

## Critical Review

# Perspectives on the Cultural Appropriateness of Assessment for Learning in the Chinese Context

by Ying Zhan ([zhanyinghku@gmail.com](mailto:zhanyinghku@gmail.com)) and Zhihong Wan ([zhwanhku@gmail.com](mailto:zhwanhku@gmail.com))

**Abstract:** *Assessment for learning (AfL) has been advocated by the government and educational agencies in Mainland China since the very beginning of the 21st century. However, its practice in Chinese classrooms is quite weak. Most teachers think that the practice of AfL is unrealistic in the present educational context, where Chinese examination-oriented culture is firmly embedded. The principal focus of this paper is on the cultural appropriateness of the implementation of AfL in Mainland China. The authors first elaborate the features of Chinese examination culture and its impact on teaching and learning and then link Chinese examination culture with AfL to identify the tensions between them. The tensions between AfL and Chinese examination culture include personal development versus competition with others, learning to learn versus rote learning and learning needs versus examination requirements. Finally the authors suggested some appropriate ways to promote the culture of AfL in Mainland China.*

## Introduction

Assessment for learning (AfL), usually known as formative assessment, refers to “frequent, interactive assessments of student progress and understanding to identify learning needs and adjust teaching appropriately” (Centre for Educational Research and Innovation [CERI], 2005, p 21). Black and Wiliam (1998a) reviewed a very wide range of published research on formative assessment up to 1988. Based on the evidence provided by the review, they argued that formative assessment raised levels of student achievement but the current practice of formative assessment was quite weak in classrooms in the countries which the review covered. Some researchers have explored the possible factors that affect the implementation of formative assessment. These factors range from the policies, politics and culture at the macro level to the teachers’ knowledge and beliefs at the micro level (Black and Wiliam, 2005; Carless, 2005). Therefore, if formative assessment is to have the desired effect, it must be ensured that what is implemented is instructionally sound as well as contextually appropriate.

In Mainland China, AfL has been advocated by the government and educational agencies over the last decade in order to promote students’ learning. For instance, at tertiary phase, the new College English Curriculum Requirements emphasise that formative assessment is particularly important in computer-based teaching, which is characterised by students’ independent learning (The Ministry of Education of PRC [MEPRC], 2004). However, most college English teachers in Mainland China only pay lip service to formative assessment and consider that its practice is unrealistic in the present educational context (Jin, 2008). Chinese examination-oriented culture, which is firmly embedded in Mainland China, makes teachers emphasise students’ marks or grades rather than their learning quality and learning needs (Han, 1997; Liu and Hu, 2005). As argued by Carless (2005), assessment culture seems to be particularly impervious to transformation. Therefore, understanding the tensions between Chinese examination culture and AfL is very important for the adoption of formative assessment in Chinese classrooms.

The principal focus of this paper is the cultural appropriateness of the implementation of AfL in Mainland China. The features of Chinese examination culture will be first elaborated on

and then be linked with AfL, with the aim of identifying the tensions between them. Finally, some measures will be discussed for the effective implementation of AfL in the Chinese context.

## **Chinese Examination Culture: Yesterday and Today**

China was the first country in the world to introduce Civil Service Examinations to select the highest officials of the land (Hu, 1984; Lai, 1970). The history of Civil Service Examinations can be traced back to the Han dynasty (201 BCE to 8 CE). The system of the examinations was elaborated and consolidated in the Tang dynasty (618 to 905 CE) and underwent continuous modification through the Song dynasty to the Qing dynasty before finally being abolished at the end of the nineteenth century (Lai, 1970).

Participation in the Civil Service Examinations was open to almost all civilians – poor or rich. There were five stages of examinations: Xian and Fu examination (a country examination), student selection examination, provincial examination, metropolitan examination and court examination. Candidates passing each of these examinations were titled Xiucai (cultured talent or budding genius), Juren (exalted man), Gongshi (candidates for court examination), and Jinshi (finished scholar). To avoid malpractice, all essays were marked anonymously, and the Emperor personally supervised the final stage. What was measured in the examinations was a set of ethical principles which were derived from Confucianism. Therefore, the representative work of Confucianism – the Four Books and Five Classics – must be understood and memorised by the examination candidates (Turner and Acker, 2002). The essays tested in the examinations, commonly known as Ba-gu wen or eight-legged essays, had a highly formalised form.

