A multilevel meta-analysis of gender differences in learning orientations

Severiens, S.E.; ten Dam, G.T.M.

DOI
10.1111/j.2044-8279.1998.tb01315.x

Publication date
1998

Published in
British Journal of Educational Psychology

British Journal of Educational Psychology (1998), **68**, 595–608 *Printed in Great Britain*

A multilevel meta-analysis of gender differences in learning orientations

Sabine Severiens* and Geert ten Dam
Graduate School of Teaching and Learning, University of Amsterdam, The Netherlands

Background. Reviewing gender differences in learning processes does not result in a clear and coherent picture of the magnitude and direction of gender differences (Severiens & ten Dam, 1994). The results not only differ on several dimensions but at times they are actually contradictory.

Aims. The aim of the present study is to obtain a more coherent picture of gender differences in learning orientations.

Samples. To provide this picture, a meta-analysis was performed on 22 studies which investigate gender differences in learning orientations. The studies used Entwistle's Approaches to Studying Inventory, and were conducted in a variety of higher education settings.

Methods. The results of these studies were analysed in a multilevel approach with two levels: the respondents are nested within studies.

Results. Results show significant mean gender differences on the Reproduction Orientation (women score higher), and on the Non-academic Orientation (men score higher). Furthermore, gender differences appeared on 11 of 16 scales (e.g., relating ideas, operational learning, fear of failure, negative attitude to studying, extrinsic motivation). The effect scores of 10 of 16 scales of the ASI are heterogeneous. The observed variation in gender differences across studies could not be explained by the small number of available study characteristics.

Conclusions. More research is needed which explores factors in the educational context 'responsible' for the magnitude and direction of gender differences in learning orientations.

Individual differences between learners has been the topic of educational research for many decades. A wide range of constructs in which learners can differ has been covered extensively. Intelligence, motivation, interest, gender and ethnicity are but a few examples (Snow, 1996). One of the developments in educational research, important for the issue of individual differences, concerns its increasing attention for processes

*Correspondence and requests for reprints should be addressed to Sabine Severiens, Graduate School of Teaching and Learning, University of Amsterdam, Wibautstraat 4, 1091 GM Amsterdam, The Netherlands. E-mail: Severiens@ilo.uva.nl

related to learning itself (see e.g., Schunk, 1991; Verschaffel, 1995). Students have to learn how to deal with processing increasing amounts of information. A variety of ways of learning or learning strategies may serve as tools for obtaining the necessary knowledge in a rapidly changing society. The underlying idea is that students who learn to use the appropriate strategies in a variety of settings are more successful while going through secondary school, university and in their future lives.

Depending on the theoretical framework and questionnaire used, research on individual differences in student learning has shown a number of dimensions in which students differ (Messick, 1996). Some students prefer to learn the complete subject matter by heart while others are searching for deeper meanings (Entwistle, 1981), some through concrete experience, while others learn more through reflective observation (Kolb, 1984) and so on. These differences in preferred ways of learning are often assumed to be related to a number of individual factors, such as motivation, interest and perception of one's own abilities (Pintrich & de Groot, 1990). From the perspective that it could be *groups* of students sharing a certain interest, an attitude or motivation, or groups sharing particular experiences, it might be interesting to examine differences between groups of students. Relative to the large number of studies into individual differences in the way students learn, little research has been done on differences between groups of students.

In this article we concentrate on differences between women and men in terms of their ways of learning. Such research might offer opportunities to contribute to discussions on gender inequality in education. A detailed study of gender differences in learning may provide a basis for linking these differences to, for example, subject choices: especially because of their under-representation in mathematics and science, girls are still leaving school with fewer opportunities of going on to another form of education and with poorer chances on the labour market (e.g., Bennett & LeCompte, 1990; Sikes, 1992; ten Dam, van Eck & Volman, in press). Moreover, if it is possible to indicate the role of educational factors in producing gender differences in learning, it will at the same time provide ways of examining how to change this role.

The studies which have been conducted on relations between gender and student learning use a variety of definitions and concepts of learning (see Richardson & King, 1991; Severiens & ten Dam, 1994). In order to make sense of this collection of different conceptualisations of approaches to learning an organising tool, provided by Curry's (1983) onion model, turned out to be quite helpful (Severiens & ten Dam, 1994). Curry's onion model contains three layers in which the assumed influence of the educational context on learning varies. In the inner layer, learning concepts are considered to be stable and anchored in personality. The educational context has no impact on the way students learn. In the outer layer, learning is assumed to depend on the educational context. In the middle layer, the middle position is assumed: the way students learn is stable, but not completely fixed. Curry has not been the only one to organise different conceptualisations of student learning according to the assumed influence of the educational context. Messick (1996) describes a similar division: conceptualisations of learning vary according to the extent to which learning is assumed to depend on the task at hand.

