Learning and Individual Differences 22 (2012) 747–758

Contents lists available at SciVerse ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety?

Lazar Stankov [a,*], Jihyun Lee [b], Wenshu Luo [b], David J. Hogan [b]

[a] University of Western Sydney, Australia
[b] National Institute of Education, Singapore

ARTICLE INFO

Article history:
Received 23 August 2011
Received in revised form 30 April 2012
Accepted 24 May 2012

Keywords:
Confidence
Self-beliefs
Non-cognitive
Predictors of academic achievement

ABSTRACT

In this paper we report the results from a study that assessed confidence together with scales measuring self-belief – i.e., self-efficacy, different kinds of self-concepts, and anxiety – among the 15-year old students from Singapore. A distinct confidence factor was identified in the domains of mathematics (N = 1940) and English (N = 1786). Our results show that confidence is: a) a robust individual differences dimension; b) that can be combined with accuracy information to obtain bias scores that may be useful for group comparisons and for identification of misconceptions about particular topics. Confidence as studied in our work to date has been c) the best predictor of achievement in both mathematics and English; d) is related to both cognitive and self-belief measures; and e) it captures much of the predictive variance of other self-beliefs that are, in turn, among the best known predictors of achievement.

1. Broader context: the role of student variables in school achievement

Considerable educational investment goes into school- and teacher-level development such as teacher education and training, curriculum development, instructional strategies, and program evaluation. However, researchers have known all along that variables other than those related to formal education and training play a particularly important role in student achievement. For instance, Marzano (2000) concluded that 80% of the variance in achievement could be accounted for by student effects, 13% by teacher effects, and only 7% by school effects. Hattie's (2009) summary of the findings from over 800 meta-analyses shows a somewhat more promising view of the role of teachers and schools: he reports 31 meta-analyses that have produced Cohen's d greater than 1.00. Among these, 10 studies show the importance of school/teacher/curricula (average Cohen's d = 1.12) and 11 studies deal with teaching practice itself (average Cohen's d = 1.32). However, the highest average Cohen's d (1.48) is found in the 10 studies reporting on *student* effects. Thus, although teaching and school do make a difference in students' learning, the strongest effects arise mostly from psychological variables — i.e., individual differences between students themselves.[1]

Among the student variables the most important predictor of academic achievement is general cognitive ability or aspects thereof such as measures of fluid and crystallized intelligence (see Hattie, 2009; Kuncel, Hezlett, & Ones, 2004). Recently, there has been a shift in research from a strong emphasis on cognitive abilities to the possible role of non-cognitive variables. The broad aim of this shift is to uncover important non-cognitive variables that affect performance on achievement tests. The hope is that such variables may be more amenable to change, sensitive to intervention, and perhaps lead to an improved academic performance. These non-cognitive variables may also be treated as important educational outcomes on their own.

1.1. Confidence

The focus of the present paper is on the role of a student variable confidence which should be understood as a state of being certain about the success of a particular behavioral act.[2] In most of our work, the behavior in question is performance on a series of items in a cognitive test. According to Efklides (2011) confidence can be seen as task specific metacognitive experience. Our findings show that confidence has properties of a trait — a disposition to respond in a particular way, relative to other individuals, when asked to indicate the level of certainty about the accuracy of one's answers. A

* Corresponding author at: Centre for Positive Psychology and Education, University of Western Sydney, Bankstown Campus, Locked Bag 1797, Penrith South NSW 2751, Australia.
E-mail addresses: lazondi@rocketmail.com, l.stankov@uws.edu.au (L. Stankov).

[1] Cultural effects brought about through parental attitudes and behaviors may also be important, particularly in Confucian societies. Meta-analyses of parents' role in education do not appear in Hattie's (2009) book. It is possible that parental influences will also turn out to be more important than the effects of formal schooling at least in some countries.

[2] We do not think that using a similar-sounding term 'Self-confidence' instead of plain Confidence is necessary. For Confidence, Thesaurus gives several words with the prefix "self-". Dictionary definitions of both "-efficacy" and "-concept" call a prefix "self-" to provide psychological meaning of these terms for the reader. In other words, "self-" is not naturally associated with "-efficacy" or "-concept"; it is naturally associated with "confidence".

1041-6080/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.lindif.2012.05.013

measure of confidence captures cognitive aspects (i.e., the probability of being correct, which is higher if the person has high ability or the test is easy), aspects akin to personality (or self-beliefs or self-perceptions about the competencies related to performance on cognitive tests and the degree of openness to experience of the person) and motivation (i.e., the intention to make accurate self-appraisals in the given situation).

Stankov (1999) positions confidence on the "no-man's-land" *between* cognitive abilities and personality. Because of its multifaceted nature, we wish to claim that confidence is an important psychological variable that can help in the prediction and understanding of aspects of behavior and we present empirical evidence in this paper in support of this claim by showing that it has a strong relationship to achievement.

