Task demands and test expectations. Theory and empirical research on students' preparation for a teacher-made test

Broekkamp, H.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
DISCUSSION

Overview
A theoretical and empirical review of the literature indicated that discrepancies occur between teachers' intended task demands, students' test expectations and actual demands. However, to what degree these discrepancies generally occur remained unclear. For this reason, we conducted a new empirical investigation, consisting of four related studies. In this discussion chapter, we discuss the findings of the four studies in an integrated way. In addition, we discuss implications for future research and educational practice.

1 SUMMARY AND DISCUSSION OF THE RESEARCH FINDINGS

In our investigation, 22 history teachers and their 451 eleventh-grade students rated the relative importance of 26 sections of a particular textbook chapter on which these teachers would be giving a teacher-made test. Furthermore, participants rated the degree to which four types of questions (verbatim, paraphrase, inference and skill items), in principle, were to be expected in the upcoming test. Both kinds of ratings were performed on a four-point scale. In addition to ratings that indicated perceived task demands, we collected the teacher-made tests. We determined the occurrence and weight of both the sections (content demands) and the question types (processing demands) in these tests. Finally, we collected test grades as an indication of students' test performance. The relations among variables were analyzed by means of several quantitative methods, including multilevel modeling.

Our discussion of the findings is organized according to the five research questions (see Introduction), which were formulated in relation to our main research question: *Do secondary school students who prepare for a classroom test have a clear (i.e., accurate) perception of task demands?* For each of the five questions, a separate summary of the results concerning content and processing demands precedes a general discussion of the findings. After we have discussed these five research questions, we address the main question.

1.1 To what degree do teachers, who give a teacher-made test about the same learning materials, vary in their intended task demands regarding students' test preparation?

1.1.1 Content demands
Teachers, on average, deemed text sections important; they indicated that students were to expect test questions about most of the sections. However, all of the teachers
made distinctions between more and less important sections; in a minority of cases, sections were deemed very important or unimportant. Across teachers, some sections received higher ratings than other sections. The degree of agreement among teachers varied across sections, but was generally limited. That is, teachers showed considerable variation for each of the sections.

1.1.2 Processing demands

Although most teachers indicated that students were to expect all four types of questions, some types were considered more important in that students were to expect quite a lot or a great many questions of these types. Across teachers, the average ratings differed for the four question types. Whereas inference questions were considered most important, verbatim questions were considered least important. Nonetheless, teachers showed considerable variation for each of the question types.

1.1.3 Discussion

Even though their teaching and testing pertained to the same textbook chapter, teachers showed large differences among each other in their intended task demands. For researchers, this finding implies that in determining the instructional importance of text elements or processing demands, they cannot rely only on reading materials. Instead, researchers have to consider teachers’ views that relate to these materials (cf. Alexander & Jetton, 1996).

Definition of task demands is part of teachers’ instructional planning and decision making (cf. Calderhead, 1996; Clark & Peterson, 1986; Shavelson & Stern, 1981; Stark, 2000). As such it is influenced by multiple and overlapping contexts (i.e., background factors) including: (a) the school subject and its academic and societal context, (b) the organizational context of the department, the school and the state, (c) the curricular materials, (d) teacher characteristics, such as knowledge, beliefs and teaching styles, and (e) the classroom context, including student characteristics and teacher-student interactions (cf. Grossman & Stodolsky, 1994; Marker & Mehringer, 1992).

To the degree that contextual characteristics are shared, they will contribute to the agreement among teachers in their task demands. For instance, the relatively high importance that was assigned to inference and skill questions could possibly follow from a general societal and educational movement towards demands for information to be applied rather than memorized (see Bloom, 1956; Doyle, 1983; Shuell, 1996). Concerning history education in the Netherlands, there has been an increasing emphasis on the application of information, especially during the last decade. Recently it is stressed – and translated into state curriculum requirements – that procedural knowledge should be acquired in tandem with declarative knowledge to achieve the primary goal of historical awareness (De Rooij, 2001).

The variation among teachers’ task demands seems to be best explained by the characteristics of teachers and their direct classroom context (Stark, 2000). In Jetton and Alexander’s case study (1997), differences between three science teachers regard-
ing their assignment of instructional importance were explained by the different profiles of their background knowledge. Similar findings were obtained by Wilson and Wineburg (1988, 1993). In their case studies of history teachers, they distinguished several types of knowledge and beliefs, including beliefs about historical knowledge and pedagogical content knowledge (knowledge about how to transform subject content for instructional purposes). Individual differences regarding these knowledge types were shown to be related to lesson planning, teaching style and testing criteria.

Considering the influence of the classroom context, Fleming, Ross, Tollefson, and Green (1998) found that high school teachers reported using higher cognitive levels of test questions for classes with students who were more interested and had higher ability than for classes that showed lower levels of interest and ability. Although our classes were relatively homogenous regarding academic ability, it is likely that they differed according to the interest that students showed and the problems that they experienced in relation to the learning content. Depending on the inclination of teachers to adapt their teaching and testing to the classroom context, such student characteristics may have influenced teachers' definition of task and test demands.

Like the researcher who examines instructional importance, it appears that students in secondary history classes cannot rely solely on the textbook when preparing for a test. The large differences among teachers seem to make it necessary that they tune into the particular task demands that are intended by their teachers. However, such attuning seems to make sense only when teachers' intended task demands are consistent with the test demands.

1.2 To what degree do teacher-made tests correspond with teachers' intended task demands?

1.2.1 Content demands

On average, less than half of the sections that teachers considered important were included in the test. However, for those sections included in the test, the vast majority was considered important by teachers. Moreover, when tests included sections that teachers considered relatively unimportant, these sections had a relatively low weight in the test. In interpreting this general picture, we should take into account that tests varied widely in the number of sections that they covered. Similarly, the number of sections in the test that were rated unimportant, varied from teacher to teacher.

1.2.2 Processing demands

Although teachers' tests showed variation in the weight that the four question types had, tests generally included each of the four question types. Consistent with the average importance ratings, inference questions received most weight. However, tests also gave considerable weight to verbatim questions, which is inconsistent with teachers' importance ratings. Moreover, only a weak correspondence could be dem-
onstrated between teachers’ individual ratings for the four question types and the weight that these questions had in their tests.

1.2.3 Discussion

The findings underline the importance of making a distinction between teachers’ intended task demands and their actual test questions. The differences between teachers’ intended content demands and actual test content especially concerned a low content coverage of the tests. This low content coverage could follow from practical considerations. Teachers, for instance, may want to limit the time that students spend on taking tests (cf. Dochy & Moerkerke, 1997, p. 416). Especially, when higher level test questions ask students to produce new information, a lesson hour for a test falls easily short in addressing all the important topics. Nevertheless, a minimum level of content coverage is needed to obtain valid test scores (cf. Airasian, 1996; Brookhart, 1999; Dochy & Moerkerke, 1997; Dolmans, Gijselaers, & Schmidt, 1993).

The considerable discrepancy between teachers’ intended processing demands and the cognitive level of actual test questions can be explained in at least three ways: (a) teachers experience difficulties in constructing test items that are representative of their intended task demands, (b) teachers have unclear intentions regarding task demands, and (c) the discrepancy may reflect the difference between the formal and informal curricula.

Concerning (a), teachers may experience problems with constructing representative tests items because they do not know how to construct a good and large set of questions that require more than memorization of information. Daniel and King (1998) and Bol, Stephenson, and O’Connell (1998) found that teachers varied widely regarding their testing skills. Moreover, the study of Bol et al. showed that with increasing skills, teachers’ reported using test types that required higher levels of cognition.