In our narrative review (Severiens & ten Dam, 1994) we have shown that most studies on gender differences in learning use theories belonging to the middle and outer layer. A theory in the middle layer which has often been used in the 1980s to investigate gender differences in learning is Kolb's theory on experiential learning (1984). Experience plays a key role in learning and individuals differ in the ways they approach various tasks and use experience. Kolb defines learning style as a student's fairly consistent response to and use of stimuli in the context of learning (Hayden & Brown, 1985). Because it is assumed that learning styles result from hereditary factors, past experiences or socialisation and present environment (Kolb, 1984), his theory fits in the middle layer of the onion model. A meta-analysis on gender differences in learning processes using Kolb's Learning Style Inventory (LSI) (Severiens & ten Dam, 1994) shows a small gender difference on one of its four scales (Abstract Conceptualisation). Men showed a greater preference than women for the abstract conceptualisation mode of learning.

Studies which can be positioned in the middle layer of the onion model assume that the educational context has an impact on student learning, but only a small number of studies have made a beginning with investigating this impact. In these studies it is usually assumed that gender inequality in education is related to differences in ways of learning. It is argued that because gender differences in learning are not properly taken into account, differences in school careers are the result. For example[1], Bar-Haïm & Wilkes (1989) and Dippelhofer-Stiem (1989) investigated the relation between learning processes, school careers and study choices of women and men. Matthews (1991) and Thompson & O'Brien (1991) linked gender, learning processes, school careers and school performance. Some links were observed, but these were difficult to interpret. In the meta-analysis on Kolb's LSI described above, we linked the available contextual variables (discipline and country) to gender and learning styles. No statistically significant links were observed, which means that gender differences on Kolb's LSI do not vary according to discipline or country.

Theories which can be positioned in the outer layer of Curry's (1983) onion model tend to emphasise learning environment more than those in the middle layer. In this layer, the term strategies is more often used than the term styles to indicate learning processes. Learning strategies are assumed to vary according to the context, and changes in educational context imply changes in the strategies learners use. Investigating learning from the outer layer seems more in line with the recent plea for a process-oriented approach into investigating learning (Boekaerts, 1997) compared to investigating learning from the middle layer in which learning is considered to be relatively independent of context. One of the assumptions of a process-oriented approach is that both beliefs and strategies are dependent on the specific learning episodes: the way learners react in different learning contexts should be examined, for example, by examining their strategies while performing different tasks or in different kinds of assessment situations.

Instruments such as the Inventory of Learning Processes (ILP) (Schmeck, 1983), the Study Behaviour and the Study Process Questionnaire (SBQ and SPQ) (Biggs, 1987) and the Approaches to Studying Inventory (ASI) (Entwistle, 1981) measure learning strategies. A meta-analysis on a limited number of studies (seven) using Entwistle's ASI (Severiens & ten Dam, 1994) showed that gender differences mainly appear on the affective components in Entwistle's conceptualisation of approaches to studying. On the average, men score higher on the extrinsic motivation scale and the achievement motivation scale, whereas women score higher on the intrinsic motivation scale, on fear of failure and on the surface approach. On most scales, however, gender differences varied according to the studies included in the meta-analysis: the results were heterogeneous. For example, in one study it was found that men report they more often use a deep approach to learning, while in other studies women indicate they more often use this approach to learning. In this meta-analysis we did not have enough studies at our disposal to investigate reliably the effect of the available background variables.

In the present article, a meta-analysis will be discussed on 22 instead of seven studies which use Entwistle's ASI. We decided to update our 1994 meta-analysis in case of Entwistle's theory because his theory on deep and surface learning seems to fit well into recent ideas about investigating learning from a process-oriented approach. Besides, because Entwistle's theory clearly supposes the impact of the context on learning, it opens up possibilities for investigating gender differences in learning as a process in which the context plays a role. In turn, such research might point at implications which could establish more gender equality in educational practices. We make use of a multilevel approach to meta-analysis (Bryk & Raudenbush, 1992). Such an approach makes it possible to examine gender differences and the possible effect of the relevant and available background variables. The research questions of the present study are: to what extent do men and women indicate using different learning orientations and are these differences related to available background variables?

The studies

In the most important databases (SSCI, ERIC, British Education Index, EUDISED-R&D, Dissertation Abstracts Online, Sociological Abstracts, Psychinfo, ADION/DION) a search was conducted to find authors using Entwistle's Approaches to Studying Inventory (ASI) in the period 1980–1995. This search resulted in 35 authors using the ASI. A small number of authors had published statistics regarding gender differences; the remaining authors were approached to request the necessary statistics. Ten authors sent the statistics; the other authors either did not react, no longer had access to the data or did not code the data by gender. Because some authors had more than one sample available, 22 studies using the ASI were suitable for the meta-analysis (see Table 1 for a description of the studies included).

A possible problem in a meta-analysis concerns the insufficient quality of the included studies, the so-called 'garbage in, garbage out' problem (Light & Pillemer, 1984). To answer this problem, we have thoroughly examined the articles which describe the studies in the present meta-analysis. Besides, another quality control has taken place by peer review because most of the included studies have been published in scientific journals. We assume that this guarantees sufficient quality of the included studies.

