ASSESSING CONFIDENCE IN THE CHINESE LEARNER

Stephen Bruce

Assessing Confidence in the Chinese Learner

Stephen Bruce, Educational Development, Napier University,
Edinburgh, EH10 5LG

Abstract

This short paper firstly examines the cultural issues of confidence-based assessment by observing a Chinese learner group completing a multiple choice test in English as a Foreign Language (EFL). The results show no indication that the Chinese students behaved differently from their UK counterparts. The second aim of this pilot study is to examine whether good confidence judgement is a measure, and so potentially a positive mediator of academic self-concept. Student responses to the confident-based test were compared with their responses to an academic self-description questionnaire. Scores from the 'perception of ability & achievement in EFL' scale show a positive correlation with the number of correct answers selected with high confidence. Initial results indicate that as a learning tool, confidence-based assessment may have a valuable role in the development of positive academic behaviours, for younger learners in particular.

Introduction

Chinese students represent the largest international group in UK universities and their numbers are growing. However the two educational cultures have very different learning, teaching and assessment approaches, and this has posed unique challenges to educators and learners alike. It is well documented that Chinese learners show a preference for rote-learning and memorisation, so-called surface level learning strategies. From the Western perspective this approach can only reproduce the bare essentials, and the learner is unable to reach the larger meaning of the material studied. However responses to learning strategy questionnaires indicate quite the contrary and Chinese students out-perform their Western counterparts academically whether at home or abroad (Watkins & Biggs, 1996).

A further paradox arises when considering that learner self-esteem has long been known to be positively associated with academic achievement (Brookover et al., 1964; Prendergast & Binder, 1975; Song & Hattie, 1984). However, comparative research shows that Chinese levels of self esteem are lower than that of UK and American college students (Paschal and Kuo, 1973), and more recently that the level of self-esteem in young Hong Kong learners is significantly below that of their white-UK and UK-Chinese counterparts (Chan, 2000). One possible reason for this may be due to the more self-effacing and modest values in Chinese culture, strongly influenced by the Confucian tradition of a 'humble' character. Also the traditional

authoritarian style of education and the highly competitive pressures created by schools, families and society may have a bearing.

In a case study of Chinese students at Napier University, Gourlay (2004) writes, "The English Foundation Programme (EFP) was set up in the Centre for Business Languages at Napier University in 2001, and aimed to prepare Chinese students for entry to Masters Programmes, both in terms of linguistic and study skills. Students on the EFP hold offers for Napier conditional on the attainment of the required IELTS thresholds (International English Language Testing System). The experience of the first year of the EFP showed that the linguistic and study skill needs of these students were complex, the "learning curve" for them was extremely steep, and that the demands they faced on a linguistically complex and culturally unfamiliar one-year Masters would be considerable, even given the requisite IELTS pass." A sample of students from this programme form the subject of this pilot study.

Confidence-based assessment

There is a long history of research into confidence testing, with the purpose of extracting further information from objective tests (see review by Echternacht, 1972). Confidence assessment asks the learner to select an answer, and then to select his/her level of confidence that he/she has answered correctly, thus creating weighted answer and test scores. By employing a properly motivating scoring scheme, the learner is encouraged to distinguish between reliable and unreliable answers in order to gain high marks. This awareness is vital for academic knowledge to be usable, that is, knowledge which can be applied when making decisions and performing actions (Hassmen & Hunt, 1994). Gardner-Medwin & Gahan (2003) report that when choosing a confidence level, a student firstly estimates the probability that the chosen answer will be correct, and then assesses the impact of the reward or benefit when the answer turns out to be right or wrong. In discussion however, students rarely express their decisions in terms of explicit probabilities, and so is generally an intuitive matter.

Confidence-based assessment promises several improvements to the multiple choice test item, including more reflective answering, reduced guesswork, improved feedback, and increased test validity and reliability (Ahlgren, 1969; Gardner-Medwin & Gahan, 2003). In recent years web-based confidence assessment has proved popular as a self, formative and summative assessment method with medical and biomedical students (Gardner-Medwin & Gahan, 2003, Khan et al. 2001) and computing students (Davies, 2002). As a further educational tool, valuable feedback has included a comparison of the weighted and unweighted test scores and commented on the learner's perceived confidence and highlighted areas of misplaced knowledge (Khan et al. 2001).

