BACKGROUND The practice of medicine demands that its physician practitioners are self-directed, lifelong learners. The Self-Directed Learning Readiness Scale (SDLRS) intends to measure adults’ readiness to engage in self-directed learning.

PURPOSE The present study assesses the underlying factor structure of the SDLRS for a sample of entering medical students.

METHODS Over a period of 6 years, 972 first-year medical students at the Virginia Commonwealth University School of Medicine completed the SDLRS. To summarise the inter-relationships among variables, a principal axis factor analysis with oblique rotation was used on the 58 SDLRS items. A series of confirmatory factor analyses using LISREL 8.54 was performed to further examine the measurement model underlying the SDLRS.

RESULTS A 4-factor confirmatory model representing 4 correlated substantive factors and a reverse coding method factor fits these data well.

CONCLUSIONS Medical educators should hold limited expectations of the SDLRS to measure medical students’ readiness to engage in self-directed learning. The definitions and theoretical assumptions that inform readiness for self-directed learning should be reconsidered. Alternative approaches to studying self-directed learning should be explored.
change and is self-confident; one who is able to use basic study skills, organize his or her time and set an appropriate pace for learning, and to develop a plan for completing work; one who enjoys learning and has a tendency to be goal-oriented.\textsuperscript{5} (p. 73).

Guglielmino developed items that related to these qualities and produced her original version of the SDLRS consisting of 41 items. She tested it on a group of 307 US high school juniors and seniors, college undergraduates, and adults in continuing education courses. She reported that principal component analysis (PCA) with varimax rotation yielded an 8-factor structure. She labeled these factors: (1) openness to learning opportunities; (2) self-concept as an effective learner; (3) initiative and independence in learning; (4) informed acceptance or responsibility for one’s own learning; (5) love of learning; (6) creativity; (7) future orientation; and (8) ability to use basic study skills and problem solving skills.\textsuperscript{6} Later she modified the SDLRS to include 58 items.\textsuperscript{6}

The SDLRS has been widely used but also criticised. Field questioned the validity of the scale as a measure of ‘readiness’ for self-directed learning.\textsuperscript{7} Responding to Field’s criticism, Guglielmino wrote that readiness in the scale title is ‘a measure of an individual’s current level of readiness to engage in self-directed learning “with the implication that this level may change”’\textsuperscript{8} (p. 236). Bonham also questioned the construct validity of the SDLRS, suggesting that low scores on the SDLRS could mean two things: a dislike for learning or one’s preference for his or her learning to be directed by another.\textsuperscript{9} Brockett and Hiemstra stated, ‘…the evidence is rather convincing that early concerns raised about certain items of the scale are warranted’\textsuperscript{10} (p. 73).

Subsequent empirical work yielded mixed results regarding what the SDLRS actually measured. Mourad and Torrance administered the 58 item SDLRS to a random sample of 684 K-12 students enrolled in a programme for gifted children at the University of Georgia.\textsuperscript{11} Their PCA suggested an 8-factor model; they concluded, ‘more studies are needed to validate the scale using different samples’ (p. 102). Field’s SDLRS study involved 244 adult students in Sydney, Australia.\textsuperscript{7} Using common factor analysis he identified 4 factors but then concluded that the ‘scale measures a construct that is homogeneous’\textsuperscript{7} (p. 138). The single construct was love and enthusiasm for learning. Bligh\textsuperscript{12} conducted an SDLRS study with medical trainees (\(n = 216\)) during their medical education preparation in the UK. His PCA yielded 3 major factors: enthusiasm for learning, positive self-concept as a learner, and orientation to learning.