Meta-analysis and multilevel analysis

A meta-analysis serves to integrate research findings of multiple studies on the same subject. Entwistle's ASI consists of different scales measuring different aspects of learning. Meta-analysing these scales in a multilevel approach will show to what extent and in which direction gender differences exist.

Bryk & Raudenbush (1992) point out that the multilevel regression model can be applied to meta-analysis because in a meta-analysis respondents are nested in studies. In fact, a meta-analysis can be considered to be a special case of the multilevel regression model. In a multilevel approach a mean effect score (indicating gender differences) and its variance are estimated. This variance is partitioned into 'within study' (sampling) variance and 'between study' (systematic) variance. It is usually assumed that the 'between study' variance is related to characteristics of the individual studies. The advantage of using this model instead of a more standard approach to meta-analysis concerns the possibility of accounting for sample variance and systematic variance in the estimation of the mean effect on its variance, and at the same time model the study characteristics to find out whether they explain any of the systematic variance.

Usually, in multilevel analysis the data of all the respondents are needed. In a meta-analysis, however, these data often are not available. But Bryk & Raudenbush (1992) note that in the case of reasonably large data sets in each separate study (larger than 30, p. 157) it is tenable to assume a normal distribution of effect scores with known

Table 1. Studies using Entwistle's ASI and selected background variables

Researchers	N_{women}	N_{men}	MO	RO	AO	NO	Age	Discipline	Country
1 Cano-Garcia *et al.*	22	66	.08	.06	.02	.42	21	Sciences	Spain
2 Cano-Garcia *et al.*	51	45	−.13	−.35	−.28	.08	21	Sciences	Spain
3 Cano-Garcia *et al.*	17	65	−.14	−.38	−.47	.09	21	other	Spain
4 Cano-Garcia *et al.*	42	39	−.08	−.25	.10	.11	21	Medicine	Spain
5 Cano-Garcia *et al.*	75	22	.18	−.06	.21	.24	21	Soc.sc.	Spain
6 Cano-Garcia *et al.*	79	60	−.13	−.35	−.03	.20	21	Soc.sc.	Spain
7 Cano-Garcia *et al.*	47	22	−.19	−.36	.22	.01	21	Arts	Spain
8 Cano-Garcia *et al.*	69	27	.05	−.65	−.09	−.08	21	Arts	Spain
9 Cano-Garcia *et al.*	52	42	−.16	−.38	−.07	.15	21	Soc.sc.	Spain
10 Cano-Garcia *et al.*	95	22	−.41	−.31	−.38	.05	21	Arts	Spain
11 Gunn*	85	70	−.18	−.17	.11	.10	34	various	Gr.Brit
12 Meyer *et al.*	79	242	.14	−.33	.08	−.02	miss	Sciences	S.Africa
13 Meyer	128	370	−.18	−.14	.06	.11	≈18	various	Gr.Brit
14 Meyer *et al.*	144	266	−.05	−.12	.00	−.02	≈18	Sciences	Gr.Brit
15 Solomonides*	389	40	−.02	−.01	.23	.24	21	Sciences	Gr.Brit
16 Miller *et al.*	650	465	.08	−.23	.04	.21	miss	various	USA
17 Gledhill *et al.*	41	135	−.03	.14	−.09	.32	≈24	Medicine	S.Africa
18 Clarke	68	85	.03	−.19	.30	.05	miss	Medicine	Australia
19 Watkins *et al.*	132	160	−.41	−.09	−.07	.15	21	various	Australia
20 Richardson	27	68	.09	−.30	miss	miss	miss	Soc.sc.	Gr.Brit
21 Richardson	113	114	.23	.00	miss	miss	miss	Soc.sc.	USA
22 Coles	44	43	−.10	.17	−.12	miss	≈18	Medicine	Gr.Brit

MO = d-score Meaning Orientation, RO = d-score Reproduction Orientation, AO = d-score Achievement Orientation, NO = d-score Non-academic Orientation, country = country the study was performed in, miss = missing, Soc.sc. = Social Sciences

Note. In the publications of studies 17 and 18 the relevant statistics were published, statistics from the remaining studies were obtained from the researchers, in studies 20 and 21 the statistics were calculated from data other than those published.

*These two studies used the revised version of the ASI (Tait & Entwistle, 1996).

variance. In this respect, they speak of the *Level 1 Variance Known* model. If the *Level 2* variance, the 'between study' variance, reaches statistical significance, it is concluded that the effect scores vary across studies. In that case, the results are described as heterogeneous: in each particular study included in the meta-analysis, gender differences may take on a different form. In the case of heterogeneous results, the available and relevant study characteristics are modelled to try to explain some of this 'between study' variance. In a multilevel approach this can be done relatively easily. If, on the other hand, the 'between study' variance is not significant, and all variance in the observed effect scores is due to sampling variance, it is concluded that the results are homogeneous[2]. It makes no sense to try to explain this (non-existing) variance.

