have an important role to play in enhancing the scoring validity of SA and PA, although in contexts that can be characterised to have an *exam culture* (cf. Hamp-Lyons, 2007 as cited in Butler & Lee, 2010), stakeholders' perceptions of the role of assessment may undermine both the implementation as well as the outcomes of self- and peer-assessment as neither learning to assess nor assessing to learn is recognised as curricular goals for learners.

Conclusion

As has been shown above, there are a lot of benefits self- and peer-assessment entails, although we have highlighted some potential problems as well. We agree with Boud (2000 as cited in Liu & Carless, 2006), who recommends the creation of a course climate in which the giving and receiving of peer feedback is a normal part of teaching and learning processes. Whether Boud's call has been answered or not, remains to be seen. In Poland, a recent survey by Lewkowicz and Zawadowska-Kittel (2011) has revealed that in the Polish state primary and secondary school context there is a low level of popularity of alternative assessment tasks in the English language classroom as the teachers reported using peer- and self-assessment in mostly small percentages. At the primary school level, self-assessment was reported by 44% while peer assessment by 11%. At the lower secondary level these percentages changed for 31% and 16%, while at the upper secondary level 31% and 25% of the teachers reported to use these alternative forms of assessment respectively. In Hungary, where the teaching and learning culture is very similar to the Polish one, little is known about the use of alternative forms of assessment in the Hungarian state school context. Future research needs to address these issues as time has come to explore Hungarian teachers' perceptions and classroom practices in relation to self- and peer-assessment in English language learning too.

The research reported here has been supported, in part, by the TÁMOP [4.2.1/B-09/1/KONV-2010-0007](#) project, which is implemented through the New Hungary Development Plan, co-financed by the European Social Fund and the European Regional Development Fund.

References

- Blanche, P. (1990). Using standardized achievement and oral proficiency tests for self-assessment purposes: the DLIFLC study. *Language Testing*, 7(2), 202–229.
- Boud, D. (2000) Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167. <http://www.tandfonline.com/doi/abs/10.1080/713695728>
- Boud, D., Cohen, R. & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education*, 24(4), 413–426. <http://www.tandfonline.com/doi/abs/10.1080/0260293990240405>
- Bullock, D. (2011). Learner self-assessment: an investigation into teachers' beliefs. *ELT Journal*, 65(2), 114–125. <http://eltj.oxfordjournals.org/content/65/2/114.abstract?sid=66867fea-96f8-4cec-a8d0-1a7c509657ea>
- Butler, Y.G. & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27(1), 5–31.
- Chen, Y.-M. (2006). Peer and self-assessment for English oral performance: A study of reliability and learning benefits. *English Teaching and Learning*, 30(4), 1–22.
- Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235–262. <http://ltr.sagepub.com/content/12/2/235.abstract>
- Cheng, W. & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93–121. <http://ltj.sagepub.com/content/22/1/93.abstract>
- Dickinson, L. & Carver, D. (1980). Learning how to learn: steps towards self-direction in foreign language learning in schools. *ELT Journal*, 35(1), 1–7. doi:10.1093/elt/XXXV.1.1
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment & Evaluation in Higher Education*, 11, 146–165. <http://www.tandfonline.com/doi/abs/10.1080/0260293860110206>
- von Glasersfeld, E. (1989). Constructivism in education. In T. Husen & N. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies. Supplementary Volume* (pp. 162–163). Oxford: Pergamon Press. Retrieved from <http://www.vonglasersfeld.com/114>
- Hamp-Lyons, L. (2007). The impact of testing practices on teaching: Ideologies and alternatives. In J. Cummins & C. Davison (Eds.), *The International handbook of English language teaching*, Vol. 1. (pp. 487–504). Norwell, MA: Springer.
- Hanrahan, S. J. & Isaacs, G. (2001). Assessing Self- and Peer-assessment: the students' views. *Higher Education Research & Development*, 20(1), 53–70. <http://www.ingentaconnect.com/content/routledg/cher/2001/00000020/00000001/art00005>
- Heilenman, L. K. (1990). Self-assessment of second language ability: the role of response effects. *Language Testing*, 7(2), 174–201. <http://ltj.sagepub.com/content/7/2/174.abstract>

- Kim, M. (2009). The impact of an elaborated assessee's role in peer assessment. *Assessment & Evaluation in Higher Education*, 34(1), 105–114. <http://www.tandfonline.com/doi/abs/10.1080/02602930801955960>
- Kolláth, K. (1996). The importance of learner training as a tool to develop autonomous learners. *Külkereskedelmi főiskolai füzetek* 6. sz. Retrieved from http://elib.kkf.hu/okt_publ/szf_06_33.pdf
- Lewkowicz, J. & Zawadowska-Kittel, E. (2011). Classroom-based Assessment: the Polish Experience. In D. Tsagari & I. Csépes (Eds.), *Classroom-Based Language Assessment* (pp. 147–167). Frankfurt am Main: Peter Lang GmbH.
- Liu, N. & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279–290. <http://web.edu.hku.hk/staff/dcarless/Liu&Carless2006.pdf>
- Mok, J. (2011). A case study of students' perceptions of peer assessment in Hong Kong. *ELT Journal*, 65(3), 230–239. <http://eltj.oxfordjournals.org/content/65/3/230.abstract?sid=adf2bda8-fc66-4fff-8458-6d61ca82ed34>
- Oscarson, M. (1989). Self-assessment of language proficiency: rationale and applications. *Language Testing*, 6(1), 1–13. <http://ltj.sagepub.com/content/6/1/1.abstract>
- Patri, M. (2002). The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing*, 19(2), 109–131. <http://ltj.sagepub.com/content/19/2/109.abstract>
- Poehner, M. E. & Ableeva, R. (2011). Dynamic Assessment and Learner Engagement in the Activity of Development. In D. Tsagari & I. Csépes, (Eds.), *Classroom-Based Language Assessment* (pp. 15–28). Frankfurt am Main: Peter Lang GmbH.
- Pond, K., Ul-Haq, R. & Wade, W. (1995). Peer review: a precursor to peer assessment. *Innovation in Education and Training International*, 32, 314–23.
- Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1–20. <http://ltj.sagepub.com/content/15/1/1.abstract>
- Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25(4), 553–581. <http://ltj.sagepub.com/content/25/4/553.abstract>
- Shepard, L., A. (2000). *The role of classroom assessment in teaching and learning*. CSE Technical Report 517. CRESST, University of Colorado: Boulder. Retrieved from http://datause.cse.ucla.edu/DOCS/las_rol_2000.pdf
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18 (4), 373–391. <http://ltj.sagepub.com/content/18/4/373.abstract>
- Shore, T. H., Shore, L. M. & Thornton III, G. C. (1992). Construct validity of self and peer evaluations of performance dimensions in an assessment center. *Journal of Applied Psychology*, 77, 42–54. Retrieved from http://www-rohan.sdsu.edu/~lshore/reprints_pdf/construct_validity_of_self.pdf

Somervell, H. (1993). Issues in assessment, enterprise and higher education: The case for self-, peer and collaborative assessment. *Assessment & Evaluation in Higher Education*, 18, 221–233. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/0260293930180306>

Stefani, L. J. (1998). Assessment in partnership with learners. *Assessment & Evaluation in Higher Education*, 23(4), 339–350.

Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34(2), 221–230. Retrieved from <http://www.mendeley.com/research/student-perspectives-formative-peer-assessment-attempt-deepen-learning-1/>

To cite this Paper:

PRINTED: Ildikó Csépes: Self- and Peer-Assessment in Focus: Stakeholders' Perceptions and Issues of Implementation. In Gabriella Pusztai, Zoltán Tóth & Ildikó Csépes (Eds.) *Current Research in the Field of Disciplinary Didactics*. Hungarian Educational Research Journal (HERJ) Special Issues, 2. pp. 63–74. Budapest (Hungary): HERA.

ONLINE: Ildikó Csépes: Self- and Peer-Assessment in Focus: Stakeholders' Perceptions and Issues of Implementation. In Gabriella Pusztai, Zoltán Tóth & Ildikó Csépes (Eds.) *Current Research in the Field of Disciplinary Didactics*. Hungarian Educational Research Journal (HERJ) Special Issues, 2. doi: 10.5911/HERJ.SPECIALISSUES2.5.