Concerning (b), the importance rating task asked teachers to make explicit their task demands. The question is whether teachers had such an explicit representation of task demands when teaching or when making test items. As indicated by teacher planning research, teachers’ decision making often takes place unconsciously. Moreover, instructional planning is a dynamic process that may occur both prior to instructional episodes and during those episodes (Calderhead, 1996; Clark & Peterson, 1986). In principle, teachers use three methods to align task and test demands. The first method requires that the test is available at the start of the instructional episode and that the teachers attune their instruction to this test. With the second method, teachers make the test at the end of the instructional episode and try to match the test with what is taught (and how). The third method is a combination of the first two methods. We did not systematically collect or analyze information about these three methods. However, at least the first two methods, “teaching to the test” and “testing what is taught”, seem to have been used. Although we asked teachers to bring their test to the research session, four teachers had not made their test yet. On the other hand, some teachers indicated that they chose for a test that
they had used in previous years. For each of the three methods, a limited awareness of task demands may explain discrepancies between importance ratings and actual test questions.

Concerning (c), the importance rating task instructed teachers to express the likelihood that contents and processing activities would appear in the test as well as the corresponding weight when tested. Although, with this instruction, we encouraged teachers to report their intended demands in a realistic way, we should take into account the possibility that teachers reported desired demands instead of the actual demands that they imposed upon students. For instance, Doyle (1983) described the situation in which teachers want to encourage higher level processing by their students but eventually choose for reproductive tests because of the pressure that students exert on the teacher to make the test not too difficult. Whereas, in this situation, generative processing is formally required, teachers, and possibly also the students, know that the informal demands are reproductive. Such informal demands are part of the hidden curriculum, as described by Snyder (1971).

The distinction between the formal and the informal (i.e., hidden) curricula (cf. Cuban, 1992; Doyle, 1992; Erickson & Shultz, 1992; Gehrke, Knapp, & Sirotnik, 1992; Venezky, 1992) evokes the question of whether teachers’ intended task demands were more representative of the task demands imposed upon students than the actual test demands. To get more insight into this question, we have to examine students’ perceptions of task demands as well.

1.3 To what degree do individual students’ test expectations correspond with their teachers’ intended task demands?

1.3.1 Content demands

Like their teachers, students, on average, considered text sections rather important, indicating that they expected test questions about most of the sections. In addition, average section ratings showed a quite similar pattern for students and teachers. At the individual level, however, the degree of correspondence among the participants was rather limited. Although students within classes were clearly more similar than students between classes, within classes, student ratings, generally, showed considerable deviation from the average class rating. This large variation among students contributed to a quite low average correspondence between ratings of individual students and those of their teachers, with some teachers showing a higher average correspondence with their individual students than other teachers.

1.3.2 Processing demands

Across classes and students, students’ average importance ratings indicated that students expected all four question types. Furthermore, students indicated that they expected inference questions to receive the most weight in the test and verbatim questions the least weight. Compared with the teachers, the only inconsistency at the group level was that students assigned more importance to paraphrase questions than their teachers did. At the individual level, considerable variation across students
within classes was found. We also showed systematic variation between classes. Still, only a low correspondence could be demonstrated between individual students’ ratings for the four question types and their teachers’ ratings for these question types. Furthermore, we did not observe any systematic differences between classes in this respect.

1.3.3 Discussion

The wide variation in teachers’ intended task demands underlined the necessity for students to attune their perceptions of task demands to their own teachers. However, the average degree of correspondence between teachers’ and students’ ratings suggests that individual students generally did not have an accurate perception of their teachers’ intended task demands. Nonetheless, concerning content demands, the differences we showed between classes, suggest that some classes had a more accurate perception of the teacher’s content demands than other classes. Similarly, within classes, some students seem to have had a clearer perception of the teacher’s content demands than other students. This latter conclusion is strengthened by the finding that students who had a higher correspondence with their teacher showed more spreading in their ratings. This suggests that these students were more certain about task demands and therefore dared to make sharper distinctions concerning the relative importance of sections.

A clear perception of task demands depends on both the task information that is available and the way students make use of this information. Concerning the available task information, our findings suggest that some teachers conveyed task demands more effectively than other teachers did. In the case study by Jetton and Alexander (1997), such differences were attributed to differences in teachers’ content knowledge and pedagogical strategies to accentuate gradations in instructional importance. However, as Thomas and Rohwer (1993a) made clear, teachers may also vary in their willingness to convey test-related task information to their students. To avoid giving away the test questions, some teachers may refrain from giving this type of task information altogether. Moreover, we should take into account the descriptive nature of our study; systematic differences between students, for instance, regarding their academic skills, could have contributed to class-level differences in the correspondence between individual students and their teacher.

In order to effectively ascertain task demands, students must be able to perform at least three strategies: (a) infer the need to adapt studying to task demands, (b) ascertain task information, and (c) discern relations between task information elements and infer more and less important task demands (cf. Butler, 1998a; Thomas & Rohwer, 1986). Assuming that some students in our study were more effective in ascertaining the teacher’s task demands, such differences will have followed from the degree to which students used these three strategies and used them effectively. Various student factors could have influenced the use of these strategies. An important factor is students’ experiences with previous study-test episodes and the metacognitive knowledge they derived from these experiences (see Winne & Hadwin, 1998). Another factor is students’ understanding and overview of the learning con-
tent, which seems helpful, if not necessary, to judge the value and the structure of task information elements. It is likely that students in our study have shown wide variation regarding this variable. For instance, some students will have started to read the chapter only the night before the test while some others will have kept track with regular reading assignments. Apart from individual differences, the understanding and overview of the learning content could have explained why we found class-level differences in the correspondence between individual students and their teacher. Whereas in some classes the rating task was performed two days before the actual classroom test, in other classes it was performed in the last lesson preceding the test. However, closer inspection of the data revealed no clear effects of the time interval between the rating task and the test.

1.4 To what degree do students’ test expectations correspond with the actual test demands?

1.4.1 Content demands

Students’ importance ratings did show a correspondence with the actual test content, but the degree of this correspondence was lower than the correspondence between teachers’ ratings and test content. That is, students more often rated sections in the test as unimportant and, when this was the case, these sections did not have a lower weight in the test. On the other hand, for sections that teachers considered important, students more often gave high importance ratings when these sections appeared in the test than when they did not appear in the test. The same was true for sections that teachers considered unimportant. Again, this is the general picture. The correspondence between students’ ratings and test content varied across both students and classes.

1.4.2 Processing demands

Averaged over students and classes, students’ expectation that each of the question types would appear on the test was right. However, their expectations about the relative weight of these question types were not entirely fulfilled. Whereas they overestimated the relative importance of paraphrase items, they underestimated the relative importance of verbatim items. Because ratings varied across students (within and between classes) and test demands varied across classes, we compared the correspondence between individual ratings and the test that their teacher gave. This comparison yielded no indications of a relationship between these two variables.

1.4.3 Discussion

As pointed out in previous chapters, students may base their test expectations not only on information that their teachers provide during the study episode, but also on other sources, such as the task demands that are suggested by the learning materials or their experiences with previous tests. These other sources would make it possible that students’ test expectations were better aligned with the test demands than with
the teachers’ intended demands. The actual test, then, could shed light on discrepancies between students’ test expectations and their teachers’ intended task demands by showing that students compensate for differences between teachers’ intended task demands and the test demands (cf. Smith & Feathers, 1983a; Jetton & Alexander, 1997).

In our study, however, the actual test falls short in explaining the discrepancies between students’ test expectations and their teachers’ intended task demands. Although students showed compensation for some of the differences between teachers’ intended content demands and actual test content, the correspondence between students’ test expectations and test content was eventually lower than the correspondence between teachers’ intended task demands and test content. Concerning the processing demands, we found no evidence that students compensated for differences between teachers’ intended demands and test demands. The comparison between students’ test expectations and test demands, then, confirms that students generally did not have an accurate perception of the task demands that were formulated in their particular class.