In the present meta-analysis of gender differences in learning, heterogeneity of effect scores means that gender differences vary across studies. Results can be contradictory: for example, in one study women might score higher on the Intrinsic Motivation scale, while in another study the reverse is true. In the case of heterogeneity, it makes sense to investigate to what extent study characteristics can explain these differences. For example, do gender differences on the Intrinsic Motivation scale differ according to the discipline the study was conducted in? If, on the other hand, the results are homogeneous, it can be concluded that the magnitude of gender differences is similar in all studies. In that case, variance in effect scores seems to be due to sampling variance only.

The program of Bryk, Raudenbush & Congdon (1994) for analysing multilevel data (HLM) offers an option for analysing meta-analysis data. This option is called *Variance Known Hierarchical Linear Models* (VKHLM). A typical meta-analysis data set consists of an effect score for each study and its sampling variance and the study characteristics that are possibly related to the effect score. Because in the present meta-analysis some studies use relatively small samples, the unbiased effect size d was used (Hedges & Olkin, 1985, p. 81). This effect score is the standardised difference between two groups (in this case, women and men), corrected for small sample bias and defined as

$$d = \left(1 - \frac{3}{4N-9}\right) \times \left(\frac{M_{women} - M_{men}}{SD}\right)$$

(SD is the square root of the weighted average of the two variances)
($N = N_{women} + N_{men}$)

D-scores are usually interpreted according to Cohen's rule of thumb (1977): a d-score of .2 is small, a medium-sized d-score is about .5 and d-scores of .8 are interpreted to be large. The sampling variance of d is estimated in the following formula (Hedges & Olkin, 1985, p. 86):

$$s^2(d) = \frac{N_{women} + N_{men}}{N_{women}N_{men}} + \frac{d^2}{2(N_{women} + N_{men})}$$

In the meta-analysis, first the so-called *intercept-only* model (a model without any background variables) is estimated. In this model, the mean effect score is estimated and its variance partitioned into 'within study' variance and 'between study' variance.

The homogeneity of the 'between study' variance is tested in a chi-square test. In the case of heterogeneity, characteristics of studies are modelled to find out whether they explain anything in the systematic variance. In case of homogeneity, we assume that the magnitude of gender differences is the same in all studies and that the background variables are not modelled.

The relevant and available study characteristics are the country in which the study is conducted, the discipline and the mean age of the respondents in the study. The possible effect of country is examined because Hanna, Kundiger & Larouche (1990) and Baker & Perkins Jones (1993) have shown that gender differences vary according to country. Age is included as a background variable because relationships have been found between age and learning styles (Vermunt, 1992; Watkins & Hattie, 1981). It is therefore not implausible that gender differences in learning may also vary according to age. Lastly, discipline is examined because former studies have shown links between discipline, gender and learning orientations (see e.g., Dippelhofer-Stiem, 1989). Because the variables discipline and country are discrete variables, the effects are examined by making use of dummy variables. Each dummy variable compares one of the countries (or disciplines) to the remaining countries (or disciplines).

Results

Entwistle's ASI consists of 64 Likert-type items which constitute 16 scales[3]. Combinations of these scales result in four learning orientations: a Reproduction Orientation, a Meaning Orientation, an Achieving Orientation and a Non-academic Orientation (see Table 2). In general, the reliabilities of the learning orientations are reasonable, varying from .59 to .79. For a discussion of the reliability of this research instrument, we refer to Richardson (1990).

The learning orientations do not always appear in factor analyses. Meyer & Parsons (1989) compared the factor structures of the ASI from different studies: the Meaning Orientation usually appears as a factor, but the remaining learning orientations do not always. This means that the way in which the learning orientations are composed of scales may vary across studies. Therefore, we have decided to analyse the composite scores on the learning orientations, as defined by the authors of the ASI themselves (Tait & Entwistle, 1996) and estimate the scores on the separate scales in multilevel models as well. In other words, the multilevel analyses are conducted both on the orientations and on the scales. Table 1 describes the studies, the d-scores on the four learning orientations and the available background variables.

A model in which the Meaning Orientation is estimated shows that the mean effect score does not differ from zero significantly (see Table 3)[4]. To put it differently, women report this learning orientation as often as men. The 'between study' variance is not significant. Therefore we conclude that the result is homogeneous. It makes no sense to model the available background variables.

Analysing the scales in this orientation shows that two scales result in a significant gender difference: women score higher on Relating Ideas (d-score = −.18) and Operational Learning (d-score = −.17). Moreover, it is noteworthy that the effect scores of all scales of this learning orientation are heterogeneously distributed. Apparently, even though the variance in the composite score of the Meaning Orientation is not systematic, the variance on the separate scales which constitute this orientation is. The background variables are modelled in order to explain this heterogeneity.