With the exception of Field’s study, all the investigations reported herein employed PCA. While PCA and factor analysis methods may bear some similarities, they tend to produce different results. Thus it is challenging to compare Field’s factor structure with the others’ component structures. Indeed McCune criticised Field along these lines when she wrote, ‘Also, Field should realise that if Guglielmino used principal component analysis while he used common factor analysis, their results should differ\textsuperscript{13} (p. 245). Researchers often use PCA when conducting exploratory factor analysis (EFA). We believe even Field’s study was exploratory in nature though McCune offered a competing view when she wrote, ‘He states that “eight factors are sought” which I assume is his attempt to portray his analysis as a confirmatory factor analysis\textsuperscript{13} (p. 245). PCA reduces a large set of items into smaller components and accounts for all of the variance among the items, but the studies we reported using PCA were trying to identify underlying structures of the SDLRS. PCA is not designed for that purpose. Preacher and MacCallum explain the distinction between PCA and exploratory factor analysis.\textsuperscript{14}

PCA yields observable composite variables (components), which account for a mixture of common and unique variance (including random error). The distinction between common and unique sources of variance is not recognised in PCA, and no attempt is made to separate unique variance from the factors being extracted. Thus, components in PCA are conceptually and mathematically quite different from factors in EFA.\textsuperscript{14} (p. 20).

West and Bentley examined the underlying SDLRS measurement model using confirmatory factor analysis (CFA).\textsuperscript{15} Their study was conducted with 439 K-12 Tennessee teachers and administrators. The analysis concluded that an orthogonal solution to the SDLRS measurement model was not adequate. More importantly, they reported that a highly correlated 6-factor model best described the underlying theory. The 6 factors were: (1) love of learning; (2) self-confidence as a learner; (3) openness to challenge; (4) inquisitive nature; (5) self-understanding; and (6) acceptance of responsibility for learning. Interestingly, factor 3 contained mostly reverse scored items; yet, the possibility of reverse scored methods variance was not reported. Finally, they conjectured that the first order factors could be subsumed under a single factor characterising a higher order structure.
What did we learn from the literature? First, nearly every exploratory study concluded that there was a component/factor that consisted of reverse scored items including Guglielmino’s original analysis. These items included negative statements about a high self-directed learner or a positive statement about a low self-directed learner. Not 1 author fully explained this phenomenon. Second, all researchers used CFA techniques except West and Bentley who used CFA methods. CFA allows the investigator to specify a second order factor to account for relationships among first order factors. In addition, it provides the researcher with information for disentangling random and systematic (method) error variance. For example, reserves scored items may be investigated as a source of method variance.

To study the psychometric nature of the SDLRS for entering medical students, we collected SDLRS data from 972 students and conducted an EFA study with half the data and confirmed the model it produced with a CFA study using the other half of the data. Preacher and MacCallum emphasised the need for making good decisions in the process of conducting exploratory factor analysis. They were particularly concerned about using PCA, retaining components with eigenvalues greater than one, and using varimax rotation; a bundle of procedures affectionately termed ‘Little Jiffy’. This potentially limited approach was used in most of the previous SDLRS factor analysis studies we reviewed. We followed Preacher and MacCallum’s guidelines in our factor analysis.

**METHODS AND RESULTS**

**Sample and instrument administration**

The 58-item SDLRS was group administered to first year medical students at Virginia Commonwealth University School of Medicine early in the fall semester for the graduating classes of 1999–2005 (excluding the class of 2004). SDLRS items feature 5 response choices ranging from (1) almost never true of me; I hardly ever feel this way to (5) almost always true of me; there are very few times when I don’t feel this way. Seventeen of the items are reverse scored. According to Guglielmino, wording was reversed in some of the items to prevent the response set of acquiescence (agreeing with all the items). The SDLRS was scored by its publisher. Missing responses were scored as a 3 and the publisher excluded cases missing 4 or more responses from analysis.

**Exploratory factor analysis – method**

We used Principal Axis Factor Analysis SPSS 11.0 for Windows to extract factors. To decide what factors to retain we used the scree plot, results from previous studies, and our comfort with the extracted factors. We decided to use an oblique rotation (Promax) based upon the West and Bentley study and Preacher and MacCallum’s argument, ‘...if the researcher does not know how the factors are related to each other, there is no reason to assume that they are completely independent. It is almost always safer to assume that there is not perfect independence, and to use oblique rotation instead of orthogonal rotation’ (p. 26). Individual loadings of 0.30 or greater were used in the factor designation. Extracted factors were examined and named based on an analysis of the items loading on each factor. Cronbach’s was used to estimate the internal consistency of the items constituting a factor.

