This item is the archived peer-reviewed author-version of:

Do students develop towards more deep approaches to learning during studies? A systematic review on the development of students' deep and surface approaches to learning in higher education

Do students develop more deep approaches to learning during studies? A systematic review on the development of students' deep and surface approaches to learning in higher education

Abstract

The focus of the present paper is on the contribution of the research in the student approaches to learning (SAL) tradition. Several studies in this field have started from the assumption that students' approaches to learning develop towards more deep approaches to learning in higher education. This paper reports on a systematic review of longitudinal research on how students' approaches to learning develop during higher education. A total of 43 studies were included in the review. The results give an unclear picture of the development of approaches to learning and, thus, lack to provide clear empirical evidence for the assumption that students develop more deep approaches in higher education. Neither methodological nor conceptual aspects of the studies investigated explained the ambiguity of the research results in a clear way. Both theoretical and empirical implications for further research are discussed.

Introduction

Since the 2004 special issue in Educational Psychology Review on "Measuring Studying and Learning in Higher Education—Conceptual and Methodological Issues" (Lonka, Olkinuora, & Mäkinen, 2004) a lot of research-effort from a variety of research traditions has been invested in exploring the ways in which students learn in higher education. The focus of the present paper is on the contribution of the research in the student approaches to learning (SAL) tradition. Research in this tradition generally focuses on the different ways students engage in learning or handle learning tasks as reported by the students themselves (Biggs, 2003; Entwistle 2009; Lonka et al., 2004;Schmeck, 1988). Hence, the SAL-tradition has been interested in the qualitative different ways students go about learning and studying for at least 40 years now (1976a; 1976b).

Approaches to learning in higher education

Today's higher education faces the challenge of not only having to teach students a bulk of domain-specific frameworks and disciplinary insights, but also having to foster skills that will enable them to become versatile experts in their own fields and lifelong learners. In order to reach these goals, successful learning and studying in higher education is expected to involve students in deep approaches to learning towards more meaningful and critical learning instead of just repeating knowledge (Asikainen, 2014; Biggs, 2003). Although, several studies have the (at least implicit) assumption that students' approaches to learning develop towards more deep approaches in higher education (Baeten, Kyndt, Struyven, & Dochy, 2010; N. Entwistle & Ramsden, 1983; Prosser & Trigwell, 1999), there seems to be no clear theoretical foundation for this assumption and no clear empirical evidence. The rationale behind this assumption seems to be largely based on the idea that higher education is indeed increasingly required (and working towards) producing work-ready graduates that are prepared to a life as lifelong learner and that, given the scholarly nature of higher education, part of that should be the development of deep approaches to learning (Lake & Boyd, 2015). Despite a general agreement on the basic assumptions, a multitude of conceptual models can be discerned within the SAL-tradition, placing emphasis on different aspects of learning, using a variety of different but related concepts, and using a multitude methods and measures (largely self-report questionnaires) to empirically measure their concepts (Coffield et al, 2004; Richardsson, 2000; Entwistle & McCune, 2004; Vanthournout et al., 2014). The present paper aims to put some piles in what is called by some 'the swamp of student learning in higher education'. Therefore, we will systematically investigate the research on how students' approaches to learning develop during higher education and what conceptual and methodological aspects are related to this development.

In the next paragraphs we will first situate the SAL tradition within the wider literature on students learning in higher education by giving an historical overview of the research from which the

concept of approaches to learning originated and how it has been operationalized and measured longitudinally over the years.

Historical situation

In the 1970s, Marton and Säljö (1976a)) explored 30 educational psychology students' ways of processing a scientific article and the relationship with their learning outcomes. They found two different ways of processing: surface processing meant that students concentrated on memorizing the text as it was presented whereas deep processing indicated that students concentrated on the meaning of the text or the main message of it (Marton & Säljö, 1976a). In a related publication, Marton and Säljö (1976b) found similar differences in study processes but also found that students changed their study processes according to their perception of the requirements. They found that deep processing could be changed to surface processing by asking students the kind of questions that provoke it. In a further elaboration of this study Svensson (1977) found similar qualitative differences in students' study processes: students applying a holistic approach concentrated on the author's intention and tried to understand the meaning of the text and students applying an atomistic approach focused on separate details in the text and tried to memorize it. In addition, Pask (1976) also found similar categories as Marton and Säljö (1976a; 1976b): students applying a *serialistic* strategy in studying focuses on details and concentrates on learning linearly one thing at a time: A student using a *holistic* strategy tries to understand the connections between different matters and to form an understanding of matters as a whole.