On the Deep Approach scale the dummy variable comparing the science disciplines to the other disciplines has a significant coefficient. Apparently, the magnitude of gender differences in the deep approach in science differs from gender differences in other disciplines. The mean d-scores in each discipline indicate that men score higher on the deep approach in science, while in the other disciplines women score higher. After modelling the study characteristics, the variance is still heterogeneous. This means that besides discipline other study characteristics may also be related to the heterogeneity.

Gender differences on the scale Intrinsic Motivation correlate with the dummy variables for discipline: all coefficients differ from zero significantly. Gender differences in intrinsic motivation seem to vary across disciplines. The mean d-score in science is 0, in languages .15, in the social sciences 0, in medicine −.15 and in the category 'other' −.35. Thus, men seem to be more intrinsically motivated in languages, while in the other disciplines women are more intrinsically motivated, or no gender differences appear. After modelling the background variables the 'between study' variance is no longer heterogeneous. It follows that this variable explains all systematic variance in d-scores.

Table 2. Scales of the Approaches to Studying Inventory (sources: Ramsden & Entwistle, 1981; Tait & Entwistle, 1996)

Scale	Meaning
Meaning Orientation	
Deep Approach	Active questioning in learning
Use of Evidence	Relating evidence to conclusions
Inter-relating Ideas	Relating to other parts of the course
Comprehension Learning	Readiness to map out subject area and think divergently
Operation Learning	Emphasis on facts and logical analysis
Intrinsic Motivation	Interest in learning for learning's sake
Reproducing Orientation	
Surface Approach	Preoccupation with memorisation
Syllabus-boundness	Relying on staff to define learning tasks
Improvidence	Over-cautious reliance on details
Fear of Failure	Pessimism and anxiety about academic outcomes
Achieving Orientation	
Strategic Approach	Awareness of implications of academic demands made by staff
Achievement Motivation	Competitive and confident
Non-Academic Orientation	
Disorganised Study Methods	Unable to work regularly and effectively
Negative Attitudes to Studying	Lack of interest and application
Globetrotting	Over-ready to jump to conclusions
Extrinsic Motivation	Interest in courses for the qualifications they offer

The dummy variable 'country' comparing the US with other countries has a significant value on the scales Use of Evidence and Comprehension Learning. The means show that the largest difference occurs between the US and Australia. In the US men score higher on Comprehension Learning and Use of Evidence compared to US women, while Australian women more often than Australian men aim at understanding and using evidence. After modelling the background variables the systematic variance is no longer heterogeneous with regard to these scales, either.

The available background variables do not explain any variance in the other scales in the Meaning Orientation.

Table 3. Multilevel models of Entwistle's Approaches to Studying Inventory

Intercept-only models	Fixed part Coefficient (SD) intercept	*t*	*p*	Random part Variance component	d.f.	chi-^{2}	*p*	Proportion
Meaning Orientation	−.053(.042)	−1.25	.18	.011	21	28.11	.14	.14
Reproduction Orientation	−.180(.032)	−5.70	.00	.000	21	21.60	.42	.04
Achievement Orientation	.021(.032)	.63	.32	.000	19	16.93	< .50	.06
Non-academic Orientation	.132(.033)	3.97	.00	.000	18	10.07	< .50	.08

Proportion = proportion of the true parameter variance

On the Reproduction Orientation women score higher than men (d-score = −.18, see Table 3). The effect scores are homogeneous. An analysis of the separate scales shows that gender differences appear on all scales, except for the scale Syllabus-boundness. Women seem to focus more often on details, more often report using a surface approach and score higher on fear of failure compared to men. Differences on the scales for Improvidence and Fear of Failure are heterogeneous: gender differences vary across the studies. Modelling the explaining variables shows that none of the variables explains any variance. Apparently the heterogeneity of d-scores is related to other study characteristics.

Women and men do not differ with regard to the Achieving Orientation: the mean effect score does not differ from zero (see Table 3). Moreover, the effect scores are homogeneous: the systematic variance is not significant. The separate scales do have significant effect scores in opposite directions. Women seem to be more strategic in their approach (d-score = −.17) while men are more often motivated by their achievements (d-score = .19). These results are heterogeneous: the background variables turn out to explain part of the variance of the scale Strategic Study Approach. It concerns the dummy variable distinguishing the discipline medicine from the other disciplines. The means indicate that male medical students study slightly more strategically than do female medical students, while in the other disciplines women score higher on the strategic study approach. After modelling the dummy variables the variance on study level is no longer significant: the available study characteristics explain the heterogeneity.

Men score higher on the Non-academic Orientation than women (d-score = .13, see Table 3). Again the result is homogeneous: the differences across the studies are due to sampling variance. The higher scores of men appear on the scales for negative study attitude (d-score = .16) and extrinsic motivation (d-score = .29). The scales in this learning orientation are homogeneously distributed.