The partly successful compensation for differences between teachers’ intended content demands and test demands suggests that students used information other than that conveyed to them by teachers. However, even if we had found no evidence of compensation effects, it would have been likely that students used various sources of information. Students may get to know teachers’ intended task demands not only by the explicit information that teachers give but also by implicit teacher information and other indirect sources. Similarly, teachers who are effective in communicating their task demands to students do not necessarily convey these demands in a direct and explicit way. For instance, they may also develop a predictable pattern of testing.

Sources of task information can be assumed to differ in their value of signaling content and processing demands. For instance, previous tests generally are more indicative of the relevance of specific processing demands than for the relevance of specific topics, because specific topics will generally be different between tests, whereas the cognitive level of questions could well be similar. Nonetheless, previous tests can signal students about types of content. In a pilot study, we interviewed two secondary school history teachers about their task and test demands regarding a chapter about the Second World War in which several explanations were given for the causes of this war. Interestingly, one teacher stressed economical causes, the other teacher social cultural causes. Both explained their position and indicated that it was grounded in their general beliefs about the origins of major historical events. Their beliefs were also consistent with the chapter test they gave to their students. We did not examine students’ perceptions about content demands in these two classes. However, these perceptions could have been attuned to the fundamental beliefs of their teacher, not only regarding the chapters about the Second World War but also regarding other chapters that described major historical events (cf. Wilson & Wineburg, 1988, 1993).

Study materials as a source of information about the task demands may include not only textbook chapters, but also supplementary materials, including texts, practice questions, VCR-presentations and study handouts (cf. Thomas & Rohwer, 1987;
DISCUSSION

Stodolsky, 1989). We did not systematically collect or analyze these additional sources. However, our impression was that the classes in our study confirmed the conclusion of previous studies on the use of curricular materials (see Gehrke et al., 1992). These studies indicate that the textbook occupies a central role in social studies classrooms. Concerning study handouts, Thomas and Rohwer (1993a, p. 23) acknowledge that, in principle, these study aids can be very informative about task demands, but that they usually provide only the dates of tests, the topics that will be taught per lesson and a time schedule for reading a certain number of pages in preparation for the lessons and eventually the test. The study handouts that two teachers in our study spontaneously gave to us, confirmed this picture. These handouts were part of a study calendar and gave no indications about the relative importance of text units or ways to process these units (cf. Knott, 1987).

Now we have discussed each of the relations between the teachers' task demands, actual test demands and students' test expectations, we have a fair basis to decide what constituted the actual test demands of the classroom study tasks we examined. Concerning content demands, we conclude that teachers' intended demands were generally more representative of the actual demands, compared with the other two variables. They were more consistent with the test demands than were students' test expectations and at the same time more comprehensive than these test demands. However, for some demands, students' expectations were more indicative of the actual demands that teachers' intended content demands. For these demands, students' test expectations showed a compensation for differences between teachers' intended demands and the actual test. Concerning processing demands, the considerable inconsistency between teachers' intended demands and the test demands makes it hard to decide which of the two is more representative of the actual demands that were imposed upon students in the classroom. Moreover, including students' expectations of processing demands in our analysis did not shed light on this issue. Therefore, we tend to conclude that the relative importance of processing demands was generally unclear to both the teachers and the students. This does not mean that we exclude the possibility that some students or teachers did have an accurate perception of processing demands. However, since we did not have a reliable criterion to ascertain what were the more important processing demands in classrooms, we cannot examine whether some participants had a more accurate perception of processing demands than other participants.

1.5 Is a closer correspondence between individual students' test expectations and the teacher's intended task demands related to a higher test performance?

1.5.1 Content demands

The degree of correspondence between individual students' test expectations and their teachers' intended content demands did not show a direct relationship with students' test performance. However, we demonstrated an indirect relationship, indicating that students with higher test grades more consistently showed a relatively close correspondence with the teacher's intended content demands.
1.5.2 Processing demands

As our discussion of question 4 made clear, the teacher's intended processing demands could not be considered a reliable criterion for the actual task demands imposed upon students. Therefore, with regard to processing demands, we could not examine question 5 in a meaningful way.

1.5.3 Discussion

The finding that students with higher grades generally showed more correspondence with their teacher confirms that teachers' intended content demands indeed constituted a sound criterion for the content demands that were actually imposed upon students. Moreover, it confirms that some students had a more accurate perception of these demands than had other students. However, it is uncertain whether students actually used a more accurate perception of teachers' content demands to increase their learning performance. It is indeed possible that students with higher grades did not engage in selective processing, but paid equal attention to important and unimportant content. In this case, their perception of content demands would have been an epiphenomenon and would not have contributed to their superior learning performance (cf. R. C. Anderson, 1982; Reynolds, 1992).

Beforehand, it was clear that we could not demonstrate a causal relation between students' test expectations and test performance. To do this, we would also have had to examine the study process of students. Our results, however, question the existence of a causal relation; a substantial number of students had a relatively clear perception of content demands, but eventually obtained a relatively low test grade. It should be stressed that a relatively clear perception of content demands does not necessarily mean that students had an absolutely clear perception of content demands. It is even possible that content demands were insufficiently clear for any of the students in order to engage in selective processing of the learning content. As is shown in experimental studies (see Reynolds, 1992), students only use a selective attention strategy when they are convinced that some parts of the learning content are unimportant. The finding that students within classes especially showed a low consensus about the content demands that their teachers deemed unimportant, could mean that such certainty simply did not exist.

In chapter 3, three conditions are mentioned that need to be fulfilled in order for students to profit from their perceptions of content demands: (a) content demands should be clear, (b) a limited study time should encourage students to engage in selective processing, and (c) students should possess study skills that allow them to selectively attend to task information that they perceive as relevant. We want to add, now, three other important conditions to this list: (d) studying should be attuned to other demands of the study task as well (e) students should possess test taking skills that allow them to ascertain test demands and make use of their knowledge effectively, and (f) tests should be representative of the task demands.
DISCUSSION

In retrospect, we now realize that to examine the effects of a possible causal relation between students’ expectations of content demands and their test performance, we would have had to take into account each of these six conditions. Moreover, we now know that the probability of demonstrating effects was seriously reduced by the low content coverage of teachers’ tests. That is, only for some of the content demands could an accurate perception have lead to an increased test performance. Still, this situation may reflect the reality that exists in classrooms. Emphasizing studying for important contents does not guarantee a higher test grade. This fact gives reason to consider the feedback that students receive from the selective attention strategy (cf. Butler & Winne, 1995). When students experience positive effects of the selective attention strategy only in some of the study-test episodes, they might prefer to choose a more reliable strategy and study all text units with more or less equal intensity. As long as study time allows them to do this, this could be the best way to obtain a high test grade. In the second study, we tested this hypothesis by examining whether students who, on average, considered sections more important, also obtained higher test grades. Although no significant relationship could be demonstrated, our importance ratings were only a remote measure of the actual attention paid to the text materials. In general, we cannot assume that students selectively attend to task-relevant contents. Similarly, as pointed out in the following section, we cannot assume that students engage in task-appropriate processing either.

2 DO STUDENTS WHO PREPARE FOR A CLASSROOM TEST HAVE A CLEAR PERCEPTION OF TASK DEMANDS?

In our investigation, we examined students’ expectations concerning relative importance of text sections (content demands) and types of question (processing demands) in an upcoming test. According to both researchers and students, it is important that study strategies are attuned towards these kinds of task demand (cf. Anderson & Armbuster, 1984; Miller & Parlett, 1974; Thomas & Rohwer, 1987; Reynolds, 1992; Van Etten, Freebern, & Pressley, 1997). However, we could not demonstrate that students, on average, had accurate perceptions regarding the two types of demands. Although the methodological design of our study may be responsible for this result, our findings are consistent with previous studies that demonstrated, for classes from various grade levels, considerable discrepancies between students’ test expectations, teachers’ intended task demands and actual test demands (cf. Alexander, Jetton, Kulikowich, & Woehler, 1994; Jetton & Alexander, 1997; Schellings & Van Hout-Wolters, 1995; Simpson & Nist, 1997; Van Hout-Wolters, 1990a, 1990b/1997; Smith & Feathers, 1983a). Moreover, our findings are consistent with studies that failed to show relationships between the variation in students’ study strategies and the variation in the test demands of courses (Thomas & Rohwer, 1987; Thomas, Bol, Warkentin, Wilson, Strage, & Rohwer, 1993).