Deep and surface approaches to learning

These early studies concentrated on the processing, but moving forward from the early studies of Marton and Säljö, research now generally differentiates between two qualitatively different approaches to learning: surface and deep approach to learning (Biggs, 2003; Marton, 1976; Marton & Säljö, 1984; Prosser & Trigwell, 1999). The deep approach to learning is associated with

students' intentions to understand and to appropriately engage in meaningful learning, focusing on the main themes and principles and using strategies that are appropriate for creating such meaning. The surface approach to learning, on the other hand, refers to students selectively memorizing, based on motives or intentions that are extrinsic to the real purpose of the task, such as fear of failure or keeping out of trouble (Vanthournout, Donche, Gijbels & Van Petegem, 2014). A strategic (Entwistle & Ramsden, 1983) or achieving (Biggs, 1987) approach to learning has been also identified which refers to how students organized their studying or study according to the assessment of the course. Conceptually the achieving/strategic approach differs from the deep and surface approach in that the deep and surface approaches to learning describe ways in which students engage in learning whereas the achieving/strategic approach to learning deals with how students organize their learning (Vanthournout et al., 2014). Some studies on approaches to learning refer to organized studying (Entwistle & Peterson, 2004; N. Entwistle, 2009) by emphasizing the organized effort in studying more and not emphasizing the effect of the assessment (Entwistle & Petersson, 2005; Entwistle, 2009).

Marton & Säljö (1976a; 1976b) initially talked about ways of processing in reading an academic text, which were based on experimental data. That is to say, these studies did not consider the way students generally go about studying as it is assumed sometimes (Richardson, 2015) but the processing strategies of students and how these were affected by their perceptions of the teaching-learning environment. Marton (1976) first introduced the concepts of *deep* and *surface approach* to describe students' approaches to learning in everyday studies. These were built on studies based on both experimental data and student interview data about how students go about learning and studying. When talking about processing Marton and Säljö (1976) did not talk about everyday studies, but the concept *approach* referred to students' everyday studying. Svensson (1977) also used the concept of approach in his studies by referring to students' approaches in general studies. This is also how the concept mostly has been used in the large body of quantitative research.

Measuring approaches to learning

The studies from Marton, Säljö and Svensson in Götenborg discussed above influenced John Biggs in Australia and Noel Entwistle and Paul Ramsden in the UK to explore the way how students usually go about their studying on a more general level. Biggs (1978) developed a Study Process questionnaire (SPQ) which differentiated between three different dimensions of study processes which include students motivation and their study strategies for a specific course: 1) Utilizing 2) Internalizing and 3) Achieving: *Utilizing* describes meeting minimum requirements in studying and using reproduction though rote learning, *internalizing* refers to students' interest in the academic subject and relating ideas with previous knowledge and the *achieving* approach relates to competitive motives in learning and organized studying. Biggs later renamed these approaches to learning to surface, deep and achieving approaches (Biggs, 1987). Furthermore, in Lancaster, Entwistle and Ramsden (1983) also explored students' study processes on a more general level. The purpose of their studies was to explore approaches to learning in larger contexts by developing an instrument, Approaches to Studying Inventory (ASI). The work of Biggs (1978), as well as Marton and Säljö (1976a;1976b) and Pask (1976) influenced the development of ASI and they found similar dimensions of approaches to learning. The deep approach to learning was defined as *intention to understand* and a *critical process in learning* which was measured with two components: relating ideas and use of evidence (Entwistle & Ramsden, 1983).

In the early nineties, Jan Vermunt in the Netherlands developed the learning patterns model in an attempt to provide a more comprehensive and integrated account of learning by bringing together four different learning components: cognitive processing strategies, regulation strategies, conceptions of learning, and orientations to learning (J. D. Vermunt, 1996; J. Vermunt & Vermetten, 2004) We do not elaborate on this model in this contribution, but refer to the paper by Vermunt and Donche in this special issue and the chapter by Vanthournout et al., (2011) for a review on the longitudinal research using this model.