Civil Service Examination was seen as the main vehicle for social mobility. Therefore, competition among candidates was highlighted. Thus the higher the stage of examination, the fiercer the competition. To get a civil service post through the examinations became the dominant external motivation for people to be educated. Civil Service Examination then became the syllabus for study by candidates. Thus it might be said that “the examination tail wagged the education dog” (Hu, 1984, p 8).

In China, examinations have existed for two thousand years. As time passed, a characteristic examination culture formed and sunk its roots deep in the hearts of the people. Cheng (2004, p 10) vividly described this examination culture as follows:

Reading books, taking exams ... this is a straightforward road without any misunderstanding: learning the Four Books and Five Classics by heart and writing good articles. This is a cost-effective road: learning a household material by oneself. This is a fair road: no discrimination of the family background and economic conditions. This is a clean road: no examination cheating and bribe. However, this road should be built by suffering all kinds of hardship ... This channel for social mobility is widely accepted and known to all. It goes without saying that, therefore it has become a kind of culture (authors' translation).

Although Civil Service Examination has been abolished for one hundred years, it still exerts strong influence on the current assessment culture in China. Education in China is a highly selective enterprise since there is a huge testing population and limited places at each level of education (Eckstein and Noah, 1993). At the same time, Chinese parents eagerly want their children to have a bright future and become a “dragon” (Zeng, 1999). Examination results are thus seen as the way into the top streams and the top schools in the country. Distinguishing one child from another seems to be the only purpose of education. In order to meet that end, parents, students and teachers are all drawn into the competition. Parents put

more pressure on schools and teachers to prepare their children for the examination and pay to send their children to the tutorial schools after classes. Teachers become accountable for their students' performance and overemphasise students' marks or grades in their teaching. Students become more reliant on their teachers' and tutors' notes and model answers that they believe are helpful in increasing their marks or bettering their grades. They have a strong motivation to learn, but it does not mean that most of them are interested in learning itself. As argued by Harlen and Deakin Crick (2002), students have strong extrinsic orientation towards grades and social status in an education system which puts great emphasis on evaluation and selectivity.

High-stakes examination dominates educational curricula in Mainland China. What goes on in the classroom is largely dictated by what happens in the examination hall. It has been argued that examinations drive teachers to narrow curricular offerings and modes of instruction, and neglect the content and materials which are incompatible with the examination (eg Han, 1997; Yan, 1997; Zhang, 2005). Because most of these high-stakes examinations are external, norm-referenced assessment and based on uniform curricula for all, they highlight reliability and fairness. Thus it is not easy to assess high order thinking in these examinations. The high-stakes tests largely assess students' ability to recall knowledge and facts, which encourages students to learn by rote and discourages students' creativity and originality (Liu and Hu, 2005).

## **Principles of AfL**

The basic argument of AfL is that assessment would frequently provide information on children's weaknesses and strengths in relation to their progression and help teachers identify students' learning needs and adjust teaching appropriately (Torrance, 1993). Such assessment looks forward rather than backwards. Therefore, it becomes a powerful intervention to improve learning, as claimed by Black and Wiliam (1998a, p 61):

... formative assessment does improve learning. The gains in achievement appear to be quite considerable, and as noted earlier, among the largest ever reported for educational interventions as an illustration of just how big these gains are, an effect size of 0.7, if it could be achieved on a nationwide scale, would be equivalent to raising the mathematics attainment score of an 'average' country like England, New Zealand or the United States into the 'top' five after the Pacific Rim Countries of Singapore, Korea, Japan and Hong Kong.

Black and Wiliam (1998b) argue that improved formative assessment causes the rise in standards because it helps create greater equity of student outcomes. They think that "improved formative assessment helps low achievers more than other students and so reduces the range of achievement, while raising achievement overall" (p 141).

The concept of AfL is not new. It is underpinned by a strong belief that every student is unique and possesses the ability to learn, and thus the teacher should develop their multiple agencies and potentials. To promote learning, assessment should be regarded as an integral part of the curriculum, learning, teaching and feedback cycle. The curriculum sets out the learning targets and thus assessment should be a process in which learning evidence is collected based on the curriculum requirements.

