The research described in Chapters 2–4 consists mainly of qualitative investigations based on structured or semi-structured interviews with campus-based and distance-learning students. These investigations have identified certain approaches, conceptions, orientations and learning styles in both campus-based and distance education. Nevertheless, the number of such studies that have been carried out with distance-learning students is actually quite small, partly because of the sheer practical difficulty involved in arranging face-to-face interviews with students who are located at a geographical distance from the institutions where they are registered. In other words, the physical separation between students and teachers in distance education affects not only the students' experience but also how research is carried out into the students' experience.

Instead of face-to-face interviews, it is possible to carry out postal surveys that use open-ended questionnaires. This was the approach that was employed by van Rossum and Schenk (1984), for example, even with campus-based students. However, this still leaves the burden of analysing a substantial body of qualitative data, even with only a relatively modest sample of participants. If a much larger sample is required then the process of data analysis becomes quite intractable unless there are a limited number of response alternatives to each question that can be encoded in a fairly straightforward manner. In research on distance education, therefore, the adoption of inventories and questionnaires (typically ones yielding quantitative data) has been dictated by the problem of geographical distance and organizational constraints (see Morgan 1984, 1991).

Four main instruments have been used to collect quantitative data from students in distance education, and these will constitute the focus for the present chapter and the four that follow:

- the Study Process Questionnaire (Chapter 5)
- the Approaches to Studying Inventory (Chapters 6 and 7)
- the Distance Education Student Progress inventory (Chapter 8)
- the Inventory of Learning Styles (Chapter 9).
In addition, a number of other instruments will be discussed in Chapter 10. In most cases, the main function of these instruments has been to operationalize the various constructs that have emerged from qualitative investigations in order to generate quantitative scores on particular dimensions or scales that reflect different aspects of learning and studying in higher education. Most of these instruments were devised for use with students taking courses at campus-based institutions of higher education but, in principle, there is no reason why they should not be adapted for use with students who are taking courses by distance learning. This in turn means that they could then be used to make direct statistical comparisons between campus-based students and distance-learning students in their distributions of scores on the relevant dimensions or scales.

The Study Behaviour Questionnaire

This chapter is mainly concerned with the Study Process Questionnaire, but I need to begin by discussing its precursor, the Study Behaviour Questionnaire (SBQ). This was devised by Biggs (1970a, b), who, in common with other researchers at the time, characterized differences in how students went about their academic studying as the product of ‘certain enduring personality characteristics’ (Biggs 1970a: 165). Although Biggs used the expression ‘study behaviour’, the SBQ actually contains items that relate to students’ beliefs, attitudes and mental processes, as well as examples of observable behaviour such as note-taking or engaging in group discussion. In each case, respondents indicate how often a statement is true of them on a five-point scale from 1 (‘This is always or nearly always true of me’) to 5 (‘This is never or very rarely true of me’), so that a high score on a given item means that the item has been rejected. The initial version of the SBQ contained 72 items, and Biggs (1970a) administered this to 314 campus-based students who were beginning teacher training courses at an Australian university. At the end of the first year of study, 260 of these students completed the SBQ for a second time. Their responses were encoded numerically and then submitted to a factor analysis.

The latter is a technique for identifying the constructs (usually called ‘factors’ or ‘components’) that seem to underlie a set of quantitative data. If several variables are all very highly correlated with one another then it is reasonable to assume that they are all tapping the same underlying construct. For example, if a lot of people are asked which hand they prefer to use for writing, for throwing, for cutting with a pair of scissors, and so on, their responses prove to be highly associated with one another (that is, most people – although not all – will report using the same hand for most of these activities). This then makes it sensible to talk about a single underlying dimension that might be called ‘handedness’ (see Richardson 1978). In other cases, however, the pattern of correlations will imply the existence of two or more underlying constructs, and there exist different methods for
identifying these. It may then be necessary to transform (or 'rotate') these factors to achieve the most meaningful interpretation and, depending upon the analytic techniques that are used, this can result in rotated factors that are either orthogonal (that is, independent of one another) or oblique (that is, correlated with one another).

In Biggs's (1970a) study, the factor analysis generated six orthogonal factors that he interpreted as follows:

- study organization
- tolerance of ambiguity
- cognitive simplicity
- capacity for intrinsic motivation
- dogmatism
- independence of study behaviour.

The test–retest reliability of the students' scores on these factors was only moderate (less than 0.50 in all cases) and typically lower in science students than in arts students. However, this could be attributed to the fairly long test–retest interval, during which many students might have modified their approaches to studying in the light of their learning experiences. Scores on four of these factors showed statistically significant correlations with performance in the end-of-year examinations in the case of students taking arts subjects, but the relationships in question were very weak and none was statistically significant in the case of students taking science subjects. The fact that individual differences in studying consisted of a number of distinct components and could not be reduced to a single dimension of 'good' versus 'bad' in terms of subsequent attainment led Biggs to reject the view that studying could be identified with a specific set of teachable skills. It follows that the value of 'study skills programmes' is highly questionable.

In a subsequent paper, Biggs (1970b) examined the relationships between the students' scores on the six factors that he had identified and various measures of personality and attainment. Some of the correlations were statistically significant, and the pattern of correlations generally tended to reinforce his previous interpretation of the factors. However, the relationships were typically very weak, and most of these correlation coefficients were not statistically significant. In some cases, the relationships were statistically significant for arts students but not for science students, which led Biggs to conclude that the SBQ was tapping strategies that were important in arts subjects but not in science subjects. However, the discrepancies can also be attributed to the fact that he had included nearly four times as many arts students as science students in his study, so that weak relationships would be less likely to be significant in the latter group.

Biggs commented that gender differences in the correlation coefficients were 'minimal' (Biggs 1970b: 290). However, scores on the 'dogmatism' factor were found to be negatively related to performance in the end-of-year examinations in male arts students, but not in science students or in female arts students. Biggs suggested that this factor was concerned with
the insulation or ‘encapsulation’ of personal values from the university experience to protect them from change. He argued that insulating one’s value system would be irrelevant in studying science as the course content would be only marginally related to personal values, but it would be maladaptive in arts students as the course content would often confront their personal values. He went on:

The fact that this relationship between encapsulation of values and performance does not appear in Arts females is possibly due to the selectivity of the present sample. Many prospective female teachers possibly regard their academic careers in fairly superficial terms. Their ultimate career would be marriage; in the meantime, they could regard their departmental bond as a means of obtaining a degree with minimal effort and with less self-involvement than would males.

(Biggs 1970b: 294)

The views expressed in this quotation would not, of course, be acceptable nowadays.