Early confidence-based assessment studies sought to derive a measure of personality traits and largely focused on risk-taking or gambling tendencies (Swineford, 1941; Gritten & Johnson, 1941; Ziller, 1957). Recent empirical analysis do reveal a more cautious use of higher confidence levels in exams compared to formative self-assessment, but with no gender differences

appearing under either condition (Gardner-Medwin & Gahan, 2003). The first aim of this short paper is to investigate any cultural issues regarding confidence-based assessment. The Chinese educational culture is remarkably different to that of the West, and so Chinese students studying in the UK are an ideal group with which to observe responses to confidence-based assessment.

Academic self-concept, achievement and confidence judgement

Improving academic self-concept is often posited as mediating other desirable attributes such as persistence on academic tasks, motivation, self-efficacy, and self-attributions in success and failure situations (Craven 1999). Informed interventions for enhancing academic self-concept can make use of recent advances in theory. In a literature review by Shavelson et al. (1976), a model of self-concept was posited (Figure 1), and numerous studies have been based or extended from this model (Hattie, 1992). The model is multifaceted in nature, and hierarchical in structure whereby general self-concept is at the apex, and which is composed of the first-order aspects; academic self-concept (itself composed of subject-specific self-concepts) and non-academic self-concept (composed of social, emotional and physical self-concepts).
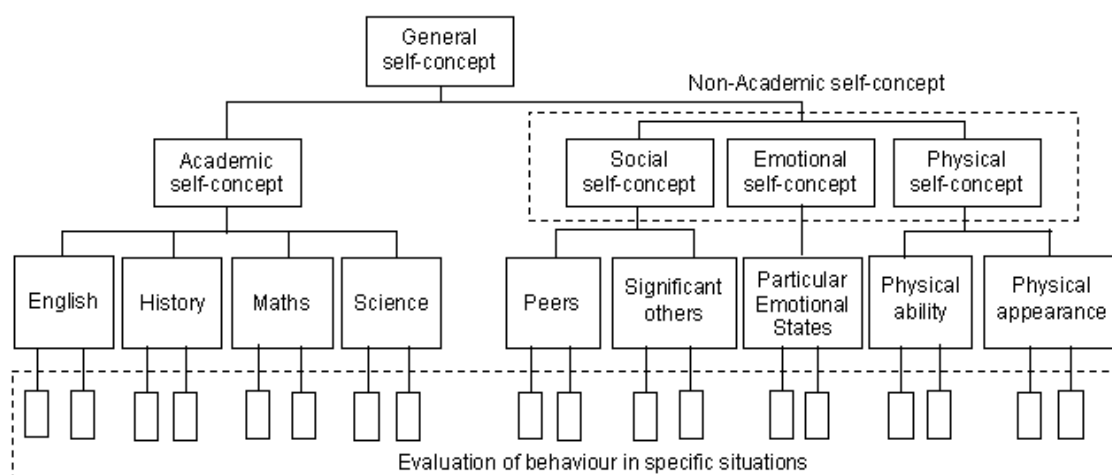

Figure 1.   A model of the hierarchical nature of self-concept (Shavelson, Hubner and Stanton,1976)

There exists strong support for the multifaceted nature. In a study of Korean adolescents, Song & Hattie (1984) report that, unsurprisingly, academic self-concept affected academic achievement much more strongly than any of the aspects of non-academic self-concept. Arguing that self-concepts will be influenced by the views of what learners would ideally like them to be, the multifaceted nature was further tested and supported by Waugh (1999) by analysing the responses from Australian undergraduates of their ideal and real self-concept. Support for the hierarchical nature is less clear, and Hattie (1992) reports of more evidence for a hierarchical model for adolescents but a unitary structure for younger children.

A study of Australian secondary school students by Watkins & Hattie (1990) found that higher levels of academic self-concept were related to the development of 'deep' and 'achieving' approaches to learning. Central to a

students approach to learning and his/her academic self-concept is the locus of control inventory. Learners with a strong internal locus of control believe that academic achievement is determined by their own abilities and efforts, whereas learners with an external locus of control believe that academic achievement is largely determined by factors outwith their control. Watkins & Biggs (1996) provide a cross-cultural summary of correlations between these scales and self-esteem and internal locus of control. A strong relationship was found between self-esteem and both 'deep' and 'achieving' approaches, and that more 'surface' approaches to learning were associated with an external locus of control.