Different modes of assessment are to be employed whenever suitable for a comprehensive understanding of students' learning in various aspects (Broadfoot, 1995; Cizek, 1997). For example, teachers' efficient feedback to students especially to low achievers can be a powerful weapon. Black, Harrison, Lee, Marshall and Wiliam (2003) suggest that feedback to any pupil should focus on the quality of his or her work with advice on how he or she can

improve it and should avoid comparisons with other pupils. Peer and self-assessment are other “essential component[s] of formative assessment” (Black and William, 1998b, p 143). In peer and self assessment, students actively build their understanding of new concepts and develop their meta-cognition rather than merely absorb information. In short, AfL should collect information about why the students are not attaining learning targets and how to help them to improve rather than focusing on what they have learnt. It can therefore be concluded that AfL emphasises personal development, cares about learning diagnosis and promotes learning.

## **Tensions between AfL and Chinese Examination Culture**

AfL offers a powerful means for meeting goals for high-performance, high-equity of students' outcomes, and for providing students with knowledge and skills for lifelong learning. Its strengths seem to be obvious but its practice is still quite weak in classrooms in different contexts, including Mainland China. There are some difficulties for the implementation of AfL such as less effective policy, large class sizes, limited instruction time, heavy workloads of teachers and students, and teachers' poor knowledge of AfL (Wang, 2007). More importantly, AfL is in conflict with the deep-rooted Chinese examination culture, which directly affects teachers' and students' beliefs and indirectly influences their behaviours.

### ***Personal Development Versus Competition with Others***

AfL focuses on learners' personal development rather than competition with others. The progress in learning is embodied in the individual difference between past and present. The teacher gives feedback to inform the students well about their strengths and weaknesses and not to damage students' self-esteem, thus encouraging students to improve their learning (Black *et al*, 2003). How does the purpose of AfL relate to the existing examination culture in Mainland China?

In traditional Chinese examination culture, the purpose of assessment was to select the excellent people to do an administrative job. Learning and social mobility were tightly buckled up and the criteria of selection motivated the scholars at that time. Competition among the scholars was extremely fierce and as Lai (1970, p 12) indicated, “the competition in these examinations was so exacting that success went only to the highly gifted. The percentage of success was normally about 5%”. Today, the degree of competition is not reduced due to limited educational resources. Teachers try to help their students to perform better than others in high-stakes examinations. They usually like ranking students after evaluating their students' learning performance. The ranking system conveys the message about which position the pupil is in by comparing with others, which is the first concern of parents, teachers and students. However, feedback by comments or marks in AfL cannot give this kind of message. If teachers only give comments to students and tell them what kind of learning targets they are not attaining, they seem to be easily questioned by students and their parents. This is because parents cannot predict their children's performance in high-stakes examinations and students may be at a loss about how well they do in comparison with others. Overemphasising competition with others, to some extent, jeopardises personal development. Under the pressure of competition, personal development seems to be a secondary objective of learning.

### ***Learning to Learn Versus Rote Learning***

AfL highlights building up learning to learn skills by peer and self-assessment since this promotes students' lifelong learning. Peer and self-assessment emphasise the process of learning and actively involving students in that process. In order to make their learning more efficient, students develop their meta-cognition by reflecting on their own work and judging

others' work. When students judge the quality of their own and their peers' work, they become aware of the desired learning goals and criteria and know well their present position. More importantly, they understand the way to close the gap between the desired standard and their actual standard (Sadler, 1989). This tenet of AfL seems to be in conflict with Chinese examination culture.

Civil Service Examination succeeded in selecting people whose chief faculty was extraordinary memory rather than good administrative ability (Lai, 1970). These scholars worked hard to rote memorise the Four books and Five Classics and imitated the established texts in examination. This tradition still exists in today's Chinese classrooms. Students seem to rely on their teachers to provide their model and right answers and then learn them by rote even without real understanding. Students are therefore not actively, but passively, involved in the teaching and learning process. Asking students to reflect on their learning and build up learning to learn skills is therefore uncomfortable and even threatening to many students.