Taken together, both the present and previous findings on students’ test preparation suggest that students, on average, do not effectively attune their task perceptions and study strategies towards the relative importance of content and processing de-
mands. There are at least two related reasons why effective attuning to these kinds of task demands would not take place: (a) students have neither the ability or the inclination to attune their studying to the task demands, and (b) the study environment makes attuning to task demands impossible or discourages this attuning. In discussing our fifth research question, we already elaborated on these two reasons with regard to content demands. Here, we wish to discuss some additional considerations that especially pertain to processing demands.

Researchers that examined study approaches in varying contexts have demonstrated that these approaches are both consistent and variable across contexts. These findings suggest that students strategically adapt their studying to contexts. However, they also suggest that students have stylistic preferences that appear to set limits on their flexible use of strategies (see Vermetten, 1999). Moreover, there is evidence that stylistic preferences are rooted in students’ beliefs about the nature of knowledge and learning (Schommer, 1994; Simpson & Nist, 1997) or even in their personality characteristics (Vermetten, Lodewijks, & Vermunt, 2001). A case study by Simpson and Nist (1997) showed how college students who considered learning as an act of memorizing factual information that their teachers provided them, did not attune their studying to their history teacher who demanded especially higher order thinking activities. Moreover, interviews with these students revealed that they did not even recognize the teacher’s higher order demands. Such rigid belief systems could explain why students across grade levels and courses most often report using nonselective, uniform and reproductive study activities (cf. Thomas & Rohwer, 1987; Thomas et al., 1993).

On the other hand, our students generally did expect higher level test questions. Moreover, their expectations that tests demanded both higher level questions and lower level questions were attuned to the actual tests they received, which contained both higher and lower order question types. The major discrepancy between students’ test expectations and actual test demands was that students often did not have an accurate perception about the relative weights that these question types would have in the test. However, we should consider whether accurate perceptions about the relative weights of the types of questions could have been helpful to the students in their studying.

If students are to expect both questions that require verbatim reproduction about a certain text fragment and questions that require the generation of new information regarding this text fragment, the higher and lower order demands could conflict with each other when preparing for such questions. The higher level of cognition requires that students integrate or transform information when studying. This way of studying could be at the cost of students’ learning of verbatim or surface features of the material (see Elshout-Mohr, Van-Hout-Wolters, & Broekkamp, 1999; Doyle, 1983; Thomas & Rohwer, 1986). Although such a conflict is not necessary (i.e., a student could make sure that both levels of learning take place effectively), it is uncertain to what extent students are generally able or inclined to alternate their study strategies. Given that Thomas and Rohwer’s study (1987) showed that students reported mostly using uniform and reproductive study strategies, it seems that such strategies allow students to answer most test questions: at least, as long as the questions do not require a too deep level of understanding. However, as described by Doyle (1983),
teachers are under pressure from students to make the test not too difficult. Consequently, teachers could anticipate students' use of a uniform reproductive strategy and refrain from asking test questions that require deep levels of understanding.

These considerations indicate that the degree to which students effectively attune to task demands depends on complex interactions between student factors and environmental factors. Moreover, in the usual classroom study-test situation, we think that conditions do not encourage students to attune studying to the relative importance of content and processing demands. This does not mean, however, that we deem attuning to task demands irrelevant to studying. On the contrary, experimental research shows that attuning to task demands can increase the effectiveness and efficiency of students' learning performance in crucial ways (e.g., see Anderson & Ambruster, 1984; Nist & Simpson, 2000). Furthermore, descriptive classroom studies have demonstrated that students do attune their studying to varying contexts (e.g., see Vermetten, 1999). An important research objective is to identify the contextual characteristics towards students actually attune their studying (cf. Butler & Winne, 1995). Moreover, research could identify study conditions under which effective attuning towards the relative importance of content and processing demands does take place. These suggestions run ahead of the following section, in which we discuss related research suggestions.

3 RESEARCH SUGGESTIONS

We have four suggestions for research: (a) research could be directed towards optimizing the methodology by which we obtained our findings, (b) studies could aim to determine the generalizability of our findings by including other types of task demands and other instructional contexts, (c) research could include other variables to examine our research questions, and (d) studies could include additional research questions to obtain a more complete picture of students' attuning to task demands in study-test episodes.

3.1 Optimizing methodology

A limitation of our investigation is that participants' perceptions of task demands were measured by a single instrument, importance ratings. There are certainly other ways to determine perceptions of task demands. For instance, Simpson and Nist (1997) and Smith and Feathers (1983a) interviewed individual students and their teacher about task demands. Jetton and Alexander (1997) asked students to write questions that they expected in an upcoming test and also collected importance ratings. Schellings and Van Hout-Wolters (1994) asked students to underline task-relevant text fragments, take notes or rate text fragments. Although preferably multiple of these methods are used, these methods are not equally suited to determine perceptions of particular demands in particular contexts. For instance, given our research purpose to examine, at the end of a realistic study episode, participants' views about the importance of broad text sections, underlining text or rating text seem less suited, because they would involve a major additional reading task. For
this reason, we chose a task in which students rated summaries of text sections. On the other hand, an interview and a question writing task (i.e., student-made test) just preceding the actual test would have been useful in complementing students’ ratings of text section summaries.

As an exception among the studies on students’ attuning to task demands, our study used multilevel modeling as a technique to determine, for instance, the average correspondence in ratings between students and their teacher, as well as the variance between teachers regarding this correspondence. The number of teachers per school was too low to disentangle school and teacher/class effects. Our study showed that in some schools, teachers of the same department each made their own test, even though the same chapter was taught. Moreover, our findings indicated that teachers of the same school who gave the same test, had quite different intentions regarding the students’ study task. These findings suggest that the differences we found at the class level can be attributed to teachers/classes instead of schools. However, to determine the generalizability of these findings, future studies need to include more teachers per school. Similarly, the number of students that rated the importance of sections – though considerable – was insufficient to examine the interrelations among the sections. Future studies could increase the number of students and attempt to determine the content structure of an instructional text, based on importance ratings that students in various classrooms give to elements of this text.

The research findings that we reported regarding the cognitive level of test questions were based on individual teachers’ classification of their own test questions. Initially, we planned to use researchers’ classification of test questions as a convergent measure. However, gradually it became clear that we did not have sufficient information available to classify the test questions accurately. First, we did not collect information that was provided supplementary to the textbook. Therefore, we could not know whether questions that apparently required higher level processing were in fact reproductive, because crucial information had been provided in the classroom. Second, to know the answers that teachers accepted from their students, we collected teachers’ model answers on paper; these model answers, however, were not informative enough to make distinctions between the four categories of test questions. For instance, the model answers generally did not indicate whether information was to be given in verbatim or paraphrased form. Similarly, it was often unclear whether students were to use a domain specific method (applying steps of skill) or just needed to combine their content knowledge with logical reasoning to answer a question (inference).