Summarising the results of the meta-analysis of the ASI, it can be concluded that gender differences appear on two learning orientations. Women more often report using the Reproduction Orientation and men more often a Non-academic Orientation. The Meaning and Achieving Orientations are as often reported by women as by men. The results on all four orientations are homogeneous: the characteristics of the individual studies do not seem to influence the magnitude of gender differences. However, if we consider the separate scales, another picture emerges. On 10 of the 16 scales of the ASI the results are heterogeneous and on 11 scales gender differences are found. In some cases the discipline, and in one case the country in which the study was carried out, explained part of the heterogeneity.

Conclusions

In this article we presented a meta-analysis on 22 studies using Entwistle's ASI. The research question was: to what extent do men and women indicate using different learning orientations and in what way are differences related to background variables? Because of the hierarchical character of the data set of a meta-analysis, the respondents being nested within the studies, a multilevel approach was used. In the multilevel models tested, the effect scores, indicating gender differences in learning orientations, served as dependent variables. The characteristics of the individual studies were modelled to try to explain the possible heterogeneity in effect scores. Before summarising the results, we would like to draw attention to the advantages of a multilevel approach for analysing meta-analysis data.

The possibility in a multilevel approach to take the difference between sampling and systematic variance in effect scores into account is an advantage. Moreover, the ease with which study characteristics can be modelled to find out whether they explain any of the systematic 'between study' variance makes this approach even more appealing. Hox (1994) points out that in most meta-analyses results turn out to be heterogeneous. Therefore, it seems that the multilevel approach is generally preferable compared to the more common approaches of meta-analysis.

The meta-analysis on 22 studies on gender and learning showed the following gender differences. Women score higher on the Reproduction Orientation, but men and women achieve similar scores on the Meaning Orientation. No gender differences appeared on the Achieving Orientation. Finally, we found that men score higher on the Non-academic Orientation. On all four learning orientations, the homogeneity test showed that the results are homogeneous. This means that the results seem to be independent from the specific context in which the studies were conducted.

If we also look at the scales of the ASI, firstly, gender differences appeared on 11 of 16 scales (e.g., relating ideas, operational learning, fear of failure, negative attitude to studying, extrinsic motivation). Secondly, the effect scores of 10 of 16 scales of the ASI are heterogeneous. This means that the variance in gender differences on the scales is partly systematic. The available background variables (age, discipline and country) could explain part of this systematic variance in a number of scales.

The trend which appeared in the 1994 meta-analysis (Severiens & ten Dam, 1994) on the results on the ASI scales of seven studies also appears in the present meta-analysis: the largest gender differences appear on the affective scales, and on most scales heterogeneity is observed. The conclusion must be that by adding even more studies a similar picture is likely to emerge. This finding prompts the question how many studies are needed in a meta-analysis in order to reliably answer the research question. This question seems especially relevant in the case of homogeneity when study characteristics are not related to the effect scores. A relatively small number of studies might suffice in order to draw a conclusion about gender differences. In the case of heterogeneity on the other hand, it is more important to include more studies which vary along important background variables. The number of background variables which might be related to gender differences will indicate the number of studies needed in the meta-analysis to reliably explain the variance in effect scores.

The meta-analysis allows for three conclusions which take the research domain of gender and learning a step further. First, the d-scores on the level of learning styles and learning orientations are small or non-existent on the average. On the separate scales of the ASI, however, gender differences appear more clearly. At this point one might ask whether the constructs as measured by Entwistle's ASI are sufficiently sensitive to measure gender differences. The literature in the area of gender and learning points to a number of constructs which do not appear in these instruments, but which are supposed to differentiate women and men (see e.g., Stone, 1994; Volman, ten Dam & van Eck, 1995). An example of such a construct is the assumed preference of women for interaction during learning. Entwistle's ASI did not measure this preference and is thus by definition unable to uncover such a difference. Another example of a possible gender-sensitive learning construct concerns the assumed preference of women for relating the subject matter to personal experiences. Again the instrument does not measure this construct. Consequently, an instrument incorporating the assumed gender-sensitive constructs may be valuable to the research domain of gender and learning. Future research in this area should explore the learning constructs which are, on the basis of theory, expected to differentiate between women and men.

Secondly, on the basis of this meta-analysis we conclude that gender differences in learning vary according to the context of particular studies. While in one study women showed, for example, more fear of failure than men, other studies did not show this result, or even showed contradictory results (see Table 2). The main question is: why do we observe such heterogeneity, why do gender differences vary across studies? The available study characteristics explained the heterogeneity in a few cases. Apparently, there are more factors 'responsible' for the magnitude and direction of the differences between women and men. Vermunt (1992, pp. 27–29) presented a survey of factors that can explain the different learning styles of students. He distinguishes individual characteristics such as intelligence, educational experiences and age, and contextual factors such as type of learning task, discipline, test demands and instruction in general. Further research on gender and learning should focus on investigating those factors that differentiate between women and men. In order to explain the heterogeneity found, we need more studies on gender-related aspects of learning processes in which possible relevant individual and contextual factors are included. Such a research programme can embody the process-oriented approach as advocated by Boekaerts (1997). A process-oriented approach may be valuable to the research area of gender and learning, because the present study has shown that it is likely that the educational context produces gender differences. Investigating the way women and men respond while performing a learning task (within a certain school subject) may uncover those aspects in the educational context which show differential effects related to gender.