In two later papers, Biggs (1973, 1976) developed the SBQ into an instrument that measured ten scales by means of 80 items concerned with motivational and strategic aspects of studying:

- Academic aspiration: ‘Pragmatic, grade-oriented, university as means’
- Academic interest: ‘Intrinsically motivated, study as end’
- Academic neuroticism: ‘Confused, overwhelmed by demands of course work’
- Internality: ‘Sees “truth” coming from within not external authority’
- Study skills and organization: ‘Works consistently, reviews regularly, schedules work’
- Fact-rote strategy: ‘Centres on facts, details, rote learns’
- Dependence: ‘Rarely questions instructors, tests; needs support’
- Meaning assimilation: ‘Reads widely, relates to known, meaning oriented’
- Test anxiety: ‘Very concerned about tests, exams, fear of failure’
- Openness: ‘University place where values are questioned’

(Biggs 1976: 72)

Although Biggs made no explicit comment about the scoring of the items in this version of the SBQ, it would seem that he had reversed the response scale so that a high score on a given item meant that the participant had accepted the item as true of them. Their responses to the relevant items were then summed to yield a total score on each of the ten scales.

Biggs (1976) obtained responses from 464 campus-based students who were starting courses in arts and science subjects at a Canadian university. Several of the scales were correlated with the students’ grade-point averages at the end of the academic year. However, academic attainment proved to be associated with different patterns of studying in men and women, especially in the science students. For men, optimal performance was apparently achieved ‘by seeing “truth” as emerging from external sources and authorities, and not worrying too much about inter-relating past knowledge with
what one is in the process of acquiring'. However, for women, it tended to be achieved 'by making up one's own mind about 'truth', by avoiding rote learning of detail, and by actively using transformational strategies' (Biggs 1976: 77). This does rather contradict Biggs's (1970b) earlier suggestion that many female students were interested only in marriage.

The 80-item version of the SPQ

Subsequently, however, Biggs (1978a) adopted a revised position in which academic attainment was a result of both 'personological' factors (such as personality, intelligence and background) and 'institutional' factors (such as academic subject, teaching method and mode of assessment). (Strictly speaking, 'personological' means 'pertaining to the study of personality', but it is also sometimes used to refer to individual differences of all kinds: that is, to variables relating to the person as opposed to their situation or context.) Biggs argued that the effects of personological and institutional factors on students' academic performance were mediated by 'study processes', which encompassed students' values, motives and strategies. In this scheme, the personological and institutional factors constitute independent variables; study processes constitute intervening variables; and measures of academic attainment constitute dependent variables (Biggs 1984).

Biggs's (1978a) account drew upon the 'presage–process–product' model of classroom teaching that had been put forward by Dunkin and Biddle (1974: 36–48). The latter researchers described the characteristics of teachers as 'presage variables' and differentiated them from the contextual variables (including characteristics of their students) that influenced the nature of their teaching. Biggs simply reversed this distinction, stressing the role of the students rather than that of their teachers and regarding characteristics of the teachers simply as part of the institutional context in which learning occurred. Biggs also subsumed both personological and institutional factors under the heading of 'presage variables' (see Figure 5.1).

Originally, Biggs characterized the causal connections as running in one direction, from presage to processes to product (that is, to academic performance) (see also Biggs 1984, 1985, 1987: 9, 96). However, in a subsequent account, Biggs (1991) incorporated feedback loops whereby observations of learning outcomes could lead students to modify their beliefs about their own capabilities and could also induce teachers and institutions to modify their teaching practices. Eventually, Biggs (1993a) allowed for the possibility of interactions among all of the different components of this system. Although this introduces a much greater degree of flexibility into the model, there is a danger that it also deprives it of any explanatory power, in so far as it can now accommodate virtually any pattern of empirical results that might be obtained.

To operationalize the domain of study processes, Biggs (1978a) developed a revised instrument in which certain of the 80 items in the SBQ had been
reclassified and the ten scales relabelled. He called this new instrument the 'Study Process Questionnaire' (SPQ) (Box 5.1). Biggs carried out separate factor analyses on the scores obtained by 420 campus-based students at a Canadian university and by two samples consisting of 150 and 148 campus-based students at an Australian university. In each case, he found that the ten scales could be subsumed under three second-order factors, which he interpreted as representing the 'reproducing', the 'internalizing' and the 'organizing' dimensions of study processes. These encompassed the values, motives and cognitive strategies that were associated with different aspects of learning (Box 5.2).

At a cognitive level, these corresponded closely to the distinction between surface approach, deep approach and strategic approach that was discussed in Chapter 2. At a motivational level, Biggs (1979) implied that they corresponded to the distinction between extrinsic motivation, intrinsic motivation and achievement motivation. The first two kinds of motivation were discussed in Chapter 4. The third had been widely discussed by researchers in the US as a personality dimension that reflected the need to succeed in contrast to a need to avoid failure (for example, see McClelland 1961; Atkinson 1964). In Biggs's subsequent writings, he referred to the reproducing dimension as an 'instrumental' or 'utilizing' dimension, and he also referred to the organizing dimension as an 'achieving' dimension (Biggs 1978b, 1979, 1982).

Biggs (1978b) administered this version of the SPQ to a class of 60 students at a campus-based university in Australia and then asked them to read two short texts that summarized educational experiments. Half of the
Box 5.1 Scales contained in the 80-item Study Process Questionnaire

Pragmatism (10 items): Grade oriented; student sees university qualifications as a means to some other end.

Academic motivation (10 items): Intrinsically motivated; sees university study as an end in itself.

Academic neuroticism (7 items): Overwhelmed and confused by demands of course work.

Internality (8 items): Uses internal, self-determined standards of truth, not external authority.

Study skills (8 items): Works consistently, reviews regularly, schedules work.

Rote learning (8 items): Centres on facts and details and rote learns them.

Meaningful learning (8 items): Reads widely and relates material to what is already known; oriented to understand all input material.

Test anxiety (6 items): Worries about tests, exams, fear of failure.

Openness (8 items): Student sees university as a place where values are questioned.

Class dependence (7 items): Needs class structure; rarely questions lecturers or texts.

Source: Biggs 1978: 268

students were instructed to read the first text for meaning (relating the purpose of the experiment to the evidence, procedures and conclusions) but the second text for fact (concentrating on facts and experimental details). The remaining students received the two sets of instructions in the reverse order. All the participants then received a list of highly factual questions about both texts, and these were repeated at a follow-up test 1 week later. When the students were split into subgroups on the basis of their scores on the three main dimensions of the SPQ, there were no significant differences between the students who had obtained high and low scores on the internalizing and achieving dimensions. Those who had obtained high scores on the instrumental (reproducing) dimension gave significantly more correct answers than those who obtained low scores, but only on the text read for fact, not on the text read for meaning.