The second aim of this short paper is to investigate whether good confidence judgement is a measure, and potentially a positive mediator of academic self-concept. Responses to an academic self-description questionnaire are examined for any correlation with responses to a confidence-based, multiple choice test. The confidence-based scoring scheme is such that a student cannot attain high marks unless their certainty of knowledge is considered. The rewards and penalties, together with the test feedback has the effect of equating internal expectation (probability that the question is answered correctly) with external performance (the question score). This process would appear to facilitate a desirable attribute that is associated with an internal locus of control, and potentially act as a significant mediator of academic self-concept enhancement.

Method

Any attempt to correlate academic self-concept measures to confidence-based assessment must consider carefully the choice of levels available, and the scoring scheme used. This study uses the scheme employed the LAPTlite system at University College London (UCL) by Gardner-Medwin & Gahan (2003), shown in Table 1. Whilst there are large penalties for incorrect answers at higher confidence, this scheme ensures that there are clear probability thresholds from which learners can select one confidence level (C) or another.

Confidence level:	C=1 (low)	C=2 (mid)	C=3 (high)
Score if correct:	1	2	3
Score if incorrect	0	- 1	- 4
Probability correct	< 50%	> 50%	> 75%

Table 2. Confidence-based scoring scheme (from Gardner-Medwin & Gahan, 2003)

The students receive two test scores; a standard %correct score (eg. 30 correct = 75%), and a confidence-based mark (CBM) which can vary from 0 to

150%. A 100% CBM is achieved with 40 correct answers at C=2 (mid) confidence, so therefore 40 correct at C=3 would be 150%. A combination of subject knowledge and good confidence judgement will see a higher CBM than the %correct score, with appropriately positive comments in the test feedback. A lower CBM provides students with guidance as to why they have lost marks, as the feedback also provides a summary of the confidence levels selected, and the number of correct answers at each level. Students are instructed to pay particular attention to incorrect answers at high confidence.

The academic self-description questionnaire consists of two sections. The first, adapted from Marsh (1992), Song & Hattie (1984) and Waugh (2001), was designed to measure the student's perceptions of their academic ability and achievement. Aligning with the posited hierarchical structure of the academic self-concept model, this first section had two scales; general academic ability and achievement, based upon their general university experiences in China; and subject-specific ability and achievement, based only on their English as Foreign Language (EFL) class experiences at Napier. Both scales consisted of 10 statements to which the students respond with their relative agreement. The hierarchical structure of academic self-concept predicts that if any correlations exist between the perception of ability and achievement and the confidence based scores (or confidence level totals), the subject-specific scale (EFL) would be stronger. The second section of the questionnaire was a locus of control inventory adapted from Trice (1985). This consisted of 28 statements to which the students responded True or False. Scores towards 0 indicate a more internal locus of control whereas scores towards 28 indicate an external locus of control.

The group consisted of 31 mainland Chinese students (54% sample), current on the English Foundation Programme, with intentions of enrolling on postgraduate courses at Napier. The students had completed their IELTS exam but at the time of the study, they had not yet received their results. There were 18 males and 13 females, and the age groupings were: 18-25 (19), 26-32 (8), 33-40 (4). The students firstly completed the online academic self-description questionnaire, receiving automatic feedback of their scores, supplemented by a verbal explanation from the author. They then completed the confidence-based, academic vocabulary test consisting of 40 multiple choice questions, again receiving automatic feedback. After further verbal feedback the students completed a short questionnaire to provide comments on their experience of test.

Results & conclusions

Chinese student performance on the confidence-based test

The %correct scores and confidence-based marks (CBM) of the Chinese students were compared to data from 930 submissions from UK students on LAPTlite (various tests, and including 40 first-time responses). Figure 2 shows that the responses from the Chinese students appear similar to those of the UK students in terms of their confidence judgement ability. Low marks are more represented in the Chinese data, which reflects the difficulty of this

language comprehension test, where small grammatical differences separated some answers and distracters. Figure 3 shows that as the question difficulty decreased, the students responded correspondingly with higher confidence. It is clear from the LAPTlite submissions in Figure 2 that % correct scores and the CBM (ie. good confidence judgement) are strongly correlated. This is also evident in the Chinese student data where the %correct and CBM correlation coefficient is 0.91 (p=0.000). There was no significant correlation between the results of IELTS exam and the %correct or CBM, although this is not surprising as the competencies measured by the IELTS exam is much broader than those of the confidence test.