**Exploratory factor analysis – results**

First year medical students from the graduating classes of 1999–2001 (N = 471) completed the SDLRS. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.90 indicating that the sample is large enough to detect the factors. The five factors, n items/factor, Cronbach’s, and factor correlations follow.

1. Learning is a tool for life (11 items; α = 0.85; r1,2 = 0.30; r1,3 = 0.59; r1,4 = 0.42; r1,5 = 0.33)
2. Self-confidence in abilities and skills for learning (15 items; α = 0.82; r2,3 = 0.43; r2,4 = 0.39; r2,5 = 0.22)
3. Responsibility for own learning (15 items; α = 0.74; 10 items; r3,4 = 0.59; r3,5 = 0.29)
4. Curiosity (8 items; α = 0.71; r4,5 = 0.25)
5. Reversed scored factor (6 items; α = 0.65)

Interestingly, the fifth factor measured all reverse scored items. This phenomenon was also reported by Guglielmino, Field, and Mourad and Torrance. Guglielmino noted her PCA showed that all items in factor 1 of her original 41 item version were negative statements. She speculated that it is possible that the factor also includes an avoidance of agreement with negative statements (p. 61). Field and Mourad and Torrance also reported a factor in which items were phrased so that they had to be reverse scored. Mourad and Torrance offered 2 explanations for this factor (1) ‘an attitude toward negative statements’ and (2) ‘preference for complex and ambitious situations’ (p. 99). Field labelled the factor.
containing all negatively worded items as ‘facility with negatively phrased items’ and argued it was not related to readiness for self-directed learning.

We found 17 items were designated as reverse scored. All 6 factor 5 items were reverse scored and had loadings of 0.31 to 0.40. Even though this factor met certain technical criteria, we dropped factor 5 because we were not sure what it meant. These six items caused us confusion:

1. I don’t like dealing with questions where there is not one right answer (0.40).
2. I’ll be glad when I’m finished learning (0.39).
3. In a classroom, I expect the teacher to tell all class members exactly what to do at all times (0.35).
4. I don’t like challenging learning situations (0.34).
5. I’m not as interested in learning as some other people seem to be (0.32), and
6. When I see something I don’t understand, I stay away from it (0.31).

Because the reverse scored phenomenon has been repeatedly reported in the literature we decided it was worth analysing in the confirmatory phase of our study.

Confirmatory factory analysis – method

Using LISREL 8.54 a series of confirmatory factor analyses was performed to further examine the measurement model underlying the SDLRS. First, the 58 items were trimmed to 41 in order to obtain items that loaded on only one factor in the specified model. Items were retained if: (a) they had factor loadings = 0.30; and (b) their secondary loadings were < 0.30. This resulted in the deletion of 17 items.

The correlation matrix among the 41 items and the item standard deviations were used as input to LISREL 8.54 and analysis was performed using the LISREL generated covariance matrix and maximum likelihood estimation. The conventional chi-square test, comparative fit index (CFI), and root mean square error approximation (RMSEA) values were used to evaluate model fit. A nonsignificant \( P > 0.05 \) \( \chi^2 \) is desirable and suggests the model adequately represents the data. The CFI can range from 0 to 1.0 and estimates the proportion of the sample variances and covariances explained by the model. The RMSEA estimates the lack of fit in a model compared to a perfect (saturated) model. CFI values > 0.90 and RMSEA values < 0.08 are considered to represent ‘good’ correspondence between observed and hypothesised factor solutions. Standardised path coefficients (factor loadings), factor correlations and second order loadings were examined to evaluate the relationship between each indicator (SDLRS item) with its associated factor.