The teaching-learning environment

From the early studies of the Marton and Säljö, it has been evident that the teaching-learning environment has effect on students' approaches to learning. Or to be more precise, students' experiences of the teaching-learning environment are related to students' approaches to learning (Ramsden, 1997). Approaches to learning are thus seen as context-specific (Entwistle & Ramsden, 1983). In general, positive perceptions of the teaching-learning environment have been found to be related to a deep approach to learning and, respectively, negative perceptions to the surface approach to learning (Entwistle & Ramsden, 1983; N. Entwistle, McCune, & Hounsell, 2003; Parpala, Lindblom-Ylänne, Komulainen, Litmanen, & Hirsto, 2010; Richardson, 2005; Richardson, 2006).

Longitudinal perspective

Learning outcomes

The first longitudinal studies concerning on students' learning outcomes were done already in the 1970's, for example, Marton (1975) and Dahlgren (1975) already explored students learning outcomes in a pre-posttest design. Also the seminal studies by Marton and Säljö (1976a;1976b) were longitudinal in nature as well. They found that students' who used deep processing also remembered the main point of the text better after a while (Marton & Säljö, 1976a). On the other hand, the students who used surface processing recalled just some details about the text in both cases even though they were provided with the information or correct answers in the first measurement. In the second study (Marton & Säljö,1976b), students' learning outcomes and *ways of processing* were measured and manipulated by giving them assignments about the text which required surface- or deep-level processing (Marton & Säljö, 1976b). It was shown that the way of processing and the learning outcome could be influenced by the assessment demands. The group that received factual questions, started to concentrate on details and tables more. In addition, some

students in the group that received questions requiring deep-level processing started to concentrate on the main ideas of the text and deep-level processing. However, not all students did so. In conclusion, it was shown that the way of processing can be impacted with changing the demands of the task.

Learning processes

Although the nature of the studies mentioned above was longitudinal, none of these studies measured students' way of processing longitudinally, but measured students' level of learning outcomes. However, Svensson (1977) used a longitudinal design in his study and also reported longitudinal data on students' normal ways of studying. His procedure build on the study by Marton and Säljö (1976b) and included three measurement points over a period of five weeks. He found that students' ways of processing the learning material were rather stable across the occasions. However, there were fewer students applying a holistic approach and some more applying an atomistic approach after the first five weeks. This drop was not so evident in the next procedure concerning the second text.

It seems that interest in the development of students learning took a pause after these studies and concentrated more on developing the questionnaires and large-scale cross-sectional studies. The interest to study the development of general level approaches to learning rose again in the mid 80's. Despite the attention for the contextual nature of the approaches to learning or processing in early studies, a lot of longitudinal studies have been set up from a more general perspective. The first longitudinal quantitative study on the development of approaches to learning was conducted by Watkins and Hattie (1985). They explored 540 Australian tertiary students' approaches to learning in their first year and in their third year using the ASI and found a decline in the deep approach during their studies. Since then, longitudinal studies have been conducted but they seem to have had contradictory results. Some longitudinal studies suggest that the deep approach does not develop

during studies in higher education (Ballantine, Duff, & McCourt Larres, 2008; Lietz & Matthews, 2010; Rodriguez & Cano, 2007; Watkins & Hattie, 1985; Watkins & Ismail, 1994; Zeegers, 2004). Other longitudinal design studies have also found a decline in the deep approach (Lietz & Matthews, 2010; Wilding & Andrews, 2006). In addition, some studies have found a decrease in the surface approach to learning during studies in higher education (Hall et al., 2004; Gordon & Debus, 2002; Rodrigues & Cano, 2007) but also an increase of the surface approach has been reported (Geitz et al 2014; Zeegers 2001). Several longitudinal studies have also suggested that students' approaches to learning are relatively stable during studies (e.g. Edmunds & Richardson, 2009; Zeegers, 2001). However, a recent review of the literature in the field of health sciences indicated that higher education institutes that adopted a curriculum-wide problem-based approach to learning would foster their students to develop more deep approaches to learning (Dolmans, Loyens, Marcq and Gijbels, 2016).

.

The present study

In the present study we will systematically review the longitudinal research on how students' approaches to learning develop during higher education and will take into account conceptual and methodological differences between the studies.