### ***Learning Needs Versus Examination Requirements***

AfL posits that teachers should offer special programmes for individual students based on the results of formative assessment. This would be realised by "relatively unstructured interaction between student and student or teacher and student rather than a planned formal assessment event" (Crooks, 2002, p 241). The unstructured interaction might be questioning in class, talk after class, and so on. Through these modes, teachers know well what students need and how to meet their needs. AfL, as described above, appears to be different from the traditional fairness model in China.

Chinese examination culture is based on fairness. The Four Books and Five Classics were the sacred books in the eyes of scholars because they were the major tested content in Civil Service Examination. The scholars whose essays in the exam did not conform to Confucianism would fail. Today, the examination requirements in high-stakes examinations are the unified learning objectives for all the students, but these do not take account of the individual differences in terms of learning ability and learning style. When teachers choose their teaching tasks and homework for their students, they often keep an eye on test requirements and ignore learning needs, which are different from test requirements (Yan, 1997). Low achievers, who cannot meet test requirements, are neglected or blamed by their teachers and further lose opportunities to convey their learning needs in class. Therefore, teachers only pay lip service to the ideal that every child can be educated and make progress (Yan, 1997).

### **Implications for AfL Practice in the Chinese Context**

As argued by Cheng (1998, p 26), "it is rare that policies that run counter to the society's cultural norms will succeed in changing educational practice". If we apply this line of thinking to AfL, the implementation of AfL seems to be problematic in Mainland China where Chinese examination culture dominates the educational system. There are some tensions between AfL and Chinese examination culture such as personal development versus competition with others, learning to learn versus rote learning and learning needs versus examination requirements.

This does not necessarily mean that AfL is impossible to implement in Mainland China. It is still possible to implement AfL in Mainland China in some form. The most important point for the implementation of AfL is how to tailor it to the needs of Chinese examination culture at the macro level as well as the realities of local classrooms at the micro level.



As suggested by Black *et al* (2003, p 3), the changes can be made step by step – a big “leap in the dark” is not necessary. This suggestion is very useful when we consider putting AfL into Chinese classrooms. Chinese classrooms are dominated by summative assessment, which is consistent with Chinese examination culture. Introducing AfL into Chinese classrooms may be confronted with strong resistance from teachers and students. Therefore we should make AfL less radical and more pragmatic in the process of implementation.

In fact, summative assessment and formative assessment are the same in nature but for different purposes. An assessment “would be called formative if it were used to help learning and teaching, or summative if it were not so utilised but only employed for recording and reporting” (Harlen, 2005, p 208). Therefore, continuous summative assessment can be regarded as the weak form of formative assessment. Since Chinese teachers and students suffer from the pressures exerted by high-stakes examinations, it is impossible to reduce their attention on summative assessment. Under this circumstance, to graft formative value onto summative procedures may be more acceptable for Chinese teachers and students.

Black *et al* (2003) introduce some pragmatic ways of using classroom tests, beyond just assessing attainment, to develop students’ understanding. For example, teachers ask students to review their work and screen past test questions to identify areas of insecure understanding. This reflection on their areas of weakness enables students to focus on their revision. Teachers can also ask students to design test questions and devise marking schemes. This helps them “both to understand the assessment process and to focus further efforts for improvement” (p 54). Besides, teachers can use the outcome of tests diagnostically. They may look for which questions are poorly answered by the majority of students and focus on amending the learning associated with those answers. Teachers can also involve students in marking each others’ tests. This encourages students to “apply criteria to help them understand how their test answers might be improved” by peer-assessment (p 55).

There are also other means of conducting formative assessment. This includes questioning, dialogue with the teacher after class, classroom observations and homework. The crucial point is that students should be provided with opportunities to express their ideas. The teacher should respond to and reorientate students’ thinking and further encourage them to reflect on their learning (Black and Wiliam, 1998b).

The above discussion hints that AfL needs to experience a process of mediation in which it is made more suitable for the Chinese context. Creative and reflective teachers have played an important role in this process. However, most Chinese teachers are confronted with challenges when reconciling the formative role and summative role of assessment. Their belief, influenced by the Chinese examination culture, often impedes their AfL practice. Teacher professional development therefore becomes very important. Series of workshops on AfL can be conducted for teachers in order to equip them with this new concept of assessment. They can also be provided with a variety of examples of AfL in practice (Black and Wiliam, 1998b) and encouraged to try them out in their classrooms or modify them to suit their students.