Biggs noted that the best performance in this study had been obtained by students who tended to adopt an instrumental or surface approach to learning under conditions in which they had been explicitly instructed to focus upon the facts and details contained in the learning materials. This remained the case even in the follow-up test 1 week later. This is perhaps unsurprising, given that the retention test in question had itself demanded the retrieval of factual information. However, Biggs pointed out that this pattern of
| **Box 5.2 Value–motive–study dimensions in the 80-item Study Process Questionnaire** |

<table>
<thead>
<tr>
<th><strong>Dimension</strong></th>
<th><strong>Value</strong></th>
<th><strong>Motive</strong></th>
<th><strong>Strategy</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reproducing</td>
<td>Pragmatism: instrumental values, university is a means to another end</td>
<td>Test anxiety, neuroticism: motive to study is fear of failing</td>
<td>Class dependence, fact-rote, minimax: goals are those defined in the course, no more; rote learned to a reproductive criterion</td>
</tr>
<tr>
<td>Internalizing</td>
<td>Openness, internality: self-growth or actualization seen as overall goal, university permits this</td>
<td>Academic motivation: intrinsic, what contributes to growth is interesting, self-motivated</td>
<td>Meaning: work unsatisfying unless understood and incorporated with existing knowledge. Class only basis for stimulation</td>
</tr>
<tr>
<td>Organizing</td>
<td>Winning through competition: university a game to show excellence</td>
<td>Achievement motivation: need for success, low anxiety</td>
<td>Structuring, organizing work, meets deadlines, plays the game</td>
</tr>
</tbody>
</table>

*Source: Biggs 1978: 276*

results was inconsistent with the assumption made by Craik and Lockhart (1972) that the use of deeper or more abstract processing would inevitably lead to better retention than the use of shallow or surface processing (see Chapter 2). As Biggs noted, there was some evidence from psychological experiments on human memory that shallow processing could in some circumstances lead to good retention, partly depending upon the aspects of the material that were being tested (see Baddeley 1978; Eysenck 1978).

Watkins and Hattie (1980) distributed this version of the SPQ in a postal survey of full-time campus-based undergraduate students at an Australian
university. Responses were obtained from 562 students, representing a 60 per cent response rate, and these were used to explore the psychometric properties of the instrument. First, they obtained measures of the internal consistency of each scale (that is, the extent to which the scores that are obtained on the individual items correlate with one another). Each of the eight scales showed a moderate but not outstanding level of internal consistency. However, there was also evidence that each of the eight scales was measuring more than one underlying dimension. Next, Watkins and Hattie carried out a factor analysis on the scores obtained by the 562 students on the eight scales. This produced clear evidence for a reproducing factor (measured by rote learning, class dependence, academic neuroticism, pragmatism and test anxiety) as well as an internalizing factor (measured by academic motivation, meaningful learning, internality and openness). Unfortunately, Biggs's (1978a) 'organizing' dimension appeared to be split into two separate factors, one measured by test anxiety and neuroticism, the other by study skills and academic motivation. Biggs himself had acknowledged difficulty in interpreting this dimension, and Watkins and Hattie suggested that this was because he had extracted too few factors from his data sets. On balance, they felt that they had obtained some support for Biggs's account of study processes, but that there was considerable scope for improving the 80-item version of the SPQ.

Watkins and Hattie (1981b) conducted further analyses of the same students' scale scores to investigate possible differences related to gender, faculty and year of study. Complete data on the latter variables were provided by 282 men and 236 women in the Faculties of Arts, Science, Rural Science and Economics. The main results were as follows:

- Men tended to achieve higher scores than women on pragmatism, academic neuroticism and class dependence, whereas women tended to achieve higher scores than men on academic motivation, study skills and openness. This suggested that men were more likely to exhibit a reproducing approach, but that women were more likely to exhibit an internalizing approach.
- Science students tended to achieve higher scores than arts students on pragmatism, academic neuroticism, study skills and rote learning. This suggested that science students were more likely to exhibit a reproducing approach than arts students.
- Finally, students in the later years of their degree courses tended to achieve higher scores on internality and openness but lower scores on study skills. This suggested that students in the later years of their degree courses were more likely to exhibit an internalizing approach than students in the earlier years of their courses. Nevertheless, Watkins and Hattie acknowledged that the interpretation of this result was ambiguous because the effects of academic level had been confounded with the possible effects of age.
The 42-item version of the SPQ

Biggs (1982) reduced the SPQ to just seven items in each of six subscales that were intended to reflect the respondents' motives and strategies on each of the three dimensions. For each item, respondents produced a response to a statement along a five-point scale between 1 ('this item is never or only rarely true of me') and 5 ('this item is always or almost always true of me'). The responses to the relevant items were summed to obtain a score on each scale and each subscale. Biggs then employed this 42-item version of the SPQ in a survey of campus-based students at five universities and nine colleges of advanced education in Australia. At the time, the latter institutions were seen as having a more vocationally oriented and less research oriented role than universities. The university students obtained higher scores than the college students on internalizing motive and strategy and on achieving strategy, but the college students obtained higher scores than the university students on utilizing motive. These differences seem to have been due, in part, to the different subject 'mix' in the two kinds of institutions. However, these data were included in a more extensive report to be discussed in a moment (Biggs 1985, 1987).

Watkins and Hattie (1981b) used the 42-item version of the SPQ in a postal survey of first-year campus-based students at an Australian university and obtained 249 responses. The main results were as follows:

- Men obtained higher scores than women on utilizing strategy, whereas women obtained higher scores than men on internalizing motive and strategy.
- Science students obtained higher scores than arts students on utilizing strategy, whereas arts students obtained higher scores than science students on internalizing motive and strategy.
- Students aged 21 years or older obtained higher scores on internalizing motive and strategy and lower scores on utilizing motive than students aged between 18 and 20 years.

These results were broadly consistent with those of Watkins and Hattie's (1981b) initial study using the 80-item version of the SPQ. They concluded that younger students, male students and science students would benefit from counselling in the use of appropriate study methods.

To simplify the terminology, and also to recognize the similarity between his own account and the ideas of other researchers into student learning in higher education, Biggs (1985, 1987: 10) relabelled his three dimensions of study processes as a 'surface approach', a 'deep approach' and an 'achieving approach', each of which had a motive component and a strategy component (Box 5.3). The 42 items making up this version of the SPQ were published as an appendix to Biggs's (1987: 132–3) book. He also produced a similar instrument, the 36-item Learning Process Questionnaire, which was intended for use with children in secondary education (see Biggs 1985, 1987: 130–2). Biggs (1987: chapter 5) conducted an analysis of the responses
Box 5.3 Scales and subscales contained in the 42-item Study Process Questionnaire

<table>
<thead>
<tr>
<th>Approach</th>
<th>Motive</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA: Surface</td>
<td>Surface motive (SM) is instrumental: main purpose is to gain a qualification with pass-only aspirations, and a corresponding fear of failure</td>
<td>Surface strategy (SM) is reproductive: limit target to bare essentials and reproduce through rote learning</td>
</tr>
<tr>
<td>DA: Deep</td>
<td>Deep motive (DM) is intrinsic: study to actualise interest and competence in particular academic subjects</td>
<td>Deep strategy (DS) is meaningful: read widely, inter-relate with previous relevant knowledge</td>
</tr>
<tr>
<td>AA: Achieving</td>
<td>Achieving motive (AM) is based on competition and ego-enhancement: obtain highest grades, whether or not material is interesting</td>
<td>Achieving strategy (AS) is based on organising: follow up all suggested readings, schedule time, behave as 'model student'</td>
</tr>
</tbody>
</table>

Source: Biggs 1985: 186

given to the 42-item version of the SPQ by 2365 campus-based students in five universities and ten colleges of advanced education in Australia. He put forward a number of conclusions:

- University students produced higher scores on deep approach and lower scores on surface approach than students in colleges of advanced education. This was not simply due to a different subject mix, because the recruitment of participants was restricted to students who were taking subjects taught in both universities and colleges, and because Biggs’s analysis controlled for the effect of subject under three broad headings (arts, education and science).
- Men tended to produce higher scores than women on surface approach, but women tended to produce higher scores than men on achieving strategy. Although there was no overall gender difference on deep approach, men tended to produce higher scores than women on this scale in university courses (and especially in education courses), while
women tended to produce higher scores than men on this scale in colleges of advanced education.