Eventually, we decided that our limited sources only allowed us to classify, in a reliable way, questions that required information that was stated verbatim or paraphrased in the textbook chapter or clearly stressed verbatim information that plausibly was provided to students supplementary to the textbook chapter. The identification of this reproductive type of questions was performed by two researchers who showed a reasonable agreement (Cohen’s Kappa = .82, \( p < .001 \)). It appeared that proportions of reproductive questions based on the researchers’ codings were correlated with the aggregated proportions of verbatim and paraphrase items as obtained by individual teachers’ test codings; the Pearson correlation was .59 \( (p = .012) \). We are not sure how to interpret this correlation value. On the one hand, it indicates that
the rank order based on teachers' own test codings of reproductive questions (verbatim and paraphrase items) was quite similar to the rank order based on researchers' identification of reproductive questions. Discrepancies could be explained by the different information sources that were available to teachers and researchers. On the other hand, closer inspection suggests that different information sources are not the only explanation. Expessed in the number of test points, on average, .13 of the test questions was classified by the teacher as an inference or skill item, whereas according to the researcher test codings the question stressed reproduction of information. Although we should acknowledge that the researchers' codings were not perfectly reliable, the proportion of .13 suggests that for part of the questions, teachers did not recognize the verbatim nature of these questions or did not carefully enough consider the relation between information required by test questions and information stated in the text.

Given that teachers' classification of test questions is not optimal, future studies could again let teachers and researchers classify test questions. However, to increase the accuracy of researchers' test codings as a convergent measure, information could be collected that is provided in the classrooms as a supplement to information stated in the textbook. Furthermore, in addition to teachers' model answers, researchers could collect the students' test answers along with the teachers' evaluations to determine the answers that teachers accept. A less laborious possibility is that researchers collaborate with teachers in coding test questions. Hereby, researchers could ensure that teachers use the categories as researchers intend and consider the relation between information required by test questions and information stated in the text or supplementary provided in classrooms.

3.2 Examining the generalizability of findings

3.2.1 Other types of task demands

We focused on students' expectations concerning the relative importance of text sections (each including three or four paragraphs) and questions types (representing various cognitive levels) in an upcoming test. Future studies could examine our research questions also in relation to other kinds of task demands that are involved in test preparation study tasks. In the preceding chapters we mentioned the following.

1) Instructional importance of text elements at lower levels (e.g., the sentence or paragraph level; cf. Jetton & Alexander, 1997; Schellings & Van Hout-Wolters, 1995)
2) Abstraction level of the required information (e.g., facts and details versus broad concepts and principles; cf. Thomas et al., 1993).
3) Required level of effort (e.g., when tests have a greater weight in determining course grade this may imply that students have to put more effort in studying; cf. Brookhart, 1997; Natriello & Dornbusch, 1984).
4) Motivational goals that students are expected to pursue (e.g., students are expected to adopt a mastery goal orientation instead of performance goal orientation; cf. Lemos, 1996; Ryan, Gheen, & Midgley, 1998).
5) Epistemological beliefs and study orientations that are appropriate in academic contexts (e.g., teachers may expect students to consider knowledge acquisition as an active process, rather than the passive intake of information; cf. Simpson & Nist, 1997).

6) Interaction among content and processing demands (e.g., regarding section A only verbatim reproduction of definition X is important, whereas regarding section B, it is important to understand the underlying principles of Y).

To this list, we would like to add five other kinds of task demands:

7) Required level of production (e.g., recognizing versus recalling information; cf. Lundeberg & Fox, 1991).

8) Sources of information (e.g., the relative importance of information given during lectures versus the information stated in the textbook; cf. Van Etten et al., 1997).

9) Difficulty of the learning materials (e.g., more difficult texts imply that students have to show more persistence in their study efforts; cf. Thomas & Rohwer, 1987).

10) Time that students have available for studying (e.g., students can be expected to keep up with regular reading assignments and not to postpone studying until one night before the test; cf. Thomas & Rohwer, 1987).

11) Information load (e.g., an increase of information load could imply that students have to spend more time, be more selective, or be more systematic in their studying; cf. Thomas & Rohwer, 1987).

Note that this list concerns a broader definition of task demands that not only includes the products of a task (1, 2, 6, 7), but also includes requirements concerning the way these products are pursued (3, 4, 5, 9), or the resources available and the constraints when generating the products (8, 9, 10, and 11) (cf. Doyle, 1983, p. 161). Instead of task demands, requirements that concern the process or the resources and constraints of tasks may also be referred to as task conditions. We use task conditions as a broad term to indicate all aspects of the task environment that may influence how students pursue the products of a given task (cf. McConkie, 1977, p. 21; Winne & Hadwin, p. 279). For study episodes, task conditions, for instance, may include the quality of teachers’ lectures (see Wilson, Lizzio, & Simons, 1997) or the classroom climate (see Wubbels & Brekelmans, 1998).

Our list is not exhaustive. An important goal for future studies, however, is to determine the kinds of task demands and task conditions that are actually in focus when students try to self-regulate their learning and increase their learning performance (cf. Butler & Winne, 1995). We assume that only demands and conditions for which students have a clear perception are likely candidates.

In the present investigation, we focused on the accuracy of perceived task demands (i.e., the correspondence that students’ perceptions show with the actual task demands). Future studies could examine two other facets of perceived task demands as well: the level of certainty that students experience regarding the correspondence between perceived and actual task demands, and the comprehensiveness of students’
perception of task demands (i.e., the degree of coverage and the level of detail; cf. Butler, 1998b).

Task demands and task conditions define the nature of test preparation study tasks (cf. Doyle, 1983). Apart from test preparation study tasks, our research questions can also be examined in relation to other kinds of academic tasks. For instance, Braet, Moret, Schoonen, and Sjoer (1993) compared students’ perceptions about the demands of writing tasks that were used in the national examinations at the secondary level in Dutch education; a wide variation among these perceptions was found. Similarly, an experimental study by Broekkamp and Van den Bergh (1996) indicated that first year psychology students who revised a text in a foreign language showed large differences in their perceptions of task demands (cf. Hayes, Flower, Schriver, Stratman, & Carey, 1985). Although probably many more studies of this kind have been conducted, our impression is that there are relatively few studies that have explored students’ task perceptions in relation to classroom demands that particular teachers define. For the future, more studies could examine such perceptions, and may even compare them across different kinds of classroom tasks (cf. Eley, 1992; Green & Weade, 1987; Hadwin, Winne, Stockley, Nesbit, & Woszczyka, 2001; Jetton & Alexander, 1997; Simpson & Nist, 1997; Smith & Feathers, 1983a; Winne & Marx, 1982).

3.2.2 Other contexts

Students’ ability or inclination to attune their perceptions and study strategies to task demands is dependent on both the student and the study environment. Both factors can be assumed to vary widely across different contexts. An important contextual variable seems to be the grade level of students. For instance, the ability to attune studying to task demands increases with grade level (Reynolds, 1992). Moreover, the need to attune studying to task demands increases up the grades, because students have to study a larger body of materials in a given time period and have to choose among an increasing number of possible processing demands (Thomas & Rohwer, 1987). Although students in our sample were in the penultimate year of the highest stream in Dutch secondary education that prepared for university, the actual study environment of university may provide very different conditions for attuning studying to task demands (cf. Thomas et al., 1993b, p. 253). Thomas and Rohwer (1987) examined students’ attuning to task demands across grade levels. Although their quantitative study included both secondary and university courses, the number of university courses was too low to examine, in a reliable way, the relation between students’ strategies and course demands across these courses. Future quantitative studies could include a sufficient number of courses at both the secondary and university level and examine possible differences in the degree to which students’ strategies and their perceptions are attuned to task demands. Moreover, longitudinal studies could follow the development of students’ attuning to task demands within students across grade levels (see also section 3.4).

Other contextual variables have probably influenced our findings. For instance, the low consensus among teachers in our study, may have followed from our choice
of school subject. History as a school subject can be characterized by its relatively long study texts and high levels of autonomy for teachers with respect to defining what they teach and assess (Stodolsky & Grossman, 1995). Although our findings were broadly consistent with previous studies in the domain of biology (e.g., Schellings & Van Hout-Wolters, 1995; Bol & Strage, 1996) and science (e.g., Jetton & Alexander, 1997), future studies could make a direct comparison between different school subjects to reveal specific differences in the way task demands are communicated and translated into test demands.