We hypothesised a series of three models (Models = A, B, C) based on the logic provided by Anderson and Gerbing. Model A is a 4-factor confirmatory model that represents the substantive factors derived from the previous EFA. Since, previous research raised concerns about effects with reverse coded items, reverse coded effects were examined by testing for a reverse coding method factor in Model B. Model B is a 5-factor confirmatory model that represents 4 correlated substantive factors and an orthogonal reverse coding method factor that loads on reverse scored items from all substantive factors. Models A and B are compared using the chi-square difference test for nested models to determine the importance of the reverse coding method factor. The null hypothesis for the comparison of Models A and B states that there is no difference in model fit. In other words, the reverse coding method factor loadings equal zero. In this study, a significant chi-square difference would indicate support for the presence of a reverse coding factor.

Next, Model C, a confirmatory higher order model was developed and evaluated. This model includes a higher order factor that is presumed to account for the 4 first order factors examined in Model A, and like Model B, includes a reverse scoring factor. The null hypothesis for the comparison of Models B and C states that there is no difference in model fit. In other words, the relationships between the first order factors can be adequately explained by a higher order construct. In this study, a nonsignificant chi-square difference would indicate support for the presence of a second order factor. In this case, the second order model preferred would be more parsimonious, describing the directions of relationships between the four factors.

Confirmatory factory analysis – results

First year medical students from the graduating classes of 2002, 2003, and 2005 (N = 501) completed the SDLRS for the confirmatory phase of this study. Table 1 presents the chi-square and associated degrees of freedom for Model A (four correlated substantive factors), as well as the overall estimates of model fit. The CFI and RMSEA values for Model A indicate good model fit (0.91 and 0.058, respectively). All of the completely standardised factor
loadings for the items associated with the 4 factors were statistically significant ($t$-values > 1.96) with the exception of item #20, and ranged in value from 0.13 to 0.72 (see Table 2). The error terms for the items ranged from 0.48 to 0.97. The correlation coefficients between these 4 substantive factors are presented in Table 3 and ranged from 0.52 to 0.82. Thus we have confirmed the four correlated factor structure generated from our EFA.

Table 1 presents the chi-square analysis for Models A and B as well as the results from the chi-square difference tests. The chi-square difference test comparing Model A and B indicates support for rejecting the null hypothesis. As shown in Table 1, the comparison yields a chi-square difference of 52.6 with 8 degrees of freedom. This exceeds the 15.5 chi-square critical value for 8 degrees of freedom. The CFI and RMSEA values for Model B indicate good model fit (0.91 and 0.058, respectively). Thus, the model that included 4 substantive factors plus a reverse coding method factor (Model B) was shown to have better fit than the model that specified only 4 substantive factors (Model A). Table 2 lists the indicator factor loadings and significance levels for items in Models A and B. In Model B, six of the 8 reverse-coded items have significant factor loadings ($t$-values > 1.96) for the reverse coding method factor. The 8 reverse-coded items are:

1. I don’t work very well on my own,
2. Even if I have a great idea, I can’t seem to develop a plan for making it work,
3. Understanding what I read is a problem for me,
4. If I don’t learn, it’s not my fault,
5. I think libraries are boring places (not statistically significant),
6. It’s better to stick with the learning methods that we know will work instead of always trying new ones (not statistically significant),
7. Constant learning is a bore, and
8. Learning doesn’t make any difference in my life.

The reverse scored factor accounted for a small amount of variance in the 8 items (4–15%). These items also maintained significant factor loadings for their respective substantive factors.

Model C, which includes a higher order factor accounting for the 4 first-order substantive factors and 1 reverse coded method factor was developed and evaluated. Table 1 presents the chi-square and associated degrees of freedom from Models B and C, as well as the results from the chi-square difference test. In Model C, all paths between the 2nd order factor and the four first order factors are statistically significant ($t$-values > 1.96) and range in magnitude from 0.72 to 0.98. Finally, the percentage of variance in the first order factors associated with the second order factor range from 53% to 95%. The chi-square difference test comparing Model B to Model C was significant, yielding a chi-square difference of 32.6, with 2 degrees of freedom. This exceeds the 5.99 chi-square critical value for 2 degrees of freedom indicating Model B is a better fit of the data. Figure 1 graphically depicts this SDLRS factor structure.