- Science students tended to produce higher scores than arts or education students on surface strategy. Science students also tended to produce higher scores on achieving strategy, with arts students being intermediate and education students lowest on this scale. Arts students tended to produce higher scores than science students on deep motive; education students tended to produce scores as high as arts students in university courses but scores as low as science students in colleges of advanced education.

- Scores on both deep approach and achieving approach tended to decline with year of study. As Biggs noted, this finding is contrary to the idea that educational experience encourages a deep approach to studying. He suggested that the trend might be due to increasing workload and concern over unemployment as students enter the later stages of a degree programme.

- Scores on deep approach tended to increase whereas scores on surface approach tended to decrease with the students' age. Scores on achieving approach showed a U-shaped function, being fairly high in the youngest students (aged 18–19 years), lowest at the age of 22 years, and highest after the age of 40 years.

Subsequent research in campus-based institutions

Although originally intended for use with students in Australia, some researchers have used the 42-item version of the SPQ in other countries. Miller et al. (1990) administered this instrument to students taking a general psychology undergraduate course at a campus-based university in the US. These researchers obtained usable questionnaires from 1119 students, which they examined for the existence of gender differences. They found no difference between men and women in their scores on deep and surface motives and strategies. Men tended to produce higher scores than women on achieving motive but women produced higher scores than men on achieving strategy. However, Miller et al. observed that even these statistically significant differences were fairly modest in their magnitude, and none of the subscale scores proved to be significantly correlated with academic performance in terms of cumulative grade-point average.

Kember and Gow (1990, 1991; Gow and Kember 1990) administered this version of the SPQ to students in different years of study at one campus-based institution of higher education in Hong Kong. For the purpose of this study, each of the 42 items was presented both in English and in a Chinese translation. In the initial stage of their research (Gow and Kember 1990), they obtained data from 1043 students and conducted a multiple regression analysis to predict their scores on the six subscales from a number
of demographic characteristics. They found, in particular, that students' scores on deep and achieving motives and strategies declined with their year of study. Gow and Kember noted that Biggs (1987) had obtained a similar trend in the case of Australian students. They inferred that students tended to become less likely to employ a deep approach as they progressed through a programme of study, and they concluded that it was most questionable whether higher education was succeeding in meeting the goals espoused by both governments and academic staff, particularly with regard to the promotion of independent learning.

Nevertheless, Kember and Gow's main aim was to test the anecdotal stereotype that students in Hong Kong relied on rote learning and memorization in their academic studies. They eventually obtained responses from 2143 students, and they compared their scores on the three main scales of the SPQ with those obtained by the students at Australian colleges of advanced education in the study by Biggs (1987). Kember and Gow found that the students in Hong Kong tended, if anything, to obtain higher scores on deep approach and lower scores on surface approach than Australian students. They concluded that any tendency for students in Hong Kong to engage in rote learning was due to the nature of the curriculum and the teaching environment rather than to any inherent characteristic of the students. Biggs (1991) presented results from several other studies using either the SPQ with tertiary students or the Learning Process Questionnaire with secondary students, and these all confirmed that students in Hong Kong tended to produce high scores on deep approach and achieving approach but low scores on surface approach.

Tooth et al. (1989) devised a shortened version of the SPQ in which just three items were used to measure each of the six subscales. They used this instrument in a postal survey of candidates who had applied for admission to medical school, and they then administered it to the students who had actually been admitted towards the beginning and the end of their first year of study. The students' scores on deep approach and achieving approach tended to decline over the three occasions of testing, whereas their scores on surface approach tended to increase over the same period. Tooth et al. also found that the students' performance in their end-of-year examinations varied directly with their scores on achieving approach and inversely with their scores on surface approach, but that it was not related at all to their scores on deep approach. They argued that the adoption of apparently maladaptive strategies was caused partly by poor performance in the mid-year examinations and partly by the students' perceptions that their assessments would place excessive emphasis on the recall of factual knowledge at the expense of understanding.

Subsequently, Wilding and Valentine (1992) administered this 18-item version of the SPQ to 263 campus-based students who had just started courses in medicine and dentistry in the UK. They found that the students' marks in their end-of-year examinations were predicted by their scores on achieving strategy and, in the case of the medical students, with their scores
on deep strategy. Finally, Wilding and Hayes (1992) found that this version of the SPQ could predict how first-year psychology students at a campus-based university in the UK took notes during a lecture: those who obtained high scores on surface motive or achieving strategy produced longer notes that recorded more of the key points in the lecture.

Volet et al. (1994) devised another short version of the SPQ to assess cross-cultural differences in approaches to studying. This contained a total of 21 items: four items drawn from each of the three motive subscales and three items drawn from each of the three strategy subscales. Volet et al. criticized earlier investigations comparing students in Hong Kong with students in Australia because these had confounded effects of individual characteristics (in other words, Hong Kong students versus Australian students) with those of contextual variables (in other words, students taking courses in Hong Kong versus students taking courses in Australia). Instead, they focused on students who were taking the same first-year unit in economics at a campus-based university in Australia. These consisted of 434 students who were described as ‘local Australian’ and 120 students who were described as ‘south-east Asian’. All of these students completed the 21-item version of the SPQ at the start of the academic year and 359 students completed it for a second time during the last week of the 13-week course. For the purpose of comparing the two groups, Volet et al. identified 63 pairs of Australian and south-east Asian students who had completed the SPQ twice and were matched in terms of their age, gender and prior study of economics.

The south-east Asian students obtained significantly higher scores than the Australian students on surface approach and achieving approach. The two subgroups did not differ in their scores on deep approach, although the Australian students obtained significantly higher scores than the south-east Asian students on deep strategy. Volet et al. claimed that these differences were due not to cultural differences but to the south-east Asian students’ successful use of survival skills in an unfamiliar educational environment. The general pattern of results was similar on the two occasions and there were no significant differences between the two subgroups in the changes in their scores over the course of time. These took the form of a significant decline in the scores on both deep approach and achieving approach, a pattern similar to that obtained by both Biggs (1987) and Gow and Kember (1990) in cross-sectional research (that is, testing different groups of students at different points in the same programme). There is, unfortunately, a basic problem in that Volet et al. did not report the cultural origins of any of the participants. Both ‘south-east Asian’ and ‘Australian’ students are culturally and ethnically extremely diverse and so this was not an ideal situation in which to evaluate cross-cultural differences in approaches to studying.