The last conclusion deals with other factors which might be important in illuminating gender differences in learning. One of the results of the meta-analysis concerned the gender differences in the affective components of learning. Fear of failure, and extrinsic and achievement motivation appear to be sensitive to gender, across a number of different studies. Investigating the affective factors in more detail and relating these factors to (cognitive) learning theories and research may provide additional insights into gender differences in learning. The work of Pintrich & de Groot (1990) and Boekaerts (1991) on motivational beliefs may provide frameworks for such research.

NOTES

[1] The examples concern studies which do not use Kolb's LSI but different instruments.

[2] Rejecting the alternative hypothesis (the results are heterogeneous) does not necessarily imply that the null hypothesis (the results are homogeneous) is true. Therefore one needs to be careful in interpreting the homogeneity test (Snijders, 1997).

[3] Two studies use a revised version of the ASI (Tait & Entwistle, 1996) (see Table 1). Tait pointed out which scales of the revised version measure similar constructs as the original version (personal communication). 'New' scales were kept out of the analyses.

[4] When considering the results of this multilevel meta-analysis it is important to note that some of the studies are nested within researchers (for example Cano-Garcia and Justicia-Justicia). Strictly speaking, this should be accounted for in the analysis. Possibilities to do this are to give weights to the studies belonging to the same researchers, to include only one study per researcher or to include average scores of all studies per researcher. To find out whether the results change by 'summarising' the studies belonging to one researcher, we performed these analyses. One result changes: we now find heterogeneity on the Meaning Orientation scale, which does not appear if the studies are analysed separately. Apparently, including the studies nested within researchers as separate studies (as in the present article) does not result in significant effects which disappear when summarising. Therefore we have decided to consider the nested studies as separate studies.

References

Baker, D.P. & Perkins Jones, D. (1993). Creating gender equality: cross-national gender stratification and mathematical performance. *Sociology of Education*, **66**, 91–103.

Bar-Haïm, G. & Wilkes, J.M. (1989). A cognitive interpretation of the marginality and underrepresentation of women in science. *Journal of Higher Education*, **6**, 371–387.

Bennett, K. & LeCompte, M.D. (1990). *The Way Schools Work. A Sociological Analysis of Education*. New York & London: Longman.

Biggs, J.B. (1987). *Student Approaches to Learning and Studying. Research Monograph*. Australian Council for Educational Research. (ERIC No. ED 308 201.)

Boekaerts, M. (1991). Subjective competence, appraisals and self-assessment. *Learning and Instruction*, **1**, 1–17.

Boekaerts, M. (1997). Personality and the psychology of learning. *European Journal of Personality*, **10**(5), 377–404.

Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical Linear Models*. Newbury Park, CA: Sage.

Bryk, A.S., Raudenbush, S.W. & Congdon, R.T. (1994). *HLM 2/3. Hierarchical Linear Modelling with the HLM/2L and HLM/3L Programs*. Chicago: Scientific Software International.

Cano-Garcia, F. & Justicia-Justicia, F. (1994). Learning strategies, styles and approaches: an analysis of their interrelationships. *Higher Education*, **27**, 239–260.

Clarke, R.M. (1986). Students' approaches to learning in an innovative medical school: a cross-sectional study. *British Journal of Educational Psychology*, **56**, 309–321.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences* (rev. ed). New York: Academic Press.

Coles, C.R. (1985). Differences between conventional and problem-based curricula in their students' approaches to studying. *Medical Education*, **19**, 308–309.

Curry, L. (1983). An organization of learning styles theory and constructs. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada. (ERIC No. ED 235 185.)

Dippelhofer-Stiem, B. (1989). The development of research-oriented learning in five European countries. *European Journal of Psychology of Education*, **4**, 489–503.

Entwistle, N.J. (1981). *Styles of Learning and Teaching*. New York: John Wiley & Sons.

Gledhill, R.F. & Van der Merwe, C.A. (1989). Gender as a factor in student learning: preliminary findings. *Medical Education*, **23**, 201–204.

Gunn, L. (1994). The state of the postgraduate swamp: modes of study and approaches to learning. Master's Thesis, University of Aberdeen.

Hanna, G., Kündiger, E. & Larouche C. (1990). Mathematical achievement of grade 12 girls in fifteen countries. In L. Burton (Ed), *Gender and Mathematics. An International Perspective*, (pp. 87–97). London: Cassell.

Hedges, L.V. & Olkin, I. (1985). *Statistical Methods for Meta-analysis*. Orlando, FL: Academic Press.

Hayden, R.R. & Brown, M.S. (1985). Learning styles and correlates. *Psychological Reports*, **56**, 243–246.