Method

Criteria for inclusion

Before searching the literature, we determined the criteria for inclusion in our analysis. First, each study needed to be empirical, meaning that each paper needs to report on the data collection and analyses of qualitative or quantitative data. Second, the studies needed to include an explicit

reference to student approaches to learning (SAL) as a research model, this could include a reference to either approaches to learning, processing strategies or study strategies. As noted earlier, research using the learning pattern model (Vermunt & Vermetten, 2004) was excluded. Third, the focus is on students in higher education. Forth, longitudinal data about students approaches to learning needed to be collected and reported upon, meaning that the study needed to report about data on student (approaches to) learning with the same students for at least two subsequent measurement moments in time. Finally, the paper needed include at least one research question or explicit results about the development of student (approaches to) learning.

For our literature search, the following databases were included: ERIC (Ebsco) PsycINFO, and the Web of Science (WoS). The search terms were "Students" and "learning approach," "studying approach," "learning strategies," "learning orientations," and "learning patterns", and "processing strategies". Given the focus of this review study, these search terms were combined with the search terms "longitudinal" or "changes over time". The search was further delimited to students participating in higher education or post-secondary education. Only peer-reviewed publications published in English language were selected. This initial search retrieved 543 articles from the ERIC database, 44 articles from PsychINFO and 653 from the Web of Science. After screening the abstracts of these papers against the criteria for inclusion, 62 studies were taken to the next step in the selection procedure. In this next step, the authors read the full papers and evaluated these papers again in the light of the five criteria for inclusion described above. One study was excluded at this stage because the full paper could not be retrieved through the subscriptions held by the authors' institutions or through requests from the authors by email and Research Gate. The systematic search provided us with 26 papers that met the criteria for inclusion. In addition to this systematic search, potential studies were also sought after by going through the references of the selected papers (snowballing). We also contacted several researchers active in the field of approaches to learning and asked them to provide relevant studies or to identify additional sources of studies. This

additional search yielded 17 extra studies. Altogether 43 studies were included for a systematic analysis.

In a final step, following the guidelines of Aveyaerd (2014) the papers were re-read and the relevant information form the papers was summarized in a review table (see Table 1). The following methodological and conceptual aspects of each paper were included in the table and will be discussed in the result section of this paper where relevant: number of measurements, level of education, time of measurement, field of study/subject, number of participants, number of courses, duration of the study, measurement instrument, theoretical framework, analysis, what was measured and significant results/changes.

Results

Before we report how students' deep and surface approaches to learning develop during higher education based on the results of our analyses of the review table in Table 1, we give an overview of the general characteristics of the studies included in our study.

A total of 43 studies met our criteria for inclusion. These studies took very different research perspectives. From those studies, three were qualitative, three were mixed method studies and the remaining 37 were quantitative studies. The duration also varied very much: 24 of the studies explored the development of the approaches to learning over a period of one year or more and 19 of the studies less than one year. A total of six studies explored the change in approaches to learning at course level. In 25 of the studies, there were only two measurement points, 12 of the studies included three measurement occasions, three of the studies had four and three had five measurement points. Also the analysis methods varied across studies. Richardson (2013) talks about traditional or old ways of analyzing longitudinal data on student learning as methods with just a single dependent variable and encourages more advanced ways of exploring student learning longitudinally by including several dependent variables. By using this division, 27 of the quantitative used fairly

simple methods with one dependent variable to measure the change (RM Anova) and 10 of the studies used more advanced methods (Manova, Path analysis, Latent Profile Transition Analysis, multivariate growth model, Structural Equation Modelling, Hierarchical linear modelling).

Development of the deep approach

In 18 of the studies a positive development of the deep approach was reported (Asikainen, Parpala, Lindblom-Ylänne, Vanthournout, & Coertjens, 2014; Ballantine et al., 2008; Biggs & Rihn, 1984; Fryer, in press; Geitz, Brinke, & Kirschner, 2016; Gordon & Debus, 2002; Hall, Ramsay, & Raven, 2004; Jackling, 2005; López, Esteban, Matero, Irene, & Rodríguez, 2015; Matthews, 2004; Muis & Duffy, 2013; Phan, 2011a; Phan, 2013; Svensson, 1977; Vu, Van, & Lacombe, 1998; Walker et al., 2010; Wilson & Fowler, 2005; Zhu, Valcke, & Schellens, 2009). In 18 of the studies, no statistically significant changes in the deep approach were found (Chan & Tang, 2006; Chen et al., 2015; De Clercq, Galand, & Frenay, 2013; Edmunds & Richardson, 2009; Geitz et al., 2016; Gijbels, Coertjens, Vanthournout, Struyf, & Van Petegem, 2009; Iputo, 1999; Jackling, 2005; Muis & Duffy, 2013; Nieminen, Lindblom-Ylänne, & Lonka, 2004; Ova, Bloomfield, & Rotem, 2012; Prat-Sala & Redford, 2010; Rodriguez & Cano, 2007; Svensson, 1977; Vu et al., 1998; Wilson & Fowler, 2005; Zeegers, 2001a; Zhu et al., 2009) and 8 studies found a decrease in the deep approach (Cleveland-Innes & Emes, 2005; English, Luckett, & Mladenovic, 2004; Lietz & Matthews, 2010; Prat-Sala & Redford, 2010; Quinnell, May, & Peat, 2012; Volet, Renshaw, & Tietzel, 1994; Watkins & Hattie, 1985; Wilding & Andrews, 2006).