**DISCUSSION**

Guglielmi no used an inductive approach to develop the SDLRS. She made clear, ‘Results of the Delphi survey were used as a guideline in the construction of items for the scale. Characteristics which emerged from the survey with a rating of desirable, necessary, or essential were considered for inclusion.’5 (p. 37). She explained ‘A one-to-one correspondence between SDLRS items and characteristics selected by the Delphi survey was not possible,'
since situational and attitudinal items were desired.\(^{5}\) (p. 38). In our study, the SDLRS instrument did not fully measure these characteristics or situational and attitudinal constructs that Guglielmino specified.

Our EFA study produced 4 acceptable factors. Since McCune cautioned that factor analysis studies of the same instrument could yield different results depending on the samples used,\(^{13}\) we were not surprised that our EFA factors varied somewhat from

\[\text{Table 2 Factor Loadings: Completely Standardised Solution}\]

<table>
<thead>
<tr>
<th>Item</th>
<th>FAC1</th>
<th>FAC2</th>
<th>FAC3</th>
<th>FAC4</th>
<th>FAC1</th>
<th>FAC2</th>
<th>FAC3</th>
<th>FAC4</th>
<th>RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>0.68*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.69*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>0.69*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.69*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>0.39*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.39*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>0.56*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.55*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>0.57*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.56*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>0.64*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.64*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>0.47*</td>
<td></td>
<td></td>
<td>0.45*</td>
<td></td>
<td>0.45*</td>
<td></td>
<td>0.42*</td>
<td>0.32*</td>
</tr>
<tr>
<td>58</td>
<td>0.42*</td>
<td></td>
<td></td>
<td>0.42*</td>
<td></td>
<td>0.42*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(*\) factor loading is significant at \(P < 0.05\). RC = reverse coding method factor. FAC1 = factor 1 FAC2 = factor 2 FAC3 = factor 3 FAC4 = factor 4.

\(\dagger\) Model A = four correlated substantive factors \(\ddagger\) Model B = four correlated substantive factors and an orthogonal reverse coding factor.
the structures reported for other medical students and other populations, especially since many of our reviewed studies used PCA. Our EFA study also allowed us to discard nonrelevant SDLRS items. Following West and Bentley’s lead we reduced the 58 item SDLRS to 41 items. Using CFA we found that a 4-factor model with a reverse coding method factor made empirical sense.

While several studies noted the presence of a factor comprising all or mostly reverse scored items, we learned that the reverse coding method variance had not been considered or further explored. The SDLRS contains 17 items that are reverse scored. A growing body of evidence concludes that reverse-coded items may weaken the internal consistency reliability of test score and impair interpretation of the factor structure.

It is unlikely that an author would choose to construct an instrument that contained method variance. In her initial study, Guglielmino purposefully included reverse coded scoring to offset ‘...the tendency to be acquiescent and agree with every statement.’ (p. 41). Assuming that negatively and positively worded items measure the same construct, experts such as Nunnally recommend that rating scales include positively and negatively worded items to ward off the effects of response set. In his empirical work with the Global Self-Esteem Scale (GSE), Marsh used CFA and identified method variance associated with negatively worded items. This effect complicated the interpretation of the GSE. Marsh suggests that the ‘easiest way to eliminate the effects of negatively worded items is to use only positively worded items...’ (p. 817). Although this approach runs counter to the recommendations of measurement experts, the problem of method variance associated with reverse coded items suggests a limitation of including reverse scored items in the SDLRS.

West and Bentley proposed the presence of a higher order model but did not provide parameter estimates, fit indices, or model comparison tests (chi square difference tests) to support their contention. To examine their assertion we tested a confirmatory higher order model presumed to account for the 6 first order factors using their 6-factor correlation matrix. All paths between the 6 first order factors and the higher order factor were statistically significant ($t$-values > 1.96) and ranged in magnitude from 0.65 to 0.93. The percentage of variance in the 6 first order factors associated with the second order factor ranged from 42% to 87%. The analysis provided evidence that a higher order factor structure could fit their data.