There were no indications that Chinese students are reluctant to select high confidence levels, and on average high confidence was selected most often (C=3, 38%; C=2, 33%; C=1, 29%). The lowest %correct scorers tended to display over-confidence resulting in low CBM, whereas higher %correct scorers showed better confidence judgement. No differences in the %correct scores or CBM were observed between the genders.
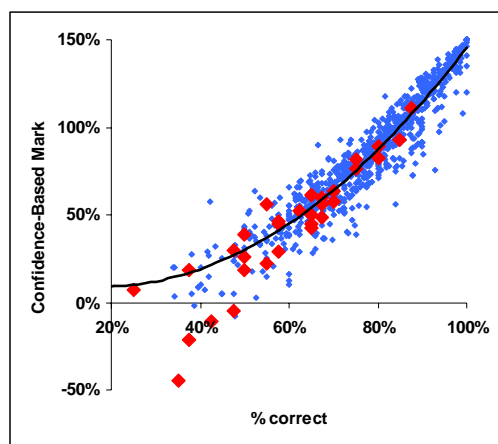

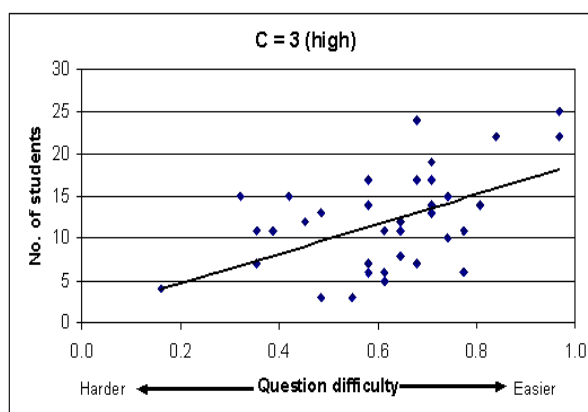Fig 2. Confidence-based test scores from Chinese (large dots) and UK students (courtesy of T. Gardner-Medwin).

Fig 3. As the question difficulty decreased, the Chinese students responded with higher confidence.

Chinese responses to the academic self-description questionnaire

The two perception of ability and achievement scales (general and EFL) showed a strong positive correlation (0.61, p=0.000). These two scales showed a strong, negative correlation with the locus of control scale, (general scale=-0.54, p=0.002 and EFL scale = -0.61, p=0.000). That is, as locus scores become increasingly internal (towards a score of 0), perception of ability and achievement scores increase accordingly.

Interestingly there were noticeable differences in the academic self-description between the genders. Females tended to respond with lower scores than males for both of the perception of ability and achievement scales, and the responses indicating the most internal of locus of control scores were male students. This is perhaps understandable given the long history of higher educational opportunities being denied to the female

population in China, particularly in rural areas and in general, women face pressure to conform to traditional social conventions (Turner and Acker, 2002). Interestingly, females generally attained higher scores in the IELTS exam.

Correlating confidence-based responses with academic self-concept

Perception of ability & achievement in EFL scores show a moderately positive correlation (0.35, $p=0.052$) with the number of correct answers at C=3 (high) confidence levels, whereas no significant correlation was observed for the general perception scale or the locus of control. These figures may indicate a trend, but may also be noise as general perception showed a moderately, negative correlation with the number of correct answers at C=2 (mid) confidence. In order that any genuine correlations can be identified, a series of confidence-based tests are required.

The student's perception of ability and achievement in EFL shows a moderate, positive correlation with their IELTS exam result (0.43, $p=0.024$), whereas there is was no significant correlation evident for the IELTS exam with the general perception scale or the locus of control. This is aligned with academic self-concept and achievement research which reports strong subject based correlation, and therefore support for the hierarchical structure.

Conclusions

This pilot study has indicated that Chinese students perform in a similar manner to confidence-based assessment as their UK counterparts. A selection of student comments are given:

> "It is a real good interesting test and I can receive different aspects of my knowledge. However, it will cost lots of time to finish this exam. As a result, I suggest we can do the exam at a regular time such an once a week."

> "It is an interesting test and I would like to do it at my university. I think that is obviously fair for students."

> "It is interesting and a little difficult to understand the result. I think I should get a higher confidence-based score because I chose (c3) nine times, and the number answered correct is 7. Maybe I have not understood this well."