In one study both a decline and stability of the deep approach was reported depending on the level of self-regulation (Prat-Sala & Redford 2010). Vu et al. (1998) and Muis & Duffy (2013) also found an increase of deep processing or critical processing in students belonging to a intervention group but found no increase in a traditional group. Geitz et al. (2016) found an increase of deep learning within Dutch students but no change with German students. Zhu et al. (2009) found no changes in

deep-level learning strategies with Flemish students but an increase with Chinese students. Jackling (2005) found a increase in the deep strategy but not in deep motive. Furthermore, six other studies found different individual differences between different students to many directions (Boulton-Lewis, Marton, Lewis, & Wilss, 2004; Fryer, 2016; Lindblom-Ylanne, Parpala, & Postareff, 2013; Saravanamuthu & Yap, 2014; Ward, 2011) . For example, Fryer (2016) found different development depending on their initial approaches. In the studies by Postareff et al (2014;2015) students were interviewed based on their change in means before and after a course. Phan (2011b) reported that the initial level of deep processing was positively related to the change in surface processing and the initial level of surface processing positively influenced the change in deep processing. A significant negative correlation was also noted between the initial levels of deep processing and the change in deep processing and between the initial levels of surface processing and the change in surface processing (Phan, 2011b) thus indicating that change was more likely with students who initially had lower scores on deep or surface approach than students with high scores.

The 18 studies, which reported a positive development in the deep approach varied a lot according to their methodological choices. These studies included one qualitative study (Svensson 1977) who also found individual differences and the remaining 17 were quantitative studies. The quantitative studies, which found a development of the deep approach to learning, varied a lot from each other. Nine of the studies used Biggs **framework** and his study process questionnaire or the revised version of it (SPQ or R-SPQ-2F) (Geitz et al., 2016; Gordon & Debus 2002, Hall et al 2004, Jackling, 2005; Lopez et al., 2015; Matthews, 2004; Phan 2011b;2013;; Wilson & Fowler, 2005). Biggs & Rihn (1984) used an early version of the SPQ. In three of the studies, Entwistle's framework was used. These studies used his ASSIST, the approaches and study skills inventory for students (Ballatine et al, 2008; Walker et al 2010;) or the Finnish version of his ETLQ, the Experiences of Teaching and Learning Questionnaire (Asikainen et al, 2014). Two studies used

Pintrichs MSLQ, the Motivated Strategies for Learning Questionnaire (Muis & Duffy, 2013; Zu et al.(2009), one study used the Inventory of Learning Processes (ILP) (Vu et al., 1998) and one used the Approaches to Study Questionnaire by Trigwell and Aswin (Fryer, 2016).

The **duration** between the measurements also varied a lot in the studies that reported a development in the deep approach. The measurements varied from one course to three academic years. In six other studies the duration was less than a year: six eight weeks (Geitz et al 2016), 9 weeks (Biggs & Rihn, 1984), 10 weeks (Svensson 1977), 12 weeks (Hall et al, 2004), 13 weeks (Wilson & Fowler, 2005), 15 weeks (Muis & Duffy, 2008) and one semester which was not further elaborated (Zhu et al.2009). In some studies the data collection took a year or longer: three years (Gordon & Debus,2002), two years (Asikainen et al., 2014; Jackling, 2005; Matthews, 2004; Vu et al., 1998), on and a half years (Phan, 2011b), 16 months (Phan 2013) and one year (Ballatine et al., 2008; Fryer, in press; Walker et al., 2010). In the study by Lopez et al., (2015) the duration was not clearly stated as students answered to the questionnaire "when teaching began and when it ended". In addition, the **number of measurements** varied also from two to five measurements. Eight of the studies comprised two measurement (Asikainen et al. 2014; Ballatine et al, 2008; Biggs & Rihn, 1984; Fryer, accepted; Hall et al., 2004; Lopez et al., 2015; Walker et al, 2010;), seven of the studies had three measurements (Geitz et al., 2016; Gordon & Debus, 2002; Jackling, 2005; Phan, 2013; Vu et al, 1998; Wilson & Fowler, 2005; Svensson 1977), one had four (Phan 2011b) and two had five measurement points (Matthews, 2004; Muis & Duffy 2013) measurement points. Thus, little similarities could be found in the studies that found the increase in the deep approach.