Although teachers in our study did demand more than reproductive processing of their students, our impression was that in the vast majority of the classes the teaching style was rather traditional, in the sense that it seemed very much centered on the textbook and on the way the teacher interpreted this textbook. Although this teaching style is still very common in content area classrooms (Alvermann & Moore, 1991; Bean, 2000), there is a general educational development going on that emphasizes the need for students to develop self-regulated skills and acquire personal understanding of the subject matter (Shuell, 1996). In the Netherlands, this development has resulted in a nationwide innovation that aims to turn the traditional secondary school for upper secondary classes into “a house of study” (Bolhuis & Voeten, 2001). We collected our data just prior to the implementation of this education policy. Furthermore, we did not attempt to differentiate between teaching styles. However, future studies could compare the role of task demands and perceptions in traditional knowledge transmission oriented classes and classes that are oriented towards independent learning and personal knowledge construction.

### 3.3 Including other variables

In this study, we examined four kinds of variables: (a) students’ perceptions of task demands (i.e., their test expectations), (b) teachers’ intended task demands, (c) actual test demands, and (d) students’ test performance. We assume that this selection of variables allows for basic answers to our research questions. However, future studies that address our research questions could increase precision or certainty of answers when two other crucial variables are included: (e) task information that is provided in the classroom, and (f) students’ study strategies and actual study behavior.

The task information that resides in classrooms (e.g., information provided in teacher-led discussions or information provided by study handouts) could show insight into the task demands that are really imposed upon students. For instance, task information may indicate that only part of the teacher’s intended task demands is available to the students. It seems reasonable to consider only the part that is communicated to the students as the actual task demands. Obviously, a major problem of including the variable task information is that covering all relevant task cues (by systematic observation and by collection of materials) is very laborious (cf. Christopoulos, Rohwer, & Thomas, 1987, p. 322). A less laborious possibility was suggested in study 2, when we examined the “class perception of task demands”. This aggregated measure of students’ perceptions was considered to reflect the salient
task information that resided in the classroom. As such, it was taken as an indicator for the actual task demands that were imposed upon students. A problem, however, is that the classroom perception does not show insight into relevant task information that is available but not used by a large part of the class. This explains why this measure appeared to be less powerful as an indication of the task demands than the teachers' intended task demands. When there is more consensus among students, however, the class perception could constitute a useful indicator of task demands (cf. Ryan et al., p. 533). Moreover, we should not dismiss collection of direct classroom task information when it does not cover all task cues. For instance, examining explicit information (e.g., oral statements by teachers about task demands) may already be illuminating.

Including measurements of students' study strategies and study behavior could indicate which kind of task demands are clear and functional enough for students to base their studying upon. For instance, when most students emphasize a certain topic in their studying, this topic can be considered as instructionally relevant.

There are direct and indirect methods to examine study strategies (see Van Hout-Wolters, 2000). An indirect method, for instance, was used by Jetton and Alexander (1997). They measured students' recall for all paragraphs of a study text to indirectly determine which of those paragraphs received most attention during studying and apparently were more instructionally relevant to the students. Although this method can provide additional insight into students' attention for particular information, it should be taken into account that learning performance does not always reflect deliberate attention processes. For instance, salient information can be well remembered despite little time being spent processing this information (see Reynolds, 1992; Hidi, 1995; Van Hout-Wolters, 1986). Another indirect method is to ask students to report their strategies at the end of the study episode, preceding the test. Again, such a method can be very useful, but may not always reflect actual strategy use (Veenman, submitted). More direct methods to determine study strategies concern on-line assessment of reading times or the collection of think-aloud protocols. Broekkamp, Schellings, and Van Hout-Wolters (1996) collected think-aloud protocols of students who performed experimental study tasks. In addition, they used underlining of task-relevant text parts and retrospective interviews as less direct methods. The three methods converged in that they showed only a few signs of task awareness. The researchers, however, were careful in concluding that students were not aware of task demands. Although the think-aloud protocols showed insight into reading processes, cognition at the level of task goals and perceived task demands may be directly verbalized only in rare cases (cf. Ericsson & Simon, 1984). Another flaw of think aloud methods are their obtrusive effects on the regular instructional-learning process (cf. Smagorinsky, 1998). Such effects may make think-aloud methods unsuitable when investigating studying under fully realistic classroom conditions. Still, as there seem to be no perfect measures, some limitations will have to be accepted. To take into account that measures may give a distorted view of strategy use or study behavior, multi method designs are preferable (cf. Van Hout-Wolters, 2000; Veenman, submitted).
3.4 Examining related questions

Our empirical investigation addressed some important questions concerning students' attuning to task demands in study-test episodes. To gain a more complete picture of this topic, related questions need to be examined. A heuristic framework or model that can generate these questions is presented in Figure 1. The model, which is based on the theoretical review presented in the introduction chapter, describes the study-test episode from the perspectives of both the student and the teacher.

From the student perspective, the model highlights the individual students' definition of the test preparation study task. Students' task definition includes the representation of the task demands and conditions, the setting of study goals and the selection or construction of study strategies for attaining study goals. To define the study task, students may use various information sources that are available during the instructional-learning process, such as study materials, notes made during lectures, and classroom discussions (teacher-student interaction). Moreover, students define tasks in interaction with their prior knowledge (e.g., metacognitive knowledge of previous study-test episodes), and study dispositions (e.g., intelligence, learning style). During the task definition process, students gradually develop a perception of the task demands and conditions (this is part of the knowledge that is acquired during the study-test episode) and a representation of goals and strategies. These goals and strategies may be based on students' perception of task demands and conditions as well as on the students' personal goals.

Students' task definition can be described by the following strategies that students ideally have to employ: (a) infer the need to adapt studying to task demands, conditions and personal goals, (b) ascertain task information, (c) discern relations between task information elements and infer more and less important task demands and conditions, (d) set personal goals, (e) compare and prioritize external requirements and personal goals, and (f) define strategies for attaining study goals (cf. Butler, 1998a; Butler & Winne, 1995; Thomas & Rohwer, 1986; Winne & Hadwin, 1998).

In addition to a depiction of the study episode, the model describes the episode of actual test taking. Again, students engage in a task definition process. That is, they develop a perception of test demands and test conditions, define goals, and construct (or select) strategies to perform the "test taking task". In interaction with other student characteristics (e.g., knowledge acquired during the study episode), students' definition of the test taking task determines the test taking process and, as a result, students' test performance. Eventually, the test taking process along with the evaluated test performance may provide feedback for future study-test episodes. The total study-test episode is framed by contextual characteristics (e.g., the influence of fellow students).
Figure 1. A heuristic model for describing study-test episodes

<table>
<thead>
<tr>
<th>teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>teaching dispositions, prior knowledge, knowledge acquired during the study-test episode</td>
</tr>
<tr>
<td>task definition (acquiring knowledge and defining goals and strategies regarding the demands and conditions of the study task)</td>
</tr>
<tr>
<td>goals and strategies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>study episode</th>
</tr>
</thead>
<tbody>
<tr>
<td>observable products of the instructional-learning process (e.g., notes)</td>
</tr>
<tr>
<td>instructional-learning process (students’ individual studying, teacher-student interaction, teacher’s lesson preparation)</td>
</tr>
<tr>
<td>study materials (e.g., textbook chapter, study handouts, previous tests)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>student</th>
</tr>
</thead>
<tbody>
<tr>
<td>study dispositions, prior knowledge, knowledge acquired during the study-test episode</td>
</tr>
<tr>
<td>task definition (representing task demands and conditions, defining goals and strategies to perform the study or test tasking task)</td>
</tr>
<tr>
<td>goals and strategies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>test episode</th>
</tr>
</thead>
<tbody>
<tr>
<td>test performance</td>
</tr>
<tr>
<td>instructional-learning process (students’ test-taking, teacher-student interaction, teacher’s test construction and evaluation of test performance)</td>
</tr>
<tr>
<td>test materials</td>
</tr>
</tbody>
</table>