Murray-Harvey (1994) evaluated the stability of students’ scores on the 42-item version of the SPQ by measuring its test–retest reliability over 8 weeks and over 12 months and by comparing it with the Productivity Environmental
Preference Survey (Price et al. 1991). This instrument is widely used in the US as a measure of physical, environmental, social and emotional learning preferences or ‘learning styles’ (for a fairly comprehensive summary, see Jonassen and Grabowski 1993: chapter 21). Murray-Harvey obtained responses to both of these questionnaires from 406 campus-based students at an Australian university and carried out a factor analysis on their subscale scores. There was very little overlap between the factors associated with the two instruments, which implied that they represented different conceptualizations of how students learn and study. A subgroup of 280 students completed the SPQ for a second time roughly a year later and the correlations between their subscale scores on the two occasions varied from 0.42 to 0.64. A separate group of 72 students completed the SPQ on two occasions with an 8-week interval, and in their case the correlations varied from 0.48 to 0.83. These results suggested that scores on the SPQ’s subscales are not wholly reliable but are nevertheless relatively stable over the course of time in the absence of any intervention.

Wilson et al. (1996) argued that previous research using the SPQ to investigate the possibility of gender differences in approaches to studying had suffered from methodological limitations. They administered the 42-item version of the SPQ to two cohorts of students who were in the first year of a psychology programme at a campus-based university in Australia. They found no sign of any difference in the scores obtained by the men and women in either of the two cohorts. Of course, by focusing on first-year psychology students, these researchers had achieved a high level of homogeneity in their participants, but there is no guarantee that similar findings would be obtained with students in other years, other programmes or other countries.

Processes or predispositions?

As mentioned earlier, the SPQ was originally an attempt to operationalize the domain of study processes (Biggs 1978a, 1985, 1987: 8–10, 95–6). However, Kember and Gow (1989) remarked that the terms ‘deep’ and ‘surface’ were used to refer both to students’ general predispositions to learn in different ways and to different strategies that they adopted in specific learning tasks. They suggested that Marton and Säljö (1976a) had used the expression ‘approaches to learning’ when classifying the ways in which students had gone about reading specific academic articles. However, Kember and Gow noted that, when items in questionnaires mentioned ‘approaches to studying’, this was characteristically in a predispositional sense. From a practical point of view, they noted that it might be possible to induce students to shift from a surface strategy to a deep strategy when carrying out a particular learning task but that the more important issue was whether it was possible to induce them to shift from a surface predisposition to a deep predisposition.
In a similar manner, though without acknowledging Kember and Gow's article, Biggs (1993b) observed that the expression 'approaches to learning' had come to have two different meanings:

- 'the processes adopted prior to, and which directly determine, the outcome of learning'
- 'predispositions to adopt particular processes'

(Biggs 1993b: 6)

Like Kember and Gow, Biggs claimed that Marton and Säljö (1976a) had used 'approaches to learning' in the former sense. Nevertheless, this is quite incorrect. As I explained in Chapter 2, Marton and Säljö consistently used the different expression 'levels of processing' in referring to students' strategies when reading isolated passages of text. In contrast, they used the expression 'approaches to studying' when referring to the different ways in which students typically went about their academic studies. It follows that Marton and Säljö (1976a) had used the expression 'approach to studying' in a predispositional sense rather than in a strategic sense. Nevertheless, like Kember and Gow, Biggs went on to point out – correctly – that questionnaires such as the SPQ were actually concerned with students' predispositions to set about learning in particular ways. He inferred that the theoretical dimensions that the SPQ was seeking to measure should more properly be located at the presage level than at the process level (see also Biggs 1988, 1991).

Biggs pointed out that questionnaires such as the SPQ could be reworded to refer to students' approaches to learning in particular contexts, although, even here, the latter expression would still seem to have a predispositional connotation. One example of this was an investigation by Eley (1992), who looked for pairs of second-year course units at a campus-based Australian university 'in which one course unit was of a reflective nature, allowing for possible student variation in elective emphasis and interpretation of content, and in which the other was of a more defined nature, with a fixed body of content to be studied' (Eley 1992: 234). Eley asked students who had completed such pairs of course units to fill out the 42-item version of the SPQ in relation to each unit separately. He found that the first kind of course unit tended to elicit higher scores on deep approach, but lower scores on surface approach, than the second kind.

The magnitude of the differences between the scores obtained on the pairs of courses was not very great. In addition, statistically significant differences were obtained only when comparing combinations of units in some academic disciplines (for instance, accounting versus law and biochemistry versus microbiology), and not when comparing combinations of units in other disciplines (for instance, chemistry versus mathematics and statistics, and English literature versus politics and philosophy). Eley ascribed this to the apparently considerable variation in how different students tended to perceive the same course units. However, this does not affect the main conclusion from this study: that students' approaches to learning according
to the SPQ depend upon the nature and the demands of each specific course.

Scouller (1998) obtained analogous results when asking students how they prepared for two different kinds of assessment: an assignment essay and an examination with multiple-choice questions. She devised a 28-item questionnaire based upon the SPQ to measure deep approach and surface approach and supplemented it with items about the intellectual skills and abilities being tapped by a particular assessment method. She then asked 206 education students at a campus-based university in Australia to respond to all the items with regard to both kinds of assessment. She found that students perceived assignment essays as assessing higher levels of intellectual processing based on their understanding of the curriculum, and were likely to adopt deep approaches when preparing their essays. In contrast, they perceived multiple-choice examinations as assessing lower, knowledge-based levels of intellectual processing and were likely to adopt surface approaches in preparing for such examinations. These findings confirm the importance of the perceived demands of assessment in determining approaches to studying.

Of course, the decision to deliver and assess a course in a particular way is not an arbitrary one. It may be constrained by institutional factors and by the demands of external accrediting bodies, but it is partly influenced by the teacher's predisposition to teach in a specific manner, which in turn depends upon the teacher's understanding of the nature of teaching and learning. Dall’Alba (1991) interviewed 20 teachers of four different disciplines at an Australian university about the teaching of their subject, and she identified seven different conceptions of teaching:

- teaching as presenting information
- teaching as transmitting information (from teacher to student)
- teaching as illustrating the application of theory to practice
- teaching as developing concepts/principles and their interrelations
- teaching as developing the capacity to be expert
- teaching as exploring ways of understanding from particular perspectives
- teaching as bringing about conceptual change.