Development of the surface approach

When considering the development of the surface approach to learning, 11 studies reported a decrease (Boulton Lewis et al., 2004; Biggs & Rihn, 1984; Chan & Tang 2008; Gordon & Debus,

2002; Hall et al, 2004; Matthews, 2004; Rodriguez & Cano, 2007; Svensson, 1977; Volet et al., 1994; Walker et al., 2010; Zeegers, 2001), nine of the studies found an increase of the surface approach (Ballatine, Duff & Larres, 2008; Fryer, in press; Geitz et al., 2016; Gijbels et al., 2009; Quinnel et al., 2012; Prat-Sala & Redford, 2010; Wilding & Andrews, 2006; Zeegers, 2001; Zhu et al., 2009) and 20 of the studies did not find changes in surface-level learning (Chen et al., 2015; Cleveland-Innes & Emes, 2005; De Clercq et al., 2013; Edmunds & Richardson, 2009; Geitz et al., 2016; Iputo et al., 1999; Jackling, 2005; Leitz & Matthews, 2010; Lopez et al, 2015; Matthews, 2004; Muis & Duffy, 2013; Nieminen et al., 2004; Ova et al., 2012; Prat-Sala & Redford, 2010; Watkins & Hattie,1985; Wilson & Fowler, 2005; Volet et al., 1994; Vu et al., 1998; Zeegers, 2001; Zhu et al, 2009). Zeegers (2001) found both an increase (from measurement one to three) and a decrease (from measurement three to five) of the surface approach. Pratsala & Reford (2010) found that students with high levels of self-efficacy demonstrated no change in surface approach but students with low self-efficacy increased in their surface approach. Zhu et al. (2009) found an increase on rehersal strategies with Chinese but not with Flemish students. Matthews (2004) and Volet et al., (1994) found a decrease on surface motive but no changes on surface strategy. Six of the studies found individual differences in the development of the surface approach (Boutlton Lewis et al 2004; Fryer, in press; Lindblom-Ylänne et al., 2013; Saravanamuthy & Yap 2014; Svensson 1977; Ward 2011). Five of the studies did not measure the change in surface approach (Asikainen et al., 2014; Phan 2011b; 2013; Vu et al, 1998; Postareff et al 2014;2015) and one study did not report the change (English et al., 2004). From the 11 studies that found a decrease in the surface approach, two were qualitative (Boulton Lewis et al., 2004; Svensson 1977) and used three measurement points. The remaining studies were quantitative of which eight studies used the framework of Biggs and one study the framework of Entwistle. Six studies had two measurement occasions, one had three and two had five measurement points.

Discussion

The main aim of this study was to systematically review the longitudinal research on how students' approaches to learning develop during higher education A total of 43 studies were included in our review of which 18 reported an increase in students' deep approach to learning and eight reported a decrease. On the other hand, eleven studies reported a decrease in students' surface approaches to learning and eight reported an increase. The conclusion based on our review could be that the results remain inconclusive as to the question whether students' develop more deep approaches during higher education. More appropriate would probably be to conclude these results give an indication that there is no empirical evidence for the assumption that deep learning should evolve during higher education (Asikainen, 2014). However, the present paper also took conceptual and methodological aspects of the studies included in the review into account in order to explain differences in the results between studies included in the review. These aspects are discussed below.