Note. The model describes a study-test episode, consisting of a study episode, in which students prepare for a test, and a test episode, in which students actually take the test. These episodes are described from both a teacher’s and a student’s perspective. From the teachers’ perspective, the model highlights teachers’ definition of the classroom study task. From the students’ perspective, the model focuses attention especially on students’ individual definition of the study task, which includes their representation of the classroom task demands and conditions. Although depicted separately, teachers’ and students’ task definition are part of the instructional-learning process, in which both teachers and students fulfill learning functions that are involved in the study-test episode.
CHAPTER 6

From the perspective of teachers, the model focuses especially on individual teachers’ definition of the classroom study task. *Teachers’ task definition* takes place in interaction with contextual variables (e.g., the role of departments and state curriculum requirements), teacher dispositions (e.g., teaching style), prior knowledge (e.g., beliefs about the learning content), the information sources that are available during the study-test episode and the knowledge that is acquired in relation to these information sources (e.g., the teacher learned that students have problems with a particular topic). One of the products of teachers’ task definition is teachers’ intended task demands, which may develop during the test-study episode and can be considered as part of teachers’ goals. To communicate intended task demands, teachers can shape their interactions with the students (e.g., providing them with explicit information about the upcoming test) and provide materials pertaining to the current episode (e.g., study handouts) or previous episodes (e.g., making available previous tests). The model also can be used to describe teachers’ construction (or selection) of the test materials at the start of a study-test episode or during this episode. Finally, the model describes teachers’ evaluation of students’ test performance.

Although depicted separately in the model, teachers’ and students’ task definition can be considered as part of the instructional-learning process associated with the study-test episode. The term instructional-learning process expresses the joint responsibility that teachers and students have for learning in classrooms (Elshout-Mohr et al., 1999). The responsibility for separate “learning functions” will mostly be divided (Shuell, 1996; Vermunt & Verloop, 1999; Van Hout-Wolters, Simons, & Volet, 2000). Although teachers can leave more or less latitude to students, teachers will generally have the responsibility of defining the classroom task demands and conditions. Similarly, they will be responsible for constructing or choosing the test materials, and evaluating students’ test answers. Despite teachers’ important responsibilities in the study-test episode, the eventual tasks that students perform are defined by the goals and strategies that individual students adopt, construct or select themselves.

From the model presented in Figure 1, we can derive the theme that was at the focus of our investigation: the degree to which students have a clear perception of task demands. Keeping our focus on task demands, at least three other themes can be derived from the model: (a) student and contextual factors that determine students’ perception of task demands (b) teacher and contextual factors that determine teachers’ intended task demands and their construction and use of test materials (c) student and contextual factors that influence the way students attune their study and test taking strategies to perceived demands.

Concerning (a), future research could examine the degree to which the accuracy of students’ perception of task demands varies within or across study tasks, task types and contexts. For instance, when longitudinal studies reveal that this accuracy increases during a study-test episode, this could reflect that students gain more overview of the learning content and get to know better what is instructionally important (cf. Alexander, Jetton, & Kulikowich, 1995). Similarly, when this accuracy increases within a course or study year, this could indicate that students get accustomed to their teachers (cf. Van Etten et al. 1997). Furthermore, the rank order within a class regarding students’ accuracy in perceiving task demands could be
DISCUSSION

examined across tasks or even task types. The level of consistency of this rank order indicates the degree to which a clear perception of task demands is a stable or non-stable attribute of students. Moreover, the consistency of students’ accuracy in perceiving task demands could be examined in relation to study style, intelligence, personality traits or other dispositional student characteristics to reveal the degree to which skills that relate to representing task demands are susceptible to change (cf. Veenman & Beishuizen, submitted; Vermetten, Lodewijks & Vermunt, 2001). Training studies by Butler (1998b, 1999) indicate that task representation skills can be effectively instructed, at least to some degree.

Concerning (b), teachers could be compared with regard to three methods of aligning their task demands with test demands: teaching to the test, testing what has been taught and a combination of the two methods. Experimental studies could determine the effect of these three methods on the correspondence between task and test demands. Furthermore, descriptive studies could explore the factors that determine teachers’ choices and effective application regarding these methods.

Concerning (c), research could examine the ways in which students’ attuning to task demands is dependent on the clarity or certainty that students have about task demands, the time pressure that students experience, their motivation, and the various skills they have available for executing task-appropriate and selective study strategies (cf. Reynolds, 1992). Furthermore, research could examine the interplay between personal study goals and demands that are imposed by others. For instance, when studying a text, students may both emphasize both information that they deem interesting or that they do not understand yet and at the same time take into account which information is relevant to the upcoming test (cf. Alexander & Jetton, 1996). However, internal and external study goals may also conflict with each other. For instance, students’ attention for interesting but irrelevant text units (“seductive details”) may decrease their attention for structurally important information (Garner, Gillingham, & White, 1989; Harp & Mayer, 1998; Schraw, 1998). The trade-off between personal and task goals may also be the other way around. For instance, a qualitative study by Entwistle & Entwistle (1991) showed how reproductive examinations distorted students’ efforts to achieve personal understanding of the learning materials. Future studies could extend this kind of research and, for instance, examine under which conditions trade-off effects take place among students’ attention for task- and personal relevant information (cf. Alexander et al., 1994; Jetton & Alexander, 1997; Schellings, Van Hout-Wolters & Vermunt, 1996a, 1996b).

4 SUGGESTIONS FOR EDUCATIONAL PRACTICE

Although tests are most often associated with studying in academic situations, performance tasks in these settings may also refer to, for instance, writing a paper or giving a presentation. Furthermore, outside academic contexts individuals study in preparation for performance tasks and events as well. For instance, “job seekers study the firms of prospective employers in preparation for interviews. Lawyers study rulings in prior cases to prepare written or oral arguments. Advertising agents study market surveys to prepare publicity campaigns. Consumers study product review peri-
odicals in preparation for major purchases. Prospective tourists study tour guides to enhance their travel experiences" (Thomas & Rohwer, 1993b, p. 241).

In all these situations, studying is directed towards performance goals and criteria, which may influence the kind of contents and processing activities the learner emphasizes. Moreover, in many situations, goals, criteria and conditions are imposed by others than the learner. This type of situation requires that learners attune to task demands and conditions, and additionally attune to goals that they have set for themselves. Education could prepare students for these kinds of attuning by encouraging students to develop the necessary skills. One of these skills is students' representation of task demands and conditions.

Students' ability to develop a clear perception of task demands and conditions is both an educational goal in itself and a condition for effective and efficient learning in classrooms. Teachers could provide students with sufficient, clear and valid task information. Furthermore, teachers could ensure that students make use of the information provided, by encouraging them to use and develop skills of attuning studying to task demands. Finally, teachers could encourage students to use and develop skills of attuning to personal study goals as well (cf. Pressley et al., 1997; Ramsden, 1992; Simpson & Nist, 1997, 2000).