Similar findings were obtained by other investigators (for reviews, see Kember 1997; Prosser and Trigwell 1999: 142–7). Moreover, Trigwell et al. (1994) identified five different approaches to teaching first-year science courses, expressed in terms of teachers' intentions and strategies:

- Approach A: A teacher-focused strategy with the intention of transmitting information to students
- Approach B: A teacher-focused strategy with the intention that students acquire the concepts of the discipline
- Approach C: A teacher/student interaction strategy with the intention that students acquire the concepts of the discipline
• Approach D: A student-focused strategy aimed at students developing their conceptions
• Approach E: A student-focused strategy aimed at students changing their conceptions

(Trigwell et al. 1994: 78)

Trigwell and Prosser (1996a) went on to show that the teachers’ use of particular approaches to teaching was strongly associated with their holding corresponding conceptions of teaching and also with their attributing corresponding conceptions of learning to their own students.

Gow and Kember (1993) interviewed teachers at a campus-based institution in Hong Kong and identified different analytic categories in their beliefs about teaching and learning. They used these to construct sub-scales in a questionnaire on conceptions of teaching. On the basis of the responses given in a pilot study involving teachers in five departments, this was revised into a final instrument containing 46 items that measured nine subscales under two orientations:

<table>
<thead>
<tr>
<th>Learning facilitation</th>
<th>Knowledge transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem solving</td>
<td>Training for specific jobs</td>
</tr>
<tr>
<td>More interactive teaching</td>
<td>Greater use of media</td>
</tr>
<tr>
<td>Facilitative teaching</td>
<td>Imparting information</td>
</tr>
<tr>
<td>Pastoral interest</td>
<td>Knowledge of subject</td>
</tr>
<tr>
<td>Motivator of students</td>
<td></td>
</tr>
</tbody>
</table>

They then administered this questionnaire to teachers from 15 departments whose students had participated in longitudinal research in which they had completed the SPQ soon after starting their course and again just before their final examination. There were a number of statistically significant relationships between the mean subscale scores obtained by the students on the SPQ and the mean subscale scores obtained by the staff on the teaching questionnaire. In particular, in departments where the predominant teaching orientation was towards knowledge transmission, the students’ use of a deep approach tended to decline over their course of study. In contrast, in departments with an orientation towards learning facilitation, the students were less likely to report the use of a surface approach (see also Kember and Gow 1994).

Trigwell and Prosser (1996b) developed a similar questionnaire, the Approaches to Teaching Inventory, which contains 16 items measuring teachers’ intentions and strategies with regard to two approaches to teaching: a conceptual change/student-focused approach and an information transmission/teacher-focused approach (see also Prosser and Trigwell 1999: 176–9). Trigwell et al. (1999) administered this instrument to members of staff teaching 48 first-year courses in chemistry and physics, and they gave a modified version of the SPQ to the students who were taking those courses. Both the teachers and the students were asked to complete the
appropriate questionnaire in relation to the particular topic being taught. Trigwell et al. found a relationship between the teachers' approach to teaching and the students' approach to learning: an information-transmission/teacher-focused approach to teaching was linked with higher scores on surface approach and with lower scores on deep approach on the part of their students; in contrast, a conceptual-change student-focused approach to teaching was associated with lower scores on surface approach and (to a lesser extent) with higher scores on deep approach.

These findings are, of course, purely correlational in nature, and Trigwell et al. were careful not to make assumptions about the underlying causal relationships. Although a teacher's approach to teaching may constitute part of the institutional context that influences a student's approach to learning, it is equally possible that teachers modify their approaches to teaching to respond to the preferences or predispositions of their students. Trigwell et al. cited an example, which they themselves had observed, of tutors adapting their approach to teaching in response to requests from students to go through problems in an information transmission/teacher-focused manner. However, the implication of their findings, as well as those obtained by Eley (1992), by Gow and Kember (1993) and by Scouller (1998), is that students could be induced to adopt desirable approaches to studying by the use of appropriate forms of course design and assessment.

McKay and Kember (1997) had an opportunity to test this idea at a campus-based institution in Hong Kong where there was a widespread assumption that students relied on rote learning and expected to be spoon fed by their teachers. Students who were taking an established three-year diploma course had been given the 42-item version of the SPQ at the beginning and the end of their course. Consistent with the findings obtained by Biggs (1987), there was a non-significant decline in their scores on deep approach between the two occasions. In the light of a decision to develop a new degree course to replace the diploma, the opportunity was taken to introduce an alternative student-centred curriculum that stressed independent learning. Quite contrary to the prevailing stereotype, the students taking this course appeared to prefer the new curriculum and there was a significant increase in their scores on deep approach over the 3-year degree.

All of these findings are entirely consistent with the conclusion drawn from the interview-based research discussed in Chapter 2 that students' approaches to studying depend upon the content, the context and the demands of the relevant learning task. As McKay and Kember (1997) pointed out, although students may have a predisposition towards one particular learning approach, the latter can be influenced by extrinsic factors such as the choice of curriculum and mode of assessment. Nevertheless, this poses a fundamental problem for the presage-process-product model:

- On the one hand, if, following Biggs (1978a, 1985, 1987), the SPQ is assumed to measure study processes, it implies that 'process' variables are inherently predispositional in nature.
On the other hand, if, following Biggs (1988, 1991, 1993b), the SPQ is assumed to measure predispositions, it implies that 'presage' variables can be influenced by extrinsic factors.

The solution to this dilemma is presumably to decompose or 'fractionate' the process level into two components. One component represents the student's general predisposition towards a particular approach; this is determined partly by 'personological' factors and partly by the persisting quality of the institutional context including the nature of the curriculum and the conceptions of teaching that are prevalent in their department. The other component reflects the actual processes or strategies employed in a specific learning situation; this is determined partly by the student's general predisposition towards a particular approach but also partly by the contingent properties of that situation. In other words, Biggs's original theoretical account needs to be elaborated into a series of four different stages: presage, predisposition, process and product. Indeed, a model with precisely this structure had been put forward by Newble and Entwistle (1986).

Research in distance education

Biggs (1987: 102–23, 1988, 1989) proposed that the SPQ could be used both to inform teachers about how students respond to their teaching and to inform counsellors when helping individual students. Indeed, some researchers have employed the SPQ as a diagnostic instrument in order to identify approaches to learning in individual students taking courses by distance learning in Australia (Parer 1988; Parer and Benson 1989; Relf and Geddes 1992).

The SPQ was also employed by Ekins (1992a,b) in an investigation of distance-learning students in the Hong Kong/Macau region who were being taught in English. This investigation yielded a number of interesting findings:

- The students' scores on deep motive increased and their scores on surface motive decreased with the number of years for which they said they had been studying.
- The students' scores on deep and achieving motives and strategies increased with their self-ratings of their command of English, but their scores on surface motive and strategy were largely unrelated to their self-rated language competence.
- The students' scores on both deep motive and achieving strategy increased with the number of credits they said they had obtained.
- The students' scores on both deep and achieving motives and strategies increased and scores on surface motive and strategy decreased with the grades they said they usually obtained.

Ekins checked the reports of credits and grades produced by a small sample of students against official institutional records and found that they were
accurate. Nevertheless, as I mentioned in Chapter 3, it remains the case that students’ retrospective accounts of their past capabilities and performance are vulnerable to reconstructive biases based on their implicit theories of personal change, and it is conceivable that these theories are in turn related to their habitual approaches to studying (as might be suggested by the account given by Conway and Ross 1984).