Conceptual considerations

When considering the conceptual factors, it seems that all and all, the SAL framework lacks a clear conceptual grounding. There are a lot of different frameworks, evident also in this review, which all conceptualize approaches to learning a bit differently. It has been suggested that the usage of different frameworks is one reason for the ambiguity of results concerning the quality of student learning (Dinsmore & Alexander, 2012). Most of the studies included in our review were conducted with Biggs' Study Process Questionnaire, Entwistle's framework was second common. In Biggs' Study Process Questionnaire, measures of the deep approach include measures of both *deep motive* comprising items measuring intrinsic interest and students' personal satisfaction or excitement in studying and *deep strategy* measuring maximizing meaning (Biggs, 1978). However, Entwistle's framework does not necessarily emphasize the deep motive as such but focuses on the students' *intention to understand* the main message of what they are learning together with thinking processes such as *relating ideas* and *use of evidence* which are similar to Pasks (1976) holistic and serialistic strategies (Entwistle et al., 2003). Thus, the focus is quite different between these frameworks.

However, in our review, consistency even between studies using the same framework was not evident. Most of the studies were conducted with Biggs' framework and no consistency was found within the results of those studies or for the studies using the framework of Entwistle. These results suggest that conceptual agreement is not enough to get consistent results about the development in the SAL framework.

A very likely responsible factor for the inconsistency of the results can be found in the different contexts of the studies. Across the different frameworks, most studies measured the development of the approaches to learning at a very general level. Only five studies measured the change within one specific course. None of the studies explored the development of students approaches to learning at a the level of a specific task although Svensson (1977) measured students' general way of studying between tasks. Most studies (N=24) measured change in students' approaches to learning over a timespan of one year or longer. There has clearly been a conceptual shift from the earlier studies and a somewhat surprising development from studying and measuring task- and context-specific ways to learn that are referred as the starting point of the SAL framework (Marton & Säljö 1976a;1976b) towards studying and measuring more general orientations to learning in longitudinal studies. Although there has been some evidence that approaches to learning can be relatively stable across time and courses (N. J. Entwistle, 1991; Zeegers, 2001b), there is the fundamental assumption that approaches to learning are context-specific (Biggs, 2003; N. Entwistle & Ramsden, 1983; Marton & Säljö, 1984) and can vary across different courses (Nieminen et al., 2004). Hence, when exploring the general development, there is the problem that although the domain of the study remains largely stable, the contextual variables of the course (the topic, the course design, the assessment, the teacher,...) probably vary to a great extent on the different measurement moments. Thus, the effect of the teaching learning environment is not taken into account so much despite of the largely accepted theoretical assumption in the SAL tradition that students' approaches to learning are not stable but change as a result of the interaction between the contextual aspects of the

learning environment and the characteristics of the learners (Biggs, 2003; Entwistle & Ramsden, 1983). It can be, thus, concluded that the way most studies measure change in students' approaches to learning is not in line with the theoretical assumptions of the theory under study. Furthermore, studies did not necessarily make it explicit whether the intention was to measure approaches to learning at a general level or in a particular course or context. For example, the SPQ (Biggs 1987) was used in most of the studies, although this instrument has been developed to measure approaches to learning at course level. Some studies did not bring this issue into question and used the questionnaire to measure the development of approaches to learning at a general level within a long time period, across different courses. Possible modifications of the questionnaire was not discussed. This raises also a question about the validity of these studies and brings us to the issue of methodological considerations.

Methodological considerations

The unclarity of our results could also be due to methodological factors. One factor that could impact the diversity of the results is that the analyses to grasp change varied a lot between the different studies in our review. Of the quantitative studies, most (n=27) used fairly old methods to measure the change (repeated measures methods or Anovas) and only a minority (10) of the studies used newer or more advanced methods such as Latent Profile Analyses. In addition, most of the quantitative studies included only two measurement points. With only two measurement points the development or growth is hard to estimate and more sophisticated analysing methods that could be more appropriate to capture longitudinal development, such as multi-indicator latent growth analyses, cannot be used. Nevertheless, most of the studies exploring the development have only two measurement points. One of the potential problems with two measurement points is the phenomenon of regression to the mean (RTM) which means that the scores in the second

measurements tend to be closer to the mean (Barnett, Van der Pols, & Dobson, 2005). One reason for the high number of studies measuring the change with only two measurement points could be problem of subject attrition, meaning that the participants of the study do not participate in the next measurement but leave the study (Gustavson, Soest, Karevold & Røysamb, 2012). The risk of subject attrition grows with more measurement and researchers may be more tempted to use less measurement points although three or more measurement moments would be more appropriate when the aim is to measure development (Singer & Willet, 2003). To conclude, one might question whether the studies with only two measurement points are really longitudinal in nature as methodologically studies which use only two measurement points, may not necessarily grasp the development. However, the studies in our review which used more advanced measures and had more than two measurement points did not have consistent results either which could support the claim that deep approach to learning develops during studies.