4.1 Providing students with task information

Teachers may directly tell students about the demands of the study task or the test. "Giving [students] information about what will be on the test is not the same as giving them the very items of information they will be required to produce on this test" (Thomas, 1988, p. 270). Giving away the test questions would lead students only to prepare for these questions and would therefore "short circuit" their studying (Thomas, 1988, p. 270). More sensible provision of test-related task information is shown in a case study by Simpson and Nist (1997). The teacher told his students, for instance: "An essay is to measure conceptual and organizational skills. I allow 5 points for the introduction. 5 points for the conclusion, and 10 points for the essay... Since I give you some direction and specifics, I expect a lot in your answers - I expect precision" (p. 370, 372). A direct way of conveying content demands, for instance, is teachers' instruction to leave certain text parts aside when studying or read these parts only for global understanding. Next to direct oral provision of test-related task information, teachers may model the strategies that students should perform (e.g., different ways of note-taking, summarizing) and the way these strategies can be shaped to study demands. Consequently, teachers could give feedback to students when applying such strategies. Furthermore, teachers may give students opportunities to interview them about task demands. Finally, information about task demands can be explicitly provided by study handouts. When these handouts are detailed enough and include sample questions, they can serve an important modeling function as well.

In addition to direct methods, teachers may more indirectly give notice to their students about what is instructionally important. First, during the lecture, they can place "important" information on the board, slow down for emphasis and spend
most time on information that has instructional relevance. Second, they can provide students with practice questions or make available test questions and corresponding answers and evaluations of previous study-test episodes. Third, teachers can point students to the kind of information and processing activities that are emphasized by the textbook author (e.g., directions to pay attention to summaries or definitions stated in italics; study the model answers of the textbook practice questions). Fourth, teachers can encourage students to discuss task demands with peers. As is indicated by university students that took part in the interview study of Van Etten et al. (1997), studying with fellow students, one runs less risk of under- or over-studying than when studying alone (p. 205). In using several sources, teachers have to take care that these sources are aligned, and point their students to possible inconsistencies. Moreover, teachers should regularly inform themselves whether their students’ perceptions still match with their own perceptions of the task. Finally, teachers should take care that task and test demands are aligned. Among other reasons, this seems necessary in order that students obtain positive feedback of attuning efforts.

4.2 Encouraging students to attune their studying to external and personal goals

4.2.1 Teaching attuning skills

Several instructional models have been proposed to teach students skills associated with self-regulated learning (see Butler, 1998a). Among these models, the Strategic Content Learning is of special relevance to us, because it assigns an important place to students’ attuning to both task demands and personal goals (Butler, 1998b, 1999). According to the SCL model, self-regulated learning processes are guided and encouraged, instead of directly modeled. Furthermore, students are to develop personalized strategies while working on meaningful school tasks and discussing these tasks with others, the teachers and peers. The discussion with others enables students to make explicit perceptions of task and personal learning goals, strategies to pursue these goals and beliefs underlying goals and strategies. Important objectives of the SCL model are that students’ perceptions and strategies gradually become more focused on task requirements or personal goals, that students become aware when this is not the case and solve inconsistencies by adapting strategies, adjusting goals or seeking additional task information. Although implementation of this model by Butler especially involved global task goals (e.g., keeping up with regular reading assignments), the SCL model seems to be a promising point of departure for teaching students to define and attune to more specific goals of classroom tasks.

4.2.2 Changing task conditions

As we discussed previously, in particular study contexts, students may not be inclined to attune studying to task demands, because task conditions do not encourage them to do so. Consequently, students miss opportunities to develop attuning skills. To encourage students to attune studying to task demands, task information should be sufficient, valid and clear. In addition, more fundamental changes in the conditions of study tasks could be carried through. Such fundamental changes could also
stimulate students' attuning to personal goals. We discuss three kinds of changes: (a) changes that encourage students to use selective attention strategies, (b) changes that encourage students to use task-appropriate processing strategies, and (c) changes that give students more latitude in defining study goals and therefore encourage students' attuning to both external and personal goals.

To stimulate selective processing of information, the number of pages that have to be studied within a certain period could be increased. Furthermore, students could be instructed to select information from different sources (e.g., newspaper articles, various textbooks, internet; cf. Guthrie & Cox, 2001). Finally, selective processing may be encouraged by reducing the number of text study aids. “Texts that abbreviate, consolidate, and signal (e.g., highlighting) important points should prompt minimal levels of autonomous selective processing on the parts of the students” (Thomas & Rohwer, 1993b, p. 15; see also Rijlaarsdam & Couzijn, 1996; Van Hout-Wolters, 1986). These measures could be crucial in order to prepare students for a society that demands its members to select from an increasing amount of available information (cf. Alexander & Jetton, 2000).

To encourage students to use specific processing strategies that are of special value for some study goals, but not for other study goals, tests could clearly emphasize one level of cognitive demands, instead of various levels. For instance, history teachers could give a test that requires students to analyze historical sources with the textbook and other sources available as a support for background knowledge. When this test is presented as an authentic history problem solving task, it can be considered as a form of performance assessment (cf. Greene, 1994; Ruiz-Primo & Shavelson, 1996; Wineburg, 1997). On a different occasion, the same teacher could assess students' global understanding of texts by asking them to reproduce the main points from memory. In yet another test, this teacher could exclusively require students to reproduce relevant facts, dates and names of historical events. Separately testing particular levels of demands, instead of combining them in a single test would take more testing time. However, this need not to be a problem when teachers and students consider teaching and testing as integral parts of their instructional-learning process (cf. Dochy & Moerkerke, 1997). If teachers sufficiently prepare their students for the changing demands of tests, these tests will stimulate students to adopt study strategies that are attuned to the specific processing requirements. Experience in this kind of attuning will lead to students who flexibly adapt their studying to optimize their learning performance.

In traditional study-test situations, the responsibility for defining tasks lies very much in the hands of the teacher. However, we have pointed out that teachers refrain from giving detailed task information, because they do not want to give away the test questions. Moreover, even when they are willing to communicate test-related task information, the demands that teachers have to convey may simply be too large in number or too subtle to pass on to their students. As acknowledged by Ramsden (1992), “numerous decisions about its [the task] requirements have to be faced” (p.160). Ramsden advocates the use of study tasks that decenter teachers' authority. That is, the teacher should transfer some of the responsibility of defining goals to the student. Such transfer of control could decrease the number of task demands that teachers need to communicate. Moreover, when more latitude is given to students in
defining study goals, they will be encouraged to attune studying to personal goals. This is likely to enhance their motivation for studying (cf. Guthrie & Cox, 2001; Wade & Moje, 2000). Finally, when students are given more freedom in defining their personal goals, they may be more willing to accept and internalize classroom goals (cf. Hodgson, 1997; Schraw, Flowerday, & Lehman, 2001).

Decentering teachers’ authority in defining study goals may prepare students better for studying outside schools. “The continual presence of courses that place decisions regarding the content and purposes of studying in the hands of the instructor, may prepare adults inadequately for assuming the responsibility for learning in the later grades and post-school life” (Thomas & Rohwer, 1993, p. 268). Like Ramsden, Thomas and Rohwer advocate the “provision of authentic self-directed learning responsibilities and tasks” (p. 268).

To enable students to develop the necessary task definition skills, the transfer of responsibilities in defining study goals should take place gradually. Moreover, teachers could scaffold students’ task definition processes and gradually decrease their instructional support (cf. Brown & Palinscar, 1989). Eventually, students should be able to independently attune studying to task demands and personal goals in fully authentic study tasks (cf. Guthrie & Cox, 2001; Simpson & Nist, 2000; Wade & Moje, 2000).

The goal to equip students with skills that they can apply independently in authentic study tasks conforms to current educational developments in the Netherlands (Bolhuis & Voeten, 2001) as well as in other countries (Van Hout-Wolters, Simons, & Volet, 2000; Newmann, Marks, & Gamoran, 1996). The stress on independent learning could lead to the belief that task demands become less important to academic learning. We consider this a misconception. Also in classes that emphasize students’ independence, task demands remain crucial, because they are inherent to classroom learning (cf. Doyle, 1983, p. 187; Doyle, 1992, p. 105). The important role of task demands should be acknowledged and shaped in a way that enhances effective, efficient and pleasant studying as well as the development of skills that enable students to attune to relevant learning goals.