A more serious problem is that Ekins did not appear to have adapted the SPQ for use in distance education. She employed the bilingual (English and Chinese) format that had been used previously with campus-based students in Hong Kong, but she apparently made only minimal and incidental changes to the items that explicitly referred to face-to-face teaching in classes, lectures or laboratories. It is simply not clear, for example, what students taking courses by distance learning were supposed to make of the following (slightly adapted) items in the SPQ:

I learn best from teachers/lecturers who work from carefully prepared notes and outline major points neatly on the blackboard.

After a class/lecture or lab I reread my notes to make sure they are legible and that I understand them.

(Ekins 1992a: 342–3)

In fact, both of these items were significantly correlated with the students’ reported grades. This suggests that the students might have produced hypothetical responses based upon their academic performance rather than veridical accounts of their actual approaches to studying.

Unfortunately, none of the research studies carried out with distance-learning students using the SPQ has included a comparison group of traditional, campus-based students, and none has directly addressed the issue of whether the SPQ possesses reliability, validity and a coherent structure when it is employed in either campus-based education or distance education. In fact, researchers have tended to assume that the SPQ can be taken from campus-based education and applied in the context of distance learning in a wholly unproblematic manner.

Problems with the SPQ

Hattie and Watkins (1981) administered the 42-item version of the SPQ to 255 campus-based students in their first year of study at an Australian university and 173 campus-based students in their first year of study at a university in the central Philippines. The internal consistency of each of the three main scales and the six subscales was satisfactory for the Australian students and adequate for the Filipino students. A factor analysis of the Australian students’ responses to the individual items yielded a six-factor solution in which the subscales outlined by Biggs were said to have been ‘clearly evident’ (Hattie and Watkins 1981: 243). A factor analysis on the students’ scores on the six subscales generated three factors that could be
identified with the three major approaches to learning. In the case of the Filipino students, however, both kinds of factor analysis yielded only a two-factor solution differentiating between motives and strategies. This suggested, at the very least, that the 42-item SPQ was not suitable for use with Filipino students.

However, subsequent attempts to reproduce the constituent structure of the 42-item SPQ proved less successful. In his own analyses of the subscale scores obtained by campus-based Australian students, Biggs (1987: 16; Biggs and Rihn 1984) consistently failed to retrieve three factors and instead found only two factors, which Biggs and Rihn described as a generalized deep approach to learning and a generalized surface approach to learning. A very similar pattern was obtained by Murray-Harvey (1994) in her factor analysis of the subscale scores obtained by campus-based Australian students on the 42-item SPQ and the Productivity Environment Preference Survey (Price et al. 1991). Kember and Leung (1998) analysed the subscale scores on the 42-item version of the SPQ obtained by 3254 students at a campus-based university in Hong Kong. They used confirmatory factor analysis to test whether Biggs’s model involving three major scales and six subscales gave a satisfactory fit to their data in comparison with six other theoretical models. In fact, Biggs’s model showed the worst fit, whereas the best fit was achieved by a model based upon a generalized deep approach and a generalized surface approach, both of which were partially tapped by the achieving motive and strategy subscales.

In fact, Biggs and Rihn (1984) made use of this factor structure with students who had opted to take a learning skills intervention programme at Stanford University in the US. This programme had the explicit goal of encouraging a deep approach to learning and discouraging a surface approach. In a pilot study, Biggs and Rihn found that students taking this programme tended to produce relatively high scores on all three motive subscales of the SPQ but that (in comparison with other Stanford students) they produced higher scores on the surface strategy subscale and lower scores on the deep and achieving strategy subscales. In the main study, 58 students completed the SPQ both before and after they had taken the intervention programme and were assigned scores on deep approach and surface approach on both occasions as follows:

- deep approach = deep motive + deep strategy + achieving motive + achieving strategy
- surface approach = surface motive + surface strategy + achieving motive.

Biggs and Rihn showed that there was a clear and highly significant increase in the students’ scores on deep approach as a result of taking this course, accompanied by a less pronounced but nevertheless statistically significant decrease in their scores on surface approach. (This is, it should be said, one of the few instances in which a ‘learning skills’ programme has been shown to generate desirable and significant changes in students’ approaches to learning.)
In any particular learning task it would seem that a surface approach and a deep approach are mutually exclusive, in which case one might expect that the two dimensions identified in these factor analyses should be negatively correlated with one another. In fact, in Biggs's (1987: 16; Biggs and Rihn 1984) analyses, the two factors were independent of one another, but this was simply because he had used orthogonal rather than oblique rotation. Similarly, in the 'best fit' model identified by Kember and Leung (1998), the two factors in question had been stipulated to be independent of one another. In both cases, then, the apparent independence of the factors stemmed from an artefact which the researchers imposed upon their data and which prejudged the question of whether the factors were actually correlated in any way.

Watkins and Regmi (1990) and Watkins and Akande (1992) administered the 42-item SPQ to campus-based students in Nepal and Nigeria, respectively, and carried out factor analyses with oblique rather than orthogonal rotation. They obtained two-factor solutions that were similar to those obtained by Biggs with Australian students, but they did not indicate the magnitude of the correlation between the two factors in question. O'Neil and Child (1984) administered the 42-item SPQ to 277 campus-based students in the UK and reported the full set of correlation coefficients among their participants' scores on the six subscales. These data indicate the existence of two factors and Table 5.1 shows the results of factor analyses carried out using both orthogonal rotation and oblique rotation. The data shown are the factor loadings, which are (very roughly) analogous to correlation coefficients between the subscale scores and the underlying factors. Both solutions are once again similar to those obtained by Biggs with Australian students and confirm the existence of a generalized deep approach and a generalized surface approach to studying. The correlation coefficient between the two factors obtained in the oblique solution is +0.08, which implies that they are effectively independent of each other.