Previous work on confidence was inspired by studies from outside education and in particular studies of decision making and calibration of forecasting (e.g., weather, economics and medical diagnosis; see e.g., Crawford & Stankov, 1996; Stankov & Crawford, 1996a,b, 1997). However, our argument for the importance of confidence also rests on its conceptual links to several areas that have been of interest to educational psychologists. Thus, measures of confidence have been used successfully in studies of metacognition to assess persons' abilities to know what they know and what they do not know (Krebs & Roebers, 2010; see also Tobias & Everson, 2000) and, in particular, to assess metacognitive monitoring processes (see Kleitman & Moscrop, 2010; Stankov, 2000). Confidence is also assessed within the feeling of knowing (FOK) paradigm that examines the distinction between memory processes of recall and retrieval (see Dunlosky & Bjork, 2008) and, as we shall see shortly, it is frequently assessed in studies of self-efficacy (see Pajares, 1997).

Our aim in this paper is to report on the use of measures of confidence in educational research with Secondary school students and to examine its concurrent validity for predicting achievement scores in mathematics and English. Although we shall examine the use of confidence for the prediction of future performance with a smaller sample of students, most of the work to be reported here is based on simultaneous assessments. The validity of confidence will be compared to the validity of three well-established academic self-belief constructs: self-efficacy, self-concept and (math) anxiety.

1.2. Measurement issues: the relationship between accuracy and confidence

Some of the desirable properties of confidence in our work derive from the nature of its measurement — i.e., its operational definition is linked to the measurement of accuracy in typical testing situations. Briefly, after the administration of a test item (multiple-choice or any other type) the respondent is asked to provide an answer and immediately afterwards indicate on a scale, usually an 11-point rating scale ranging from 0% to 100%, how confident she/he is that the answer is correct. In typical circumstances, two scores are calculated. First, the accuracy score is simply the usual number of correct scores divided by the number of items in the test (proportion) and multiplied by 100 to arrive at percentage correct. Second, the confidence score is the average confidence rating over all items in the test, and is also on a percentage scale. These two scores are sometimes referred to as measures of objective and subjective probabilities respectively.

In our approach, accuracy and confidence scores are derived from the same cognitive act. Both scores have been used in psychophysical studies since the dawn of measurement in psychology. The third, measure – speed of mental operations – was also a part of this early work but it is not employed in the present study. Thus, performance on any aptitude test can be measured with respect to the success of cognitive activity (accuracy scores), speed (time to provide an answer) and metacognitive skill (confidence in the accuracy of the answer provided and/or bias and other scores derived from calibration studies that are used extensively in decision making literature). It is important to keep in mind that the extensive literature on accuracy, mental speed and confidence indicates that these three scores define different dimensions in individual differences (Stankov, 2000). This means that despite experimental and temporal dependencies of measurement, the three scores assess different aspects of behavior.

1.2.1. Bias scores and group differences

Frequently, the relationship between accuracy and confidence is expressed as a derived score – simple difference between confidence and accuracy scores, labeled "bias" or "overconfidence" – which can be calculated for each participant. This difference score is also sometimes referred to as the "realism of confidence" score and when its value is zero, the person (or a group of people) is said to be perfectly calibrated.[3] On most cognitive tests, however, at least two thirds of participants displays overconfidence (positive bias score) and the average bias score is usually positive in value.

Intuitively, bias scores have an appeal as a measure of metacognition or "knowing what you know and what you do not know", but this interpretation must be made with caution. Bias scores are not an indication of whether the person knows that he/she has answered each item correctly or incorrectly. They tell us whether, on the average over all items in the test, the person was able to detect those items that are easier or more difficult for him/her and produce ratings that are reflective of his/her level of performance on these items.

Stankov and Crawford (1996b) report psychometric properties of different scores derived from calibration curves. From among the different indices suggested in the literature, the bias score had the most satisfactory reliability estimates.[4] Nevertheless, being a difference score, bias scores' reliability tends to be somewhat lower than reliabilities of its constituent components of accuracy and confidence. For this reason, the use of bias scores in correlational studies is questionable. Our preference is to employ separate confidence and accuracy scores in studies of individual differences.

Bias scores can however be used as a short-hand description of group differences. In this paper we report on bias scores' differences with respect to gender but, clearly, any other group comparisons can be carried out. Pallier et al. (2002) reported that in university student samples males tend to be more overconfident than females but the differences are often small and not significant (see Stankov & Lee, 2008). We shall employ the same procedure with the 15-year-olds in this study.

1.2.2. Ability, item difficulty and confidence

Confidence, as defined and measured in our studies, is tied to a particular behavioral act. Since the third component of confidence mentioned above – i.e., motivation to participate in cognitive activity – is usually assumed in both high- and low-stake testing situations, the effects of variations of both person's ability and item difficulty on confidence need to be fully understood. Stankov, Lee, and Paek (2009) employed Item Response Theory (IRT, Rasch model) to examine these relationships. In their approach, the IRT model is fitted in the usual way to accuracy scores to obtain typical Item Response Curves. A separate confidence curve is obtained by plotting for each

[3] It is clear that a particular average confidence score may arise either from giving high and low confidence ratings to groups of items in a test or by being consistent across all items in the sense of being realistic in self-assessment. In our experience, instances of high and low ratings to groups of items are extremely rare — participants tend to be consistent in their confidence ratings.

[4] In addition to bias scores, a commonly used metacognitive measure is "discrimination" score. This is simply the difference in confidence ratings between items that were solved correctly and those that were solved incorrectly. In our work, this measure tends to have low reliability. Some other measures of calibration and discrimination are described in Boekaerts and Rozendaal (2010).