> "I think it is quite interesting and helpful. It is also a good way to show me the link between confidence and academic study."

An accepted measure of academic self-concept is presented, which is moderately correlated with the IELTS exam results. Correlations with the confidence-based scores and responses may be identified with greater certainty from a series of confidence-based tests. This is planned for the Chinese students enrolling on the next English Foundation programme.

Confidence-based assessment has a potentially valuable role for the enhancement of academic self-concept, and the development of other positive academic behaviours, particularly in late primary/early secondary school learners. The scoring scheme and the test feedback provide by this assessment format seems applicable to recent research that posits a reciprocal relationship between academic self-concept and academic achievement (Marsh, 2003).

Acknowledgments

References

Brookover, W.B., Thomas, S. and Paterson A. (1964) Self-concept of ability and school achievement, Sociology of Education, 37, pp. 271-279.

Chan, Y.M. (2000) Self-esteem: A cross-cultural comparison of British-Chinese, white British and Hong Kong Chinese children, Educational Psychology, 20, pp. 59-74.

Craven, R. (1996) Enhancing Academic Self-Concept: A Large-Scale Longitudinal Study in an Educational Setting, PhD thesis, University of Sydney

Davies, P. (2002) There's no confidence in multiple-choice testing, Proceedings of the 6th International CAA conference, Loughborough, pp. 119-130.

Echternacht, G.J. (1972) The use of confidence testing in objective tests, Review of Educational Research, 42:2, pp. 217-237.

Gardner-Medwin, A.R. and Gahan M. (2003) Formative and summative confidence-based assessment, Proceedings of the 7th International CAA conference, Loughborough, pp. 147-155.

Gourlay, L. (2004) Crossing Boundaries: A Case Study Masters Level Chinese Students, LTSN (In Press)

Gritten, F. and Johnson, D.M. (1941) Individual differences in judging multiple-choice questions, Journal of Educational Psychology, 32, pp. 423-430.

Hassmen, P. and Hunt, D.P. (1994) Human self-assessment in multiple-choice tesing, Journal of Educational Measurement, 31, pp. 149-160

Hattie, J. (1992) Self-Concept, Lawrence Erlbaum Associates, New Jersey

Khan, K.S., Davies, D.A. and Gupta, J.K. (2001) Formative self assessment using multiple true-false questions on the Internet: feedback according to confidence about correct knowledge. Medical Teacher, 23, pp. 158-16

Marsh, H.W (1992a) Self-Description Questionnaire (SDQ)III: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and research monograph, MacArthur, New South Wales, Australia, University of Western Sydney, Faculty of Education.

Marsh, H.W. (2003) A reciprocal effects model of the causal ordering of academic self-concept and achievement, Paper presented at NZARE AARE, Auckland, New Zealand.

Paschal B.J. and Kuo Y.Y, (1973) Anxiety and self-concept among American and Chinese college students, College Student Journal, 7, pp. 7-13.

Prendergast, M.A. and Binder D.M. (1975) Relationships of selected self-concept and academic achievement measures, Measurement and Evaluation Guidance, 8, pp. 92-95.

Shavelson, R.J., Hubner, J.J., Stanton, G.C. (1976) Self-concept: Validation of construct interpretation, Review of Educational Research, 46, pp. 407-441.

Song, I.S. and Hattie, J. (1984) Home environment, self-concept and academic achievement measures, Journal of Educational Psychology, 76, pp. 1269-1281.

Swineford, F. (1941) Analysis of a personality trait, Journal of Educational Psychology, 32, pp. 438-444.

Turner, Y. and Acker, A (2002) Education in the New China: Shaping Ideas at Work, Ashgate, England (ISBN 0 7546 19141).

Watkins, D.A. and Biggs, J.B. (Eds, 1996) The Chinese Learner: Cultural, Psychological and Contextual Influences, CERC & ACER, (ISBN 0 86431 182 6).

Watkins, D.A. and Hattie, J. (1990) Individual and contextual differences in the approaches to learning of Australian secondary school students, Educational Psychology, 10:4, pp. 333-342.

Waugh, R.F. (2001) Measuring ideal and real self-concept on the same scale, based on a multifaceted, hierarchical model of self-concept, Educational and Psychological Measurement, 61:1, pp. 85-101.

Ziller, R.C. (1957) A measure of the gambling response-set in objective tests, Psychometrika, 22, pp. 289-292.