Hox, J.J. (1994). *Applied Multilevel Analysis*. Amsterdam: TT-Publicaties.

Katz, N. (1988). Individual learning style: Israeli norms and cross-cultural equivalence of Kolb's LSI. *Journal of Cross-Cultural Psychology*, **19**(3), 361–379.

Kolb, D.A. (1984). *Experiential Learning. Experience as a Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall.

Light, R.J. & Pillemer, D.B. (1984). *Summing Up: The Science of Reviewing Research*. Cambridge, MA: Harvard University Press.

Matthews, D.B. (1991). The effects of learning style on grades of first year college students. *Research in Higher Education*, **32**, 253–268.

Messick, S. (1996). Cognitive styles and learning. In E. De Corte & F. Weinert (Eds), *International Encyclopedia of Developmental and Instructional Psychology*, (pp. 638–641). Oxford: Elsevier Science.

Meyer, J.H.F. (1995). Gender group differences in the learning behaviour of entering first-year students. *Higher Education*, **29**, 201–215.

Meyer, J.H.F., Dunne T.T. & Richardson, J.T.E. (1994). A gender comparison of contextualised study behaviour in higher education. *Higher Education*, **27**, 469–485.

Meyer, J.H.F. & Parsons, P. (1989). Approaches to studying and course perceptions using the Lancaster Inventory – a comparative study. *Studies in Higher Education*, **14**(2), 137–153.

Miller, C.D., Finley, J. & McKinley, D.L. (1990). Learning approaches and motives: male and female differences and implications for learning assistance programs. *Journal of College Student Development*, **31**(2), 147–154.

Petrakis, E. (1981). Cognitive styles of physical education majors. *Perceptual and Motor Skills*, **53**(2), 574.

Pintrich, P.R. & De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, **82**(1), 33–40.

Richardson, J.T.E. (1990). Reliability and replicability of the Approaches to Studying Questionnaire. *Studies in Higher Education*, **15**(2), 155–168.

Richardson, J.T.E. & King, E. (1991). Gender differences in the experience of higher education: quantitative and qualitative approaches. *Educational Psychology*, **11**(3–4), 363–382.
Riding, R. & Cheema, I. (1991). Cognitive styles – an overview and integration. *Educational Psychology*, **11**, 193–215.
Schunk, D. (1991). *Learning Theories. An Educational Perspective*. New York: Macmillan.
Schmeck, R.R. (1983). Learning styles of college students. In R. Dillon & R.R. Schmeck (Eds), *Individual Differences in Cognition. Vol. 1*, (pp. 233–79). New York: Academic Press.
Severiens, S.E. & ten Dam, G.T.M. (1994). Gender differences in learning styles: a narrative review and a quantative meta-analysis. *Higher Education*, **27**(4), 487–501.
Sikes, P.J. (1992). *Gender and Schools*. London: Cassell.
Snijders, T. (1997). De populaties van de meta-analyse [The populations of a meta-analysis]. *Tijdschrift voor Onderwijsresearch*, **22**(1), 3–15.
Snow, R. (1996). Individual differences, learning, and instruction. In E. De Corte & F. Weinert (Eds), *International Encyclopedia of Developmental and Instructional Psychology*, (pp. 649–660). Oxford: Elsevier Science.
Stone, L. (Ed.) (1994). *The Education Feminism Reader*. New York & London: Routledge.
Tait, H. & Entwistle, N. (1996). Identifying students at risk through ineffective study strategies. *Higher Education*, **31**(1), 97–116.
ten Dam, G.T.M., van Eck, E. & Volman, L.L. (in press). Research programmes on gender and education: results and conceptualizations. *European Journal of Educational Research*.
Thompson, M.J. & O'Brien, T.P. (1991). Learning styles and achievement in postsecondary classrooms. Paper presented at the Annual Conference of the American Educational Research Association. (ERIC No. ED 331 554.)
Vermunt, J.D.H.M. (1992). *Leerstijlen en Sturen van Leerprocessen in Hoger Onderwijs. Naar Procesgerichte Instructie en Zelfstandig Denken* [Learning styles and regulation of learning processes in higher education]. Amsterdam/Lisse: Swets & Zeitlinger.
Verschaffel, L. (1995). Beïnvloeden van leerprocessen [Influencing learning processes]. In A. Lowyck & N. Verloop (Eds), *Onderwijskunde. Een Kennisbasis voor Professionals*, (pp. 153–187). Groningen: Wolters-Noordhoff.
Volman, M.L.L., ten Dam, G.T.M. & van Eck, E. (1995). Girls in science and technology: the development of a discourse. *Gender and Education*, **7**(3), 283–292.
Watkins, D. & Hattie, J. (1981). The learning processes of Australian university students: investigations of contextual and personological factors. *British Journal of Educational Psychology*, **51**(3), 384–393.

(*Manuscript received 3 September 1997; final version 12 January 1998*)