The latter finding means that individual students might score high or low on both factors. This has the practical implication that interventions aimed at improving student learning should be concerned specifically with encouraging a deep approach and not necessarily with discouraging a surface approach, because achieving the latter need itself have no implications for the former (see Trigwell and Prosser 1991a). Nevertheless, the remarkable consistency between the results of factor analyses of the 42-item SPQ carried out in Australia, Hong Kong, Nepal, Nigeria and the UK casts doubt upon the existence of a separate 'achieving' approach. Indeed, Biggs had himself expressed reservations about the nature of the achieving approach:

It should be noted that deep and surface approaches are different in kind from the achieving approach. The strategies involved in the first two describe ways in which students engage the actual content of the task, while the achieving strategy describes the ways in which students organise the temporal and spatial contexts in which the task is carried
Table 5.1 Results of factor analyses of the data reported by O’Neil and Child (1984)

<table>
<thead>
<tr>
<th>SPQ subscale</th>
<th>Orthogonal solution</th>
<th></th>
<th>Oblique solution</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>Factor 1</td>
<td>Factor 2</td>
</tr>
<tr>
<td>Achieving motive</td>
<td>0.38</td>
<td>0.60</td>
<td>0.36</td>
<td>0.59</td>
</tr>
<tr>
<td>Achieving strategy</td>
<td>0.58</td>
<td>0.28</td>
<td>0.57</td>
<td>0.25</td>
</tr>
<tr>
<td>Deep motive</td>
<td>0.64</td>
<td>-0.07</td>
<td>0.64</td>
<td>-0.10</td>
</tr>
<tr>
<td>Deep strategy</td>
<td>0.76</td>
<td>-0.11</td>
<td>0.76</td>
<td>-0.14</td>
</tr>
<tr>
<td>Surface motive</td>
<td>-0.08</td>
<td>0.67</td>
<td>-0.06</td>
<td>0.67</td>
</tr>
<tr>
<td>Surface strategy</td>
<td>-0.13</td>
<td>0.71</td>
<td>-0.15</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Technical note: Both the eigenvalues-one rule and the scree test indicated the existence of two factors that explained 65.1 per cent of the variance among the six variables. A common factor analysis was carried out using squared multiple correlations as initial estimates of commonality. The orthogonal solution used a varimax rotation; the oblique solution used an oblimin rotation.

out. There is, then, no inconsistency in rote learning in a highly organised way (‘surface-achieving’) or reading for meaning in an organised way (‘deep-achieving’).

(Biggs 1985: 187)

This concurs with the implication of the interview-based research discussed in Chapter 2 that what seems to be a ‘strategic’ or ‘achieving’ approach actually reflects a sophisticated response in students with an extrinsic motivation to adopt either a deep approach or a surface approach, depending on which appears likely to maximize their grades or marks in any particular course. This is how Biggs (1978b) had originally conceived of this ‘approach’. Moreover, in his most recent accounts, Biggs (1999a: 13–19, 1999b) made no mention of an achieving approach at all.

Other problems have been raised in reports that have described factor analyses carried out on the responses produced by campus-based students to the 42 individual items in the SPQ. O’Neil and Child (1984) carried out separate factor analyses on data that had been obtained by Biggs from 827 university students and 1638 students at colleges of advanced education in Australia. In both cases, there were six interpretable factors, four of which could be readily identified with the scales relating to deep and achieving motives and strategies. However, the scales relating to the surface approach were much less clearly supported. O’Neil and Child then carried out factor analyses on the data that they themselves had obtained from 277 students at an institution of higher education in the UK and these produced very similar results. They argued that the robustness of the scales relating to a surface approach was highly questionable.

As mentioned above, Kember and Gow (1990, 1991) employed the 42-item version of the SPQ with campus-based students at an institution of
higher education in Hong Kong. In spite of the different cultural context, a factor analysis of the students' responses to the individual items of the SPQ produced clear support for the scales relating to achieving motive, achieving strategy and deep strategy. However, the items relating to deep motive were split across two different factors, and it was very difficult to associate any factors with surface motive and strategy. The consistency between these results and those reported by O'Neil and Child (1984) suggests that the SPQ is 'portable' between Western cultures and Hong Kong but is unsatisfactory in either setting. Finally, Christensen et al. (1991) administered the SPQ to 328 students at an Australian college of advanced education and carried out a factor analysis on the responses to the 42 items. They did identify six orthogonal factors but, once again, found only weak support for the scales relating to a surface approach.

Biggs (1993b) responded to the latter article with a spirited defence of the SPQ. In particular, he argued that seeking an orthogonal factor structure was inappropriate as the six subscale scores were meant to constitute the motive and strategy components of the three approaches to learning. He also claimed that previous independent evaluations of the original SPQ supported its basic theory and construction in this regard. These comments are rather disingenuous as, according to O'Neil and Child (1984), Biggs himself had initially evaluated the SPQ by using analytic techniques that imposed an orthogonal structure on the factor solution. Indeed, it was precisely in order to replicate Biggs's methods that Kember and Gow (1990, 1991) had used the same techniques and generated orthogonal solutions from their data.

In addition, O'Neil and Child (1984) had explicitly allowed for the possibility that the factors underlying responses to the 42 items of the SPQ might not be independent. They carried out a separate factor analysis on their data, which led to an oblique (that is, non-orthogonal) solution. The pattern of factor loadings was very similar to that obtained in an orthogonal solution and the intercorrelations among the factors in the oblique solution were described as 'low'. This implies that, even to the extent that the six subscales of the SPQ were confirmed in the factor solution, these subscales were 'virtually independent' of each other (see also Volet et al. 1994). Finally, O'Neil and Child conducted a second-order factor analysis to investigate whether there was a coherent structure underlying the factors identified in their oblique solution. However, they found that this 'gave no discernible pattern' (O'Neil and Child 1984: 232).

In short, even using analytic techniques that accommodate the theoretical assumptions in the presage–process–product model, there appear to be problems in the composition of the subscales that were intended to measure surface motive and surface strategy, and there is no evidence that the factors that do result constitute associated components of three underlying approaches to learning. They cannot, therefore, be said to provide distinctive, homogeneous and appropriate measures of the theoretical constructs on which the SPQ was supposed to be based. Moreover, as mentioned
earlier, even when the integrity of the subscales is taken for granted by carrying out factor analyses on the total scores obtained on the subscales themselves, there is support only for a generalized surface approach and a generalized deep approach to learning. On both theoretical and practical grounds, then, the SPQ cannot be recommended as a useful research instrument.

Concluding summary

- The SPQ has been used in a number of studies concerned with campus-based education, and it has been shown to be sensitive to differences between individual students related to their age, year of study and academic discipline. In contrast, gender differences in students' scores on the SPQ are inconsistent in direction, small in magnitude and sometimes not statistically significant at all. However, the SPQ has also been used to refute an anecdotal stereotype that students in Hong Kong rely upon rote learning and memorization.

- Students' scores on the scales of the SPQ are most naturally interpreted as measuring their predispositions to adopt different approaches to learning. However, the SPQ also appears to be sensitive to contextual factors, such as the demands of the specific courses that students are taking and in particular the conceptions of teaching that are held by their teachers. This seems to demand modification or elaboration of the original presage-process-product model of student learning that was put forward by Biggs (1978a, 1985).

- Only a few studies have been carried out using the SPQ in the context of distance education. The researchers in question have tended to assume that the constructs underlying the SPQ can be applied to the context of distance learning in an unproblematic manner, and they do not even appear to have adapted the constituent items of the SPQ to ensure that they were meaningful and appropriate for students taking courses by distance learning.

- The findings of research studies carried out with campus-based students in Australia, Hong Kong, Nepal, Nigeria and the UK indicate that the SPQ simply measures a generalized surface approach and a generalized deep approach to learning. The findings of research studies carried out with campus-based students in Australia, Hong Kong and the UK have also cast doubt upon the integrity of the subscales concerned with a surface approach. As a result, the SPQ cannot be recommended as a research instrument.