Related to the kind of analyses that are reported in the studies included in the review, the idea that lack of changes in the whole group level does not necessarily mean that there are no changes in subgroups of students with similar characteristics, could be raised. Earlier, Haggis (2003) argued that it could be almost impossible to promote the development of the deep approach to learning if it is not already present. On the other hand, McCune and Entwistle (2011) have suggested that some students have a disposition to understand for themselves, which means that these students deep approach to learning is high and stable across courses because they have a strong willingness to fully understand the subject they study. According to Richardson (2011), the stability of students' approaches to learning over time could indicate also the similarities in their learning environments during studies. All and all, this suggestion that the deep approach to learning is more stable than the surface approach was not confirmed by the results of our review. On the contrary, more studies found no significant changes in the surface approach than in the deep approach.

Recent studies (e.g. Fryer, 2016; Vanthournout et al., 2013) making use of person-oriented rather than variable oriented longitudinal analyses indicate that the stability or volatility of students approaches to learning are associated to the initial subgroup they belong to, suggesting that students who start higher education with a deep approach are more likely to continue to do so. Studies that explored the development from a person-oriented perspective in our review had all somewhat consistent results: different subgroups of students develop differently and the development is individual in nature (Fryer, 2016; Lindblom-Ylänne et al 2013; Postareff et al 2014; 2015, Svensson 1977; Saravanamuthy & Yap, 2014). Thus, the individual nature of the development of approaches to learning seems to be supported by our review. This could also be a reason why the group level studies have such different results. The early study of Marton & Säljö (1976b) and Svensson (1977) already showed that different students level of learning outcomes develop differently and different students respond to the requirements of the context more easily than others. It can be suggested that in order to study the development in approaches to learning, it would be worthwhile to look further than the whole group level as indicated by some studies in our review.

A final methodological issue that is clear from this review is that all 43 studies are based on self-report data and that one way to advance the field would be to invest in other measures or at least to invest in data-triangulation (cfr. Catrysse et al., 2016) which has been more common in the Self Regulated Learning (SRL) tradition in the past years (see the paper by Zushi in this special issue).

The limitations of this review could be at least partly responsible for these results. The literature search was limited to three databases: the ERIC database, PsychINFO and the Web of Science and although relevant researchers in the field were consulted and snowballing was used to find relevant papers, widening the scope of databases and/or search terms might have had an impact on the conclusions.

To conclude, we argue that longitudinal research within the SAL-tradition aiming to investigate how students change their approaches to leaning at group level during higher education is a 'dead

end'. First of all there is no clear theoretical or empirical foundation in the SAL tradition for the claim that students' should develop more deep or less surface approaches to learning across different years. In addition, there is unclarity and lack of consistency within the SAL-framework and, thus, more work should be done to identify the differences and similarities in different traditions. Second, ignoring the contextual differences between different courses and study years within higher education is in contradiction with the theoretical assumption that students' approaches to learning change as a result of individual characteristics of the learner and the characteristics of the context. Meaningfully investigating how students change their approach to learning in higher education from the SAL perspective should therefore always take this context into account and is therefore probably best done at course/topic or even the task level. From this perspective it seems to be a bit ironic that other theoretical models, such as the model of domain learning (Alexander, 2003) are theoretically probably better elaborated compared to the SAL-model in order to explain both stability and variability in deep and surface learning within a particular domain over a period of time. For longitudinal research across different courses, domains and study years, other and more general models such as the learning pattern model (Vermunt & Vermetten, 2004) are probably much more appropriate (see also the paper on the learning pattern model in this special issue). These issues brings us back to the metaphor of the swamp of student (approaches) to learning in higher education and our failed attempt with this review to put some piles in it. We might be able to reach the other side of the swamp and make some progress in our understanding of student learning and its development in higher education by going back to the roots of the concept of approaches to learning on the one hand through studying approaches to learning within a clear context (e.g. at course or task level) and relying on other theoretical models on the other hand to strengthen the theoretical foundations (e.g. on the model of domain learning for the assumption that students develop more deep approaches with increasing domain expertise) and to explore new ways to measure students approaches to learning (e.g. inspired by research in SRL).