## References

- Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003) *Assessment for Learning: Putting it into Practice*. Berkshire: Open University Press.
- Black, P. and Wiliam, D. (1998a) Assessment and Classroom Learning. *Assessment in Education* 5, 1, 7-74.

- Black, P. and Wiliam, D. (1998b) Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan* 80, 2, 139-148.
- Black, P. and Wiliam, D. (2005) Lessons from Around the World: How Policies, Politics and Cultures Constrain and Afford Assessment Practices. *The Curriculum Journal* 16, 2, 249-261.
- Broadfoot, P. (1995) Performance Assessment in Perspective: International Trends and Current English Practice, in H. Torrance (ed.) *Evaluating Authentic Assessment: Problems and Possibilities in New Approaches to Assessment* (p 9-43). Buckingham: Open University Press.
- Carless, D. (2005) Prospects for the Implementation of Assessment for Learning. *Assessment in Education* 12, 1, 39-54.
- CERI (2005) *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: OECD.
- Cheng, K. M. (1998) Can Education Values be Borrowed? Looking into Cultural Differences. *Peabody Journal of Education* 73, 2, 11-30.
- Cheng, K. M. (2004) Examination Culture. *Education (Jiao Shu Yu Ren)* 11, 10.
- Cizek, G. J. (1997) Learning, Achievement, and Assessment: Constructs at a Crossroads, in G. D. Phye (ed.) *Handbook of Classroom Assessment: Learning, Achievement, and Adjustment* (p 1-32). San Diego: Academic Press.
- Crooks, T. (2002) Educational Assessment in New Zealand Schools. *Assessment in Education* 9, 2, 237-253.
- Eckstein, M. A. and Noah, H. J. (1993) *Secondary School Examinations: International Perspectives on Policies and Practice*. USA: Yale University Press.
- Han, J. X. (1997) On Quality Education and the Reform of National Matriculation Tests in China. *China Examination* 3, 50-52.
- Harlen, W. (2005) Teacher's Summative Practices and Assessment for Learning - Tensions and Synergies. *The Curriculum Journal* 16, 2, 207-223.
- Harlen, W. and Deakin Crick, R. (2002) A Systematic Review of the Impact on Students, Teachers and the Curriculum of the Process of using Assessment by Teachers for Summative Purposes (EPPI-Centre Review). *Research Evidence in Education Library* 4. London: EPPI-Centre; Social Science Research Unit, Institute of Education.
- Hu, C. T. (1984) The Historical Background: Examinations and Controls in Pre-Modern China. *Comparative Education* 20, 1, 7-26.
- Jin, Y. (2008) Perspectives on College English Teaching Innovation: Assessment and Teaching. *Foreign Language Education in China* 1, 3, 57-66.
- Lai, C. T. (1970) *A Scholar in Imperial China*. Hong Kong: Kelly and Walsh.
- Liu, Z. and Hu, Z. (2005) On the Opposite of Quality-Oriented Education to Exam-Oriented Education. *Curriculum, Teaching Material and Method* 10, 3-8.

- MEPRC (2004) *College English Curriculum Requirements (for Trial Implementation)*. Shanghai: Shanghai Foreign Language Education Press.
- Sadler, R. (1989) Formative Assessment and the Design of Instructional Systems. *Instructional Science* 18, 119-144.
- Torrance, H. (1993) Formative Assessment: Some Theoretical Problems and Empirical Questions. *Cambridge Journal of Education* 23, 3, 333-344.
- Turner, Y. and Acker, A. (2002) *Education in the New China: Shaping Ideas at Work*. Aldershot: Ashgate.
- Wang, X. F. (2007) A Case Study of Transition from Summative to Formative Assessment in Chinese Context. Paper presented at the World Congress on Engineering and Computer Science (WCECS) conference. San Francisco, USA. 24th – 26th October 2007.
- Yan, Z. Y. (1997) On the Harm of Examination-Oriented Education. *Shanxi Education* 5, 11.
- Zeng, K. (1999) *Dragon Gate: Competitive Examinations and their Consequences*. London: Cassell.
- Zhang, Z. Y. (2005) A Historical Review of the College English Test Band IV and Band VI. *Modern Higher Education* 2, 100-103.