In our study, we searched for the presence of a second order factor, based on the moderate to high correlations found between the 4 factors produced by our CFA (See Table 3) and our analysis of West and Bentley’s 6-factor correlation matrix. Although the SDLRS has been widely used as a single scale of readiness for self-directed learning, our analysis did not support this scheme. Instead, a model that includes 4 substantive factors plus a reverse coding method factor accurately describes the SDLRS factor structure. Despite confirmation of this structure, the reader should note the substantial amount of measurement error associated with each item.

What can we conclude from our study of the SDLRS? We acknowledge Guglielmino’s efforts to develop a practical instrument for measuring self-directed learning readiness. Her carefully constructed approach to generating the instrument appears appropriate; yet, the SDLRS apparently falls short of measuring characteristics that Guglielmino determined were associated with self-directed learning. With our two samples, the SDLRS measured the respondents’ perceptions of how often they felt positively about: (1) learning being a tool for life; (2) their self-confidence in their abilities and skills.

Table 3 Correlations between factors for the SDLRS ($n = 501$)

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Learning is a tool for life</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Self confidence in abilities and skills for learning</td>
<td>0.52</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Responsibility for own learning</td>
<td>0.82</td>
<td>0.67</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>4. Curiosity</td>
<td>0.53</td>
<td>0.68</td>
<td>0.67</td>
<td>1.00</td>
</tr>
</tbody>
</table>

These correlations are obtained from Model A = 4 correlated substantive factors.
for learning; (3) taking responsibility for their own learning; and (4) their curiosity. While it makes intuitive sense that self-directed learners would perceive themselves as holding these characteristics in abundance, there is no reason to believe that these perceptions would predict self-directed learning.
behaviour. Guglielmino identified only characteristics of self-directed learners from her original Delphi study; she offered no theory of self-directed learning or of readiness. The theoretical constructs of self-directed learning, of readiness, or of self-directed learning readiness may harbour among or between the 4 primary factors discovered in our samples; however, our model comparison tests did not support an overarching construct.

Even as conferences and books exploring self-directed learning proliferated during the past 25 years, there appears to be no change in the SDLRS. In considering self-directed learning in medicine, one must account for the current research as well as conditions that influence the performance of self-directed learning. Newer studies extend from the neurobiology of aging and the role of cognition in making practice changes\(^\text{30}\) to the social psychology of the physician’s changing work environment, the importance of physicians’ peers, and the accountability schemes and financial incentives built into medical practice.\(^\text{31}\) Students and practitioners are not being asked merely to take on new skills or to adjust their attitudes toward learning. Many are being asked to rethink the way they see themselves, their work, and their ongoing professional development. We agree with Baveye\(^\text{32,33}\) who suggests that the study of self-directed learning should be reoriented to an entirely new direction, away from simple measures of perceptions of self-directed learning to observed self-directed learning endeavours, apropos of the 21st century. It is time for medical educators to progress beyond the current 58 item SDLRS as a measure of self-directed learning readiness. They need to return to theoretical foundations and investigations to further develop measures that predict students’ and practitioners’ adjustment to ever changing practice environments.

**Contributors:** all authors contributed to the conceptual development of this manuscript. DH led author team, conducted initial exploratory factor analysis study, and interpreted results. SL conducted confirmatory factor analysis and interpreted results. PM initiated data collection and interpreted results of factor analysis studies. AB provided statistical and data analysis expertise. HS collected data and consulted on interpretation or early results.

**Acknowledgements:** we are grateful to Dr Larry Williams for his careful readings of our manuscript and insightful feedback & we thank the anonymous Medical Education referees for their important suggestions.

**Source of funding:** none

**Ethical approval:** none

**Conflicts of interest:** none

**REFERENCES**


Received 18 December 2003; editorial comments to authors 5 March 2004, 21 April 2004, 14 May 2004; accepted for publication